Understanding the interplay of lifestyle and deprivation to support policy and intervention development: a mixed methods study

Dr Hamish Michael Elliot Foster
BSc(hons) MBBS MRCS(2010) MRCGP DTMH

Submitted in fulfilment of the requirements for the Degree of
Doctor of Philosophy (PhD)

School of Health and Wellbeing
College of Medical, Veterinary and Life Sciences
University of Glasgow

May 2024

Abstract

Introduction

Unhealthy lifestyle factors (e.g., smoking, high alcohol intake, poor diet) are among the leading risk factors for adverse health outcomes (e.g., mortality). Socioeconomic deprivation is also a risk factor for the same adverse health outcomes. Unhealthy lifestyle factors tend to cluster within individuals and more deprived populations tend to have higher prevalence of multiple unhealthy lifestyle factors.

Much prior research on lifestyle and deprivation has focussed on explaining health inequalities via the prevalence and severity of a few unhealthy lifestyle factors finding that lifestyle can explain a moderate or large amount of inequity, but not all. There has been far less research into whether there is an interaction in the association between combinations of unhealthy lifestyle factors and deprivation. For example, deprivation might influence the association so that deprived populations experience disproportionate rates of adverse health outcomes from combinations of unhealthy lifestyle factors. Further, no prior studies have examined the influence of deprivation on lifestyle associations with adverse health while measuring lifestyle using a wide combination of unhealthy lifestyle factors that includes a measure of social connection (e.g., infrequent social contacts), an example of a ‘new’ or ‘emerging’ lifestyle factor, alongside more traditional factors.

Examining how deprivation influences the association between wide combinations of lifestyle factors (including ‘newer’ lifestyle factors) and adverse health outcomes could reveal higher risk groups that could guide policy and interventions.

Aims:

1) To improve understanding of the association between combinations of unhealthy lifestyle factors, socioeconomic deprivation, and adverse health outcomes.

2) To explore which measures of social connection could be included in a broad measurement of lifestyle i.e., a lifestyle score.

3) To develop a lifestyle score that encapsulates the risks associated with a wide combination of lifestyle factors (including measures of social connection) and socioeconomic deprivation.

4) To understand how combinations of unhealthy lifestyle factors are perceived in the context of socioeconomic deprivation.

Methods

This was a mixed methods study harnessing quantitative and qualitative methodology. There were three main parts or work packages (WP) for this thesis:

WP1 - a systematic review of prospective cohort studies that report on the effect of socioeconomic status (SES) on the association between a combination of lifestyle factors and adverse health outcomes.

WP2 - statistical analyses of UK Biobank, a large prospective cohort of around half a million adults, to a) estimate the associations between measures of social connection and all-cause and cardiovascular (CVD) mortality and b) create a weighted lifestyle score and examine the effects of deprivation, sex, and ethnicity on the association between the lifestyle score and adverse health outcomes.

WP3 - qualitative analysis of key stakeholder perspectives from four focus groups with 25 members of the public and 18 interviews with community, health, and policy professionals.

Results

Systematic review

Six prospective studies were identified that examine the interaction between a combination of lifestyle factors and deprivation for adverse health outcomes. The studies were widely heterogenous in their definitions of unhealthy lifestyle factors, their markers of socioeconomic position, their methodology to assess interactions, and their results. However, there remained a clear suggestion that mortality risks associated with a combination of unhealthy lifestyle factors are seen in more deprived groups, which suggests lifestyle support and policy interventions targeting these populations might usefully be evaluated.

Social connection

Structural (objective) components of social connection, such as the frequency of friends and family visits or living alone had stronger associations with all-cause and CVD mortality than functional (subjective) components, such as feelings of loneliness. However, combinations of both functional and structural components of social connection had the strongest associations highlighting potential benefits that could follow from measuring and supporting both types of social connection. For example, compared with participants with higher levels of both components of social connection (e.g., not living alone and not often lonely) those with lower levels of both components had higher CVD mortality hazard ratios (HRs [95% CI] 1.63 [1.51, 1.76]) than each component alone (functional isolation - 1.17 [1.06, 1.29], structural isolation - 1.27 [1.18, 1.36]).

Weighted lifestyle score and socioeconomic effects

A weighted lifestyle score comprising 11 lifestyle factors (including frequency of friends and family visits and weekly group activity as a measure of social participation) was created using the relative all-cause mortality effects sizes for each individual factor. Smoking had the highest score weighting while social participation contributed a similar weighting to that of more traditional lifestyle factors such as low physical activity and low intake of fruit and vegetables. The

association between the weighted lifestyle score and all-cause and CVD mortality was stronger both among more deprived groups and among men. For example, within each quartile of deprivation and compared with those with the healthiest scores, all-cause mortality HRs (95% CI) for those with the unhealthiest scores were 2.55 (2.35, 2.77), 2.92 (2.70, 3.16), 3.27 (3.02, 3.54), and 3.54 (3.27, 3.82) in the least to most deprived quartiles, respectively.

Stakeholder perspectives

There was wide and detailed appreciation of the socioeconomic barriers to healthy living and perceptions captured here de-emphasised the importance of individual-level responsibility for healthy choices. However, it was felt there was always some level of individual agency or choice and, among professionals, there was a perceived duty to support and encourage healthy change in even the most arduous of socioeconomic circumstances. This appeared to create a tension between perceptions of agency and structure whereby clinical and public health practitioners felt, on one side, a duty to support agency of those with unhealthy lifestyle factors and, on the other side, a deep understanding of the structural forces of the social determinants of health behaviours. Innovative policy and legislation are needed to tackle upstream determinants of numerous unhealthy lifestyle factors simultaneously and across populations.

Conclusion

Deprived populations may experience disproportionate adverse health outcomes from a wide combination of lifestyle factors, including emerging factors such as social participation. Tackling the disproportionate harm associated with combinations of unhealthy lifestyle factors in deprived populations will require numerous levels of innovative intervention and policy. Alongside wider structural and policy change that make healthy lifestyle choices more equitable, lifestyle interventions that incorporate our understanding of the social determinants of lifestyle and that tackle numerous lifestyle factors simultaneously could support

individuals and communities affected by deprivation to avoid multiple unhealthy lifestyle factors.

Table of Contents

List of Tables

List of Figures

List of publications

Publications/Submissions included as part of this thesis:

1. Foster HME, Polz P, Gill JMR et al. The influence of socioeconomic status on the association between unhealthy lifestyle factors and adverse health outcomes: a systematic review [version 2; peer review: 2 approved, 1 approved with reservations]. Wellcome Open Res 2023, 8:55 (doi: 10.12688/wellcomeopenres.18708.2)

2. Foster, H. M.E. , Gill, J. M.R. , Mair, F. S. , Celis-Morales, C. A. , Jani, B. D. , Nicholl, B. I. , Lee, D. and O'Donnell, C. A. (2023) Social connection and mortality in UK Biobank: a prospective cohort analysis. BMC Medicine, 21, 384. (doi: 10.1186/s12916-023-03055-7)

3. (Under peer review at BMC Public Health) Sociodemographic effects on the association between a weighted lifestyle score and mortality in the UK Biobank cohort. Foster, H.M.E., Gill, J.M.R., Mair, F.S., Celis-Morales , C., Jani, B.D., Nicholl, B.I., Lee, D., and O'Donnell, C.A.

4. (Submitted to Health Expectations) Exploring public, practitioner, and policy maker perspectives of health behaviours in the context of socioeconomic deprivation: A qualitative study. Foster, H.M.E., Mair, F.S., Gill, J.M.R., and O'Donnell, C.A.

Related publications not included in this thesis:

1. Foster, H. M.E. et al. (2022) The association between a lifestyle score, socioeconomic status, and COVID-19 outcomes within the UK Biobank cohort. BMC Infectious Diseases, 22, 273. (doi: 10.1186/s12879-022-07132-9)

2. Foster, H. M.E. , Polz, P., Mair, F. S. , Gill, J. M.R. and O'Donnell, C. A. (2021) Understanding the influence of socioeconomic status on the association between combinations of lifestyle factors and adverse health

outcomes: a systematic review protocol. BMJ Open, 11, e042212. (doi: 10.1136/bmjopen-2020-042212) (Appendix 1)

3. Foster, H. M.E. , Ho, F. K., Sattar, N. , Welsh, P. , Pell, J. P. , Gill, J. M.R. , Gray, S. R. and Celis-Morales, C. A. (2020) Understanding how much TV is too much: a non-linear analysis of the association between television viewing time and adverse health outcomes. Mayo Clinic Proceedings, 95(11), pp. 2429-2441. (doi: 10.1016/j.mayocp.2020.04.035)

4. Rey-Lopez, J. P., Ho, F. K.W., Foster, H. M.E. , Petermann-Rocha, F., Sattar, N. , Pell, J. P. , Gill, J. M.R. , Gray, S. R. and Celis-Morales, C. A. (2020) Does the association between physical capability and mortality differ by deprivation? Findings from the UK Biobank population-based cohort study. Journal of Sports Sciences, 38(23), pp. 2732-2739. (doi: 10.1080/02640414.2020.1797438)

5. Foster, H. M.E. , Celis-Morales, C. A. , Nicholl, B. I. , Petermann, F., Pell, J. P. , Gill, J. M.R. , O'Donnell, C. A. and Mair, F. S. (2018) The effect of socioeconomic deprivation on the association between an extended measurement of unhealthy lifestyle factors and health outcomes: a prospective analysis of the UK Biobank cohort. Lancet Public Health, 3(12), e576-e585. (doi: 10.1016/S2468-2667(18)30200-7)

Selected presentations arising from this thesis:

1. Which combinations of health behaviours are associated with highest risk: an exploration of UK Biobank population cohort. Annual Meeting of the North American Primary Care Research Group (NAPCRG). Arizona, USA; November 2022. (Oral presentation)

2. Associations between a weighted health behaviour score, socioeconomic status, and all-cause mortality in UK Biobank cohort. Annual Meeting of the North American Primary Care Research Group (NAPCRG). Arizona, USA; November 2022. (Oral presentation)

3. The challenges of reducing risk and severity of cardiovascular disease in socio-economically deprived communities. Scottish Lipid Forum & SHARP Annual Scientific Meeting. Edinburgh, Scotland; November 2022. (Oral presentation)

4. Associations between a weighted health behaviour score, socioeconomic status, and all-cause mortality in UK Biobank cohort: a prospective analysis. UK Public Health Science Conference. Glasgow, Scotland; November 2022. (Poster presentation)

5. Which combinations of health behaviours are associated with highest risk? An exploration of the UK Biobank population cohort. UK Public Health Science Conference. Glasgow, Scotland; November 2022. (Poster presentation)

6. Which aspects of isolation and loneliness might be most harmful? NADEGS Annual Conference (Scotland's Academic Primary Care conference). Carnoustie, Scotland; January 2023. (Poster presentation)

7. How do the public, health care professionals, and policy makers view unhealthy behaviours in the context of socioeconomic deprivation? A qualitative study. Society for Academic Primary Care (SAPC). Brighton, England. July 2023. (Oral presentation)

8. How do components of social connection interact in their associations with all-cause and CVD mortality? A UK Biobank cohort analysis. Society for Academic Primary Care (SAPC). Brighton, England. July 2023.

9. How do the public, professionals, and policy makers view unhealthy behaviours in the context of socioeconomic deprivation? Annual Meeting of the North American Primary Care Research Group (NAPCRG). San Francisco, USA; November 2023. (Oral presentation)

10. How do components of social connection interact in their associations with all-cause and CVD mortality? Annual Meeting of the North American

Primary Care Research Group (NAPCRG). San Francisco, USA; November 2023. (Oral presentation)

11. What are the effects of sociodemographic variables on the association between a weighted lifestyle score and mortality in the UK Biobank cohort? Society for Academic Primary Care (SAPC). Bristol, England. July 2024. (Accepted for poster presentation)

12. Developing the SSHARE study: Socioeconomically Sensitive Health behaviour And Risk Estimation – a community driven e-health intervention. Society for Academic Primary Care (SAPC). Bristol, England. July 2024. (Accepted for oral presentation)

List of Accompanying Material

Acknowledgements

Supervisors

I would like to thank my supervisors for their support throughout this project. I am extremely grateful to them. Their guidance through the long planning and preparation stages played a critical role in my success in obtaining funding for this project and gaining a prestigious MRC training fellowship. Further, their support during difficult circumstances like the early phases of the COVID-19 pandemic (which coincided exactly with the start of this project) and through challenging family circumstances allowed me to work flexibly and sustain productivity. I would like to offer special thanks to my principal supervisor, Prof Kate O'Donnell, for being understanding and for being a constant source of encouragement and providing near weekly advice and wisdom. Particular thanks are also due to Professor Frances Mair who has also been a frequent source of excellent advice, including invaluable tips on navigating the challenges of combining clinical and academic careers. Professor Jason Gill has been an excellent source of help around lifestyle epidemiology and quantitative analyses. And Professor Duncan Lee has provided invaluable support in overcoming technical and statistical challenges.

Colleagues and collaborators

Numerous administrative staff at University of Glasgow have also made this work possible and I would like to thank them for their assistance: Michere Beaumont, Sharleen Urquhart, Karen Penman, and Alison Purdie-Gore.

I am also very grateful for the critical contributions of colleagues who have collaborated on and facilitated numerous aspects of this and closely related projects: Dr Carlos Celis-Morales, Dr Bhautesh Jani, Dr Barbara Nicholl, Dr Fred Ho, Peter Polz, Dr Alessio Albanese, and Dr Nic Dickson. Thanks to Professor Cindy Gray, Professor Tom Nilsen, Professor Paul Mork, and Professor Susan White who provided helpful feedback as an advisory panel.

I am extremely grateful to all the staff at my clinical place of work (Foster, Logan, and Macdonald Practice, Drumchapel Health Centre) who have supported

me at times through flexible working and encouraged me by being a source of inspiration in their work with patients.

I would like to thank the members of The National Health Service Research Scotland Primary Care Patient and Public Involvement Group as well as Dr Tracy Ibbotson for facilitating the group. This PPI group have provided insightful feedback at various stages of this project. And thanks to Irene Oldfather of The ALLIANCE and Fiona McHardy of The Poverty Alliance who both assisted greatly with recruitment for the qualitative study and for providing feedback for other aspects of this thesis.

Funding

My thanks to the Medical Research Council who provided me with the funding to complete this work (grant number MR/T001585/1). And for their flexibility in permitting me a no-cost extension due to unforeseen circumstances, including the COVID-19 pandemic.

Participants

I would like to thank UK Biobank participants. The UK Biobank research reported here has been conducted using the UK Biobank Resource under Application Number 14151. This work uses data provided by patients and collected by the NHS as part of their care and support and used data assets made available by National Safe Haven as part of the Data and Connectivity National Core Study, led by the Health Data Research UK in partnership with the Office for National Statistics and funded by UK Research and Innovation (research which commenced between 1 October 2020 to 31 March 2021 grant ref MC_PC_20029; 1 April 2021 to 30 September 2022 grant ref MC_PC_20058).

I would also like to thank all those who took part in the focus groups and interviews for the qualitative study, without whom a rewarding and enriching part of this thesis would not have been possible.

Family

I would like to thank my friends and family, without whom my work would not be possible. My in-laws have been invaluable, providing indispensable childcare and babysitting. My parents deserve special mention for their ceaseless support and backing. My children, Innes and Struan, one of whom arrived just before and one during this PhD, have filled our home with laughter and more than made up for their sleep-preventing antics. Lastly, my deepest gratitude goes to Deanne, my wife, whose support and companionship have been cornerstones of my ability to complete this PhD.

Author's Declaration

This thesis was completed during my Medical Research Council funded Clinical Research Training Fellowship. I declare that I am the primary author of the work presented here and I was responsible for leading all aspects of the research with guidance form my supervisors and collaborators. The contributions of colleagues are formally acknowledged below.

This thesis is presented in alternative format where the results chapters are presented as published or submitted. I was first author for each of these papers. I wrote the first draft of each, performed all the analyses, designed the tables and figures, and managed the redrafting of the manuscripts based on comments from co-authors. The contributions of the co-authors for each paper are given below.

Systematic review (Chapter 4)

I was first reviewer for each stage of the review process. Perter Polz, a medical student, assisted as a second reviewer for screening studies and data extraction along with my primary supervisor, Professor Kate O'Donnell. Colleagues Peter Polz, Professor Jason Gill, Dr Carlos Celis-Morales, Professor Frances Mair, and Professor Kate O'Donnell provided feedback on drafts of the manuscript.

Quantitative analyses (Chapters 5 and 6)

Professor Jason Gill, Professor Frances Mair, Dr Carlos Celis-Morales, Dr Bhautesh Jani, Dr. Barbara Nicholl, Professor Duncan Lee, and Professor Catherine O'Donnell were the group of co-authors who contributed to both sets of UK Biobank analyses. Unless otherwise stated, the contributions of authors were the same for both. I developed the study concepts and designs with Professor Jason Gill, Professor Frances Mair, and Professor Kate O'Donnell. Professor Jason Gill and Professor Duncan Lee provided statistical and analytical support. All authors assisted in the interpretation of the analysis findings. For the social connection work (Chapter 5) Professor Kate O'Donnell provided support with redrafting the introduction. All authors critically revised the manuscripts.

Qualitative work (Chapter 7)

I designed this study with support from Professor Frances Mair and Professor O'Donnel. I recruited the participants and conducted the focus groups and interviews with support from Professor Kate O'Donnell. I read all the transcripts in detail and Professor Frances Mair and Professor O'Donnel reviewed a selection of transcripts to assist in noting codes and developing themes from the data. Professor O'Donnel contributed to regular analysis sessions to discuss and refine codes. Professor Frances Mair assisted with later stages of coding refinement and interpretation.

Abbreviations

AHA	American Heart Association
ACM	all-cause mortality
AP	attributable portion
BMI	body mass index
CHD	coronary heart disease
CI	confidence intervals
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CLW	community links worker
CVD	cardiovascular disease
CVDM	cardiovascular disease mortality
LF	lifestyle factor
GP	general practitioner
GVIF	generalised variance inflation factors
HR	hazard ratio
IQR	interquartile range
LCI	lower confidence interval
LTC	long-term condition
MET	metabolic equivalent of task
MI	myocardial infarction
NCD	noncommunicable disease
NHANES	US National Health and Nutrition Examination Survey
NDI	Neighborhood Deprivation Index
NHS	National Health Service
NRS PPI	The National Health Service Research Scotland Primary Care Patient and Public Involvement Group
RERI	relative excess risk of interaction
PRISMA	Preferred Reporting in Systematic Reviews and Meta-analyses
PPHPs	Policymakers and public health professionals
RQ	research question
SES	socioeconomic status
SD	standard deviation
SDH	social determinants of health
SI	synergy index
TDI	Townsend Deprivation Index
TV	television
UCI	upper confidence interval
UKB	UK Biobank
WP	work package

1 Introduction

1.1 Chapter summary

This chapter provides a general introduction to the thesis, firstly by briefly presenting the main relevant concepts, namely, 1) lifestyle factors, 2) combinations of lifestyle factors, and 3) the potential impact of combinations of unhealthy lifestyle factors on the health and wellbeing of socioeconomically deprived populations. The thesis aims and objectives are then given alongside the corresponding research questions (RQ) and data and methods used to answer them. Finally, an outline of the chapters is provided.

1.2 Introduction to lifestyle factors

This thesis is concerned with the health risks associated with unhealthy lifestyle factors. Lifestyle factors (e.g., smoking, excessive alcohol consumption, poor diet, and low physical activity levels) are among the leading risk factors for mortality and morbidity.[1] Defining lifestyle factors is not straightforward but lifestyle factors are defined here as behavioural risk factors that are potentially modifiable by individuals and that are associated with higher risk of mortality and morbidity.[2]

The mechanisms by which lifestyle factors are thought to cause adverse health outcomes vary and are dependent on the factor but, generally, their associations with adverse health outcomes can be explained by their critical role in chains of physiological events that occur repeatedly and over extended periods of time and lead to cellular damage or impaired function (e.g., altered blood pressure, blood sugar, and immune and inflammatory function) that characterises noncommunicable diseases (NCDs) (e.g., cardiovascular disease (CVD), diabetes mellitus, and cancer).[3-6] This includes NCDs that are themselves sometimes defined as lifestyle factors but otherwise recognised as intermediate or metabolic risk factors (e.g., hypercholesterolaemia, obesity, and hypertension) for ‘harder’ adverse health outcomes (e.g., CVD and cancer mortality).

Intermediate metabolic risk factors are often considered alongside lifestyle factors because of their close and likely causal associations (e.g., physical activity and dietary factors are often considered key causal lifestyle factors for obesity). However, these risk factors are not considered lifestyle factors in this thesis. Here, only 'behavioural' factors that form an aspect of daily or regular living and that are associated with adverse health outcomes are considered lifestyle factors. Examples of these 'behavioural' lifestyle factors include well known or 'traditional' lifestyle factors like smoking, alcohol consumption, and physical activity.[7] While other examples include newer or 'emerging' lifestyle factors such as time spent sitting, sleeping, or immersed in nature.[7,8]

1.3 Combinations of lifestyle factors

Evidence for the association between lifestyle factors and adverse health outcomes has existed since smoking was linked to higher rates of cancer and CVD in the 1950s.[9] While smoking continues to be recognised as the most harmful of all lifestyle factors, the number and type of lifestyle factors that are considered important to health grows as more aspects of daily living are conceptualised, defined, and measured and their associations with adverse health outcomes studied.[7,10] Further, as societies change, for example, with shifts in culture, industry, politics, legislation, and with emergence of new infectious diseases and new technologies, so too do the ways in which people live, leading to new lifestyle factors.[11,12] However, research around lifestyle factors has tended to focus on individual factors and only more recently has the impact of multiple or combinations of lifestyle factors been examined.[13,14] This more recent research shows that, compared with just one lifestyle factor, there are even stronger associations between combinations of unhealthy lifestyle factors and adverse health outcomes. This is important because not only do unhealthy lifestyle factors tend to cluster within individuals where they could interact, but also because the prevalence of combinations (or the co-occurrence) of unhealthy lifestyle factors is unevenly distributed across populations.[15] Therefore, understanding the risks associated with combinations of lifestyle factors can lead to more accurate or 'real-world' estimates of the health risks due to lifestyle factors for individuals and sections of society that are most affected.

1.4 Social connection as a lifestyle factor

One risk factor or suite of risk factors that has more recently been suggested to be considered alongside, or even as a lifestyle factor itself, is social connection.[16,17] There are many reasons for this. Firstly, in terms of associations with adverse health outcomes, types of social connection (e.g., loneliness or social isolation) have been shown to have similar or larger effects sizes than those of well-known lifestyle factors like smoking or physical inactivity.[18] Secondly, there are numerous aspects of social connection that are behavioural and form daily or regular parts of people's lives (e.g., interacting with friends or family) and therefore could meet the definition of a lifestyle factor used in this thesis. Thirdly, social connection is strongly interrelated with numerous other lifestyle factors, having both direct and indirect effects on the likelihood of having other unhealthy lifestyle factors.[17] Finally, both awareness of the importance of social connection for health and funding to improve social connection have traditionally been lower than those for similarly important risk factors (e.g., smoking, alcohol intake).[18] Therefore, considering social connection together with other traditional lifestyle factors is seen as critical in improving awareness and support for tackling the growing problem of social disconnection.[19] However, there are few studies that consider the adverse health associated with traditionally recognised lifestyle factors together with measures of social connection.[10]

1.5 Lifestyle factors combinations in more deprived populations

The prevalence of unhealthy lifestyle factors and their co-occurrence tends to be highest among those from more socioeconomically deprived backgrounds and there is widespread evidence for a socioeconomic gradient in lifestyle factors.[15] There is also a wide body of literature that examines whether and by how much the socioeconomic gradient of unhealthy lifestyle factors might explain well recognised socioeconomic inequalities in adverse health outcomes. However, while adverse lifestyle factors have been found to explain between 6-80% of inequalities in health outcomes, they are rarely found to explain 100%.[20,21] Further, socioeconomic factors are often themselves conceptualised as independent risk factors for adverse health outcomes.[22,23] But there is far less

research examining the combined association of both unhealthy lifestyle factors and socioeconomic deprivation. Understanding the risks associated with both combinations of unhealthy lifestyle factors and socioeconomic deprivation together could provide better estimates of the health risks experienced by disadvantaged sections of society.[24] For example, interactions between lifestyle factors and deprivation could suggest there was disproportionate harm from lifestyle factors in more deprived communities and, therefore, promote the targeting and focussing of policy, interventions, and support for healthy lifestyle factors towards more deprived communities.[24]

1.6 Supporting healthy living among deprived populations

Interventions or policies for improving lifestyle factors can be at the level of individuals, communities, or the wider population. Attempts to support or improve healthy living and reduce health inequalities have often focussed on individual awareness or motivation to change and have often been informed by behavioural sciences.[25-28] However, there has been long standing recognition that addressing lifestyle factors only at the level of individuals generally leads to limited success, with improvements being small in scale and/or short lived.[29] Further, a focus on individual-level solutions increases the risk of exacerbating inequalities as those with more resources are able to utilise individual-level interventions more effectively. This is particularly important for the problem of combinations of unhealthy lifestyle factors which are more prevalent in more socioeconomically disadvantaged communities. Despite this long-standing recognition, the focus on individual-level interventions, motivation, and behaviour persists.[30] Understanding how people view unhealthy lifestyle factors in the context of deprivation could help develop new ideas or theories that could inform new ways of approaching lifestyle factor policies and interventions. However, there are few studies that examine how members of the public, health professionals, or policy makers perceive lifestyle factors in the context of deprivation.

1.7 Thesis aims and objectives

This thesis aims to address some of the gaps in evidence mentioned above and improve understanding of the adverse health associated with wide combinations of lifestyle factors in more deprived populations to inform intervention and policy development.

The specific research objectives are:

1) to describe and synthesise the evidence for the effect of socioeconomic deprivation on the association between combinations of unhealthy lifestyle factors and adverse health outcomes.

2) to examine the associations and interactions between measures of social connection and adverse health outcomes to help determine which measures of social connection could be included in a broad measurement of lifestyle i.e., a lifestyle score.

3) a) to examine the strength of associations between a wide range of individual lifestyle factors (including measures of social connection) and adverse health outcomes

   b) use the magnitude of effect sizes for numerous lifestyle factors to create a weighted a lifestyle score

   c) examine the effect of socioeconomic deprivation on the association between the weighted lifestyle score and adverse health outcomes.

4) to explore how key stakeholders view combinations of lifestyle factors in the context of deprivation

1.8 Research questions

The research questions (RQ) used to guide and address the research objectives are as follows:

RQ1. What is the existing evidence for socioeconomic influences on the association between combinations of unhealthy lifestyle factors and adverse health outcomes?

RQ2. What are the associations and interactions between components of social connection and adverse health outcomes?

RQ3. How does socioeconomic deprivation influence the association between a weighted lifestyle score and adverse health outcomes?

RQ4. What are the perceptions of key stakeholders around combinations of unhealthy lifestyle factors in the context of socioeconomic deprivation?

1.9 Introduction to data and methods

This is a mixed methods thesis and the data and methods used to answer each RQ are outlined briefly below. The thesis was planned and designed around three work packages (WP1-3), which are also indicated below.

1.9.1 RQ1 – review of existing evidence (WP1)

To synthesise the existing evidence for socioeconomic influences on the association between combinations of lifestyle factors and adverse health outcomes a systematic review was conducted. A protocol for the review was designed.[31] The aim was to provide a broad description and analysis of available evidence and wide inclusion criteria were used in terms of definitions of lifestyle factors and measures of socioeconomic status. It was decided to review prospective observational studies with meta-analysis where results permitted and a specified method of systematic review without meta-analysis (SWiM) if meta-analysis was inappropriate.[32] This systematic review formed WP1.

1.9.2 RQ2 and RQ3 – social connection and weighted lifestyle score analyses (WP2)

The large (n = 503,325) prospective, population-based cohort, UK Biobank, was used both to estimate the strength of associations and interactions between components of social connection and adverse health outcomes and to examine

how socioeconomic deprivation influences the association between a weighted lifestyle score and adverse health outcomes. UK Biobank was chosen as it contains rich baseline data with numerous variables for lifestyle factors, social connection measures, and socioeconomic factors and is also linked to registries to provide data on adverse health outcomes.[33] These two sets of quantitative analyses formed WP2.

1.9.3 RQ4 – stakeholder perspectives (WP3)

Qualitative research methods were chosen to explore and analyse the perceptions of key stakeholders around combinations of unhealthy lifestyle factors in the context of socioeconomic deprivation. Members of the public were recruited to take part in focus groups; semi-structured interviews were conducted with primary care practitioners (community nurses, community pharmacists, community links workers, and general practitioners), public health professionals, and policy makers. The focus group and interview transcripts were analysed qualitatively to permit in-depth exploration of participants' views of lifestyle factors and their combinations in deprived contexts. This qualitative study formed WP3.

1.10 Outline of chapters

The chapters included in this thesis are summarised below. Figure 1-1 shows the outline and corresponding work packages and research questions. This thesis is structured and submitted by alternative format where published and submitted papers are included as chapters. Chapters 4 and 5 are published papers and Chapters 6 and 7 are submitted papers. Chapters 1-3 and 8 are written as traditional chapters.

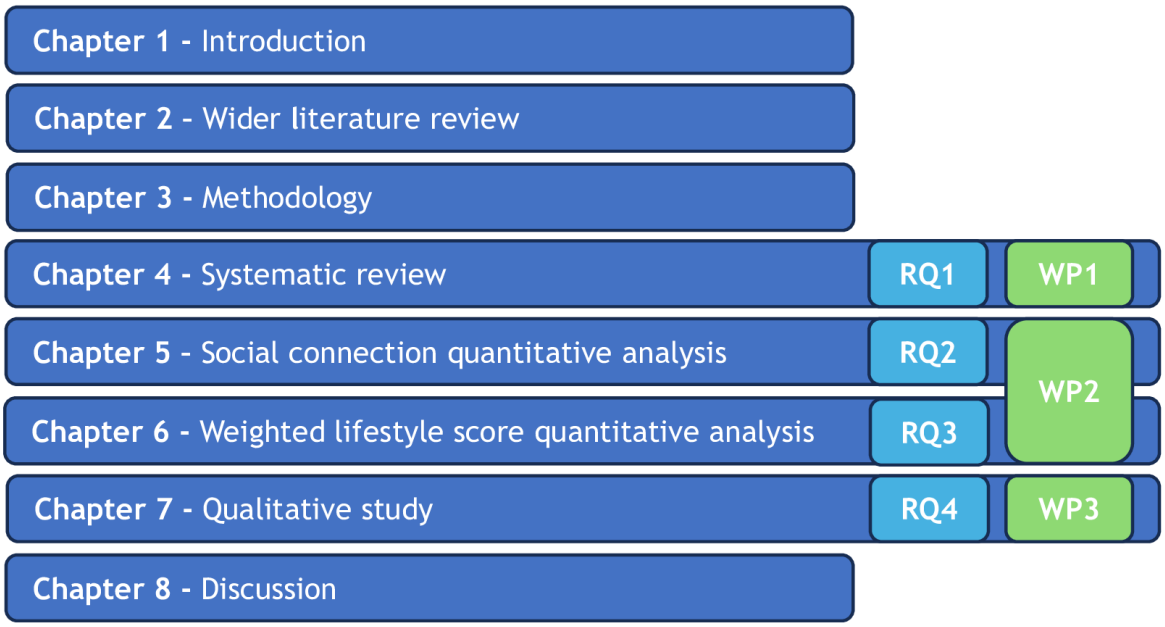


Figure 1-1: Outline of chapters with corresponding research question (RQ) and work package (WP)

Chapter 2 – Wider literature review

This chapter introduces the concept and epidemiology of lifestyle factors, lifestyle factor combinations, lifestyle factors in the context of socioeconomic deprivation, social connection (e.g., loneliness, and social isolation) and why this is considered alongside lifestyle factors, and differential vulnerability to unhealthy lifestyle factors. This chapter is structured to outline how the published and submitted papers form a comprehensive body of work.

Chapter 3 – Methodology

An overview of the wider methodological considerations and data used are given in this chapter. The strengths and limitations of the methods and data are described and concepts and frameworks that facilitates the synthesis of the results from this mixed methods thesis.

Chapter 4 – The influence of socioeconomic status on the association between unhealthy lifestyle factors and adverse health outcomes: a systematic review

This chapter is formed of a published systematic review that addresses RQ1 by providing a synthesis of available evidence from prospective cohort studies.

Chapter 5 – Social connection and mortality in UK Biobank: a prospective cohort analysis

This chapter presents the published UK Biobank analyses that answer RQ2 by examining the independent and combined associations between different components of social connection and all-cause and CVD mortality.

Chapter 6 – Sociodemographic effects on the association between a weighted lifestyle score and mortality in the UK Biobank cohort.

Chapter 6 is formed of a submitted paper containing UK Biobank analysis that addresses RQ3. The analyses examine the individual associations of 11 lifestyle factors (including a measure of social connection) and all-cause and CVD mortality. The results of those analyses are used to create a weighted lifestyle score to then examine the effects of socioeconomic deprivation, sex, ethnicity, and age on the association between the score and mortality.

Chapter 7 – Exploring public, practitioner, and policy maker perspectives of unhealthy lifestyle factors in the context of socioeconomic deprivation: A qualitative study.

This chapter presents the qualitative WP3 as submitted for publication and addresses RQ4.

Chapter 8 – Discussion

This final chapter summarises and synthesises the findings of the thesis with reference to the wider literature. The strengths and limitation of the thesis are discussed and implications for policy, practice, and future research are given.

2 Wider literature review

2.1 Chapter summary

This chapter provides an overview of the literature to set the context for the objectives outlined in the previous chapter. Firstly, the concepts of lifestyle and lifestyle factors and their associations with adverse health outcomes are discussed. The evidence for the association between combinations of lifestyle factors and adverse health outcomes and the importance of addressing them is then laid out. This is followed by a brief overview of the epidemiology of social connection and why this is included as a lifestyle factor here. Socioeconomic deprivation is then defined and the sociological and epidemiological links between deprivation and lifestyle discussed. A summary of how both lifestyle and deprivation have been considered in prior epidemiological literature in terms of 'differential exposure' and 'differential vulnerability', which are key concepts for this thesis, is then provided. The chapter concludes with putative causal mechanisms and theories that could explain how deprivation could exert a differential vulnerability to lifestyle factors.

This wider literature review is not a systematic review and therefore there may be gaps or foci of attention that may seem arbitrary. However, each topic within this literature review (i.e., each sub-heading of this chapter) comprises an enormous amount of literature and summarising each in depth is not feasible. Therefore, for transparency and to aid critique, the methods by which literature was identified and included in this section are briefly described below.

Work prior to this thesis provided much of the motivation and basis for the main research areas that were included:[24]

1. the concept of lifestyle and its relation to health

2. combinations of lifestyle factors (including newer lifestyle factors like components of social connection)

3. socioeconomic deprivation and lifestyle

4. differential exposure and vulnerability

5. mechanisms and theories for differential vulnerability

These areas were felt to be most pertinent to contextualise this project and justify the formulation of the research questions. PubMed was used to search for key terms related to each of these research topics. Key texts, such as relevant systematic reviews or meta-analyses, that were identified were then used for citation searches and Scopus was used for forward citation searching of key texts. Additional relevant literature was also highlighted by the wider research team during work for each of the analysis chapters.

As with all research, there is a level of subjectivity and interpretation that have also influenced the decisions around which literature to include for this review. As such, this literature review represents an intersection of my knowledge and experience of working as a general practitioner in an area of socioeconomic deprivation and as an early career researcher with an interest in how we live and how that affects our health. It also partly reflects the knowledge and experience of my supervisors and co-authors, which, for different parts, have included primary care researchers, epidemiologists, clinical academics, and statisticians. Finally, to summarise such an extensive range of literature, I aimed to provide broad overviews of each area and 'deeper dives' of relevant literature where helpful.

2.2 Lifestyle

2.2.1 Definition and etymology

The origin of the word 'lifestyle', 'life-style', or 'life style' in the English language is unclear and there may be several. However, as the origin of a word can influence how they are interpreted, conceptualised, or used, it is briefly reviewed here. The Oxford English dictionary suggests the word lifestyle arose from a literal translation of the German 'Lebensstil' as early as 1915 and, although not in common usage at the time, it appears to have had a similar meaning to present day usage.[34] Today, lifestyle is commonly taken to mean 'the characteristic manner or way in which a person lives'.[34]

However, the term ‘lifestyle’ is also ascribed to the psychologist, Alfred Alder, who described lifestyle as patterns of behaviour established in childhood.[35] ‘Lifestyle’ is also attributed to the sociologist Max Weber who is thought to have introduced the term to sociology although his original meaning is unclear.[36] Abel and Cockerham have argued that two of Weber’s German phrases ‘Lebensführung’ (life conduct) and ‘Lebensstil’ (lifestyle), in his widely cited work ‘Economy and Society’, have been inappropriately amalgamated through inaccurate translation of ‘Lebensführung’ into the English word ‘lifestyle’.[37] This, they argue, overlooks the idea that Weber’s concept of lifestyle comprised two separate components, life conduct (or life choices) and another term he used, ‘Lebenschancen’ (life chances, or the likelihood of realising life choices partly determined by wider socioeconomic conditions). Both because of this mistranslation and the way in which Weber changed the way he used the terms during his lifetime, Abel and Cockerham argue that the word lifestyle has been misinterpreted in sociology so that ‘life choices’ are overemphasised compared with ‘life chances’. This may, in part, explain why much discussion on lifestyle focusses on choices made by individuals, with the influence of wider life circumstances less prominent. Indeed, public health practitioners continue to highlight the problems associated with interpreting lifestyle as personal choices as opposed to opportunities.[38]

Less interestingly, but perhaps more likely, the English word ‘lifestyle’ could simply have arisen from an abridged version of the earlier phrase ‘style of life’ or ‘style of living’, which was not uncommon in the 18th century.[39] This would fit with a more organic appearance of the phrase rather than ‘lifestyle’ appearing abruptly in the English lexicon from a German translation.

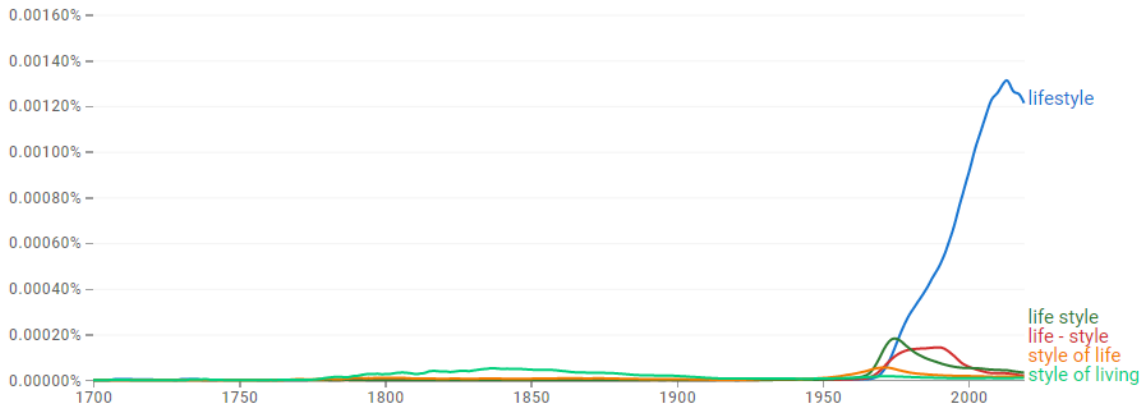


Figure 2-1: Google Books Ngram Viewer - 'lifestyle', 'life style', 'life-style', 'style of life' and 'style of living'
Google Books Ngram Viewer showing frequency of appearance of the words or phrases 'lifestyle', 'life style', 'life-style', 'style of life' and 'style of living' in >5 million digitised books from 1700s to present day.[40]

Usage of the term ‘lifestyle’ increased dramatically in English speaking popular and lay media from the 1960s onwards (Figure 2-1),[40] coinciding with parallel increases in disposable income, consumption, and leisure time, particularly in high-income countries like USA.[41,42] These parallel social changes may have been important in shaping how individuals and society consider or conceptualise ‘lifestyle’. Although, even earlier and perhaps more dramatic social changes in the late 1800s, with rapid industrialisation and the rise of a wealthy middle or ‘leisure’ class, prompted new social theories and concepts (e.g., ‘conspicuous consumption’ – purchasing status through public displays of material wealth) on how socioeconomic contexts affect the way in which people live and their health.[43] These historical examples highlight how societies undergo large scale changes alongside parallel changes in language and theory that attempt to explain and conceptualise the new ways in which people live. Therefore, understanding ‘lifestyle’ and how it relates to health requires examination in specific contexts.

2.2.2 Lifestyle or health behaviours

In the fields of medicine, medical sociology, and epidemiology, the word ‘lifestyle’ is often used interchangeably with ‘health behaviours’. While the term health behaviour also lacks a consensus definition it has long been preferred by those in public health to reduce both the affiliation with personal choice and victim blaming that the word lifestyle can imply.[38,44-46] However,

both lifestyle and health behaviours can be considered broadly as regular or daily human activities that contribute towards or detract from some aspect of health.

While there may be a sociological legacy that results in the term lifestyle overemphasising individual choice rather than life circumstances, the term lifestyle is used in this thesis as it remains well recognised in the context of potentially modifiable behaviours with strong associations with adverse health outcomes. Moreover, the interrelationship between lifestyle and wider social determinants of health is a major focus of this work. Therefore, the studies reported within this thesis have been conducted with keen consideration of the link between lifestyle and socioeconomic environment to avoid victim blaming or stigmatisation of those living with unhealthy lifestyle factors, especially those who are also from more disadvantaged socioeconomic groups.

2.3 Lifestyle factors and adverse health outcomes

Epidemiologically, the term lifestyle is often followed by 'factor' or 'risk factor', meaning those behaviours or practices that are associated with the risk of developing a disease or death.[47] The literature on lifestyle factors and their association with adverse health outcomes is vast and continues to grow. The number and types of lifestyle factors that are considered risk factors for adverse health outcomes has developed over time, and often after lengthy debate pitting health professionals and researchers against those with commercial vested interests in related 'unhealthy commodities'.[48,49] The classic example is that of cigarette smoking or tobacco, which was once considered healthy but is now clearly linked with numerous adverse health outcomes including mortality, CVD, and cancer outcomes.[50] Indeed smoking has been, and continues to be, considered among the most important and impactful unhealthy lifestyle factors.[1]

However, smoking has been joined by other unhealthy lifestyle factors as being strongly associated with adverse health outcomes, with a large body of literature examining their associations with adverse health outcomes. The World Health Organisation considers four unhealthy lifestyle factors (tobacco smoking, poor nutrition, harmful alcohol consumption, and physical inactivity – sometimes referred to as 'SNAP' risk factors) as the four modifiable risk factors with the

greatest global contributions to risk of death and NCDs.[51] However, lifestyle factors' associations with adverse health outcomes vary by the outcome of interest. For example, excess alcohol consumption is associated with oesophageal squamous cell carcinoma but not oesophageal adenocarcinoma.[52] And a range of dietary factors and physical inactivity (and overweight and obesity) are associated with up to 18 site-specific cancers but rarely with lung cancer.[53]

Further, additional lifestyle factors have emerged as having strong associations with all-cause mortality, and with CVD and cancer mortality and incidence.[10] There is no consensus as to what counts as an emerging lifestyle factor but they are defined here as those with a newer evidence base or those being considered alongside the more well-recognised factors of smoking, alcohol, physical activity, and diet.[7] Examples include sedentary behaviours or sedentary time (too much time sitting),[54] sleep (too long or too short or poor quality),[55-57] and social participation levels (infrequent or poor quality social contacts).[10,18,58]

The level of understanding of the mechanisms underlying the association between each lifestyle factor and each adverse health outcome varies. For smoking, there is a wide body of literature going back decades that has resulted in a detailed understanding of the pathophysiological mechanisms by which smoking causes lung cancer and CVD.[3,59] However, for other lifestyle factors, despite their strong associations and fit with causal criteria, questions remain over the exact mechanistic pathways by which the behaviours exert their health effects.[60] For example, while long sleep duration is associated with adverse health outcomes,[56] the underlying mechanisms to explain that association are poorly understood.[4]

2.4 Combinations of unhealthy lifestyle factors

Associations between lifestyle factors and adverse health outcomes are generally considered independent as they are often established after controlling for likely potential confounders including other lifestyle factors. However, lifestyle factors cluster within individuals, where individuals with one unhealthy behaviour are likely to have another.[15,61-68] For observational studies, this presents a potential area of residual confounding. For example, the association between physical

activity, sedentary behaviours, diet, and adverse health outcomes may be confounded by other lifestyle factors (e.g., sleep duration time) which may be unaccounted for in analyses.[69] Further, combinations of lifestyle factors may have additive (or multiplicative/synergistic) interaction effects, where the combined effect of two unhealthy lifestyle factors is greater than the sum (or the product) of the individual effects.[69,70] Hence it can be helpful to examine combinations of lifestyle factors where a larger number of lifestyle factors are taken into consideration as exposures.

Consistent with additive or synergistic interactions between lifestyle factors is the finding that combinations of lifestyle factors have even stronger associations with adverse health outcomes than individual lifestyle factors alone.[71-73] Generally, there is a clear dose response relationship where the more unhealthy lifestyle factors are present, the higher the risk of adverse health outcomes.[13,14,74] Although that relationship may be non-linear and there are patterns of clustering where certain behaviours are more or less likely to co-occur, which suggests that the specific combination as well as number of unhealthy lifestyle factors is important.[65,68,75,76] Also, the prevalence of specific combinations of unhealthy lifestyle factors may vary across sub-groups of society, which has implications for interventions that address combinations of unhealthy lifestyle factors.[77]

The literature on combinations of lifestyle factors continues to develop. Three recent systematic reviews by Zhang et al, have combined the evidence from prospective studies finding evidence for clear associations between three or more lifestyle factors and the adverse health outcomes of: all-cause mortality (87 studies), CVD incidence (25 studies) and mortality (55 studies),[72] cancer incidence (21 studies all cancer types; 35 studies site-specific cancers) and mortality (38 studies),[73] and type two diabetes (14 studies).[71]

Addressing multiple lifestyle factors has therefore been cited as an urgent and unmet challenge for improving population health.[78] However, most interventions addressing multiple lifestyle factors focus on the individual and success has proved challenging with many remaining unknowns on how to address combinations of lifestyle factors.[79] For example, a systematic review of randomised trials of interventions that target multiple lifestyle factors has

shown that the order in which to address each factor remains unclear – there were few studies that compared interventions that address the lifestyle factors sequentially versus simultaneously.[80] However, that review also suggested that addressing multiple lifestyle factors (either sequentially or simultaneously) were generally more effective than a usual care comparator. A subsequent scoping review comparing single versus multiple lifestyle factor change interventions for weight loss also found that the literature was limited but that there was a suggestion that demographic characteristics and use of theory in intervention design were associated with greater effectiveness.[81]

The level of effectiveness of interventions attempting to address numerous lifestyle factors has also been modest. A meta-analysis of 69 randomised controlled trials of non-pharmacological interventions attempting to improve at least two lifestyle factors (most often diet and physical activity, but also smoking and alcohol consumption) found that interventions (which were mainly based on education and training) had only modest effects on lifestyle factors.[82] Interestingly, although there were reductions in smoking prevalence in the studies targeting this factor, smoking was negatively associated with the intervention when there were improvements in other lifestyle factors. This could suggest that smoking should be addressed sequentially with other lifestyle factors rather than simultaneously. Finally, existing studies of interventions addressing multiple lifestyle factors also lack diversity of research participants (few studies examined minority ethnic groups, younger adults, or socioeconomically disadvantaged groups) and examine only a small range of lifestyle factors.[79] This is important as the prevalence of lifestyle factors follow socioeconomic gradients (e.g., prevalence of unhealthy lifestyle factors are higher in minority ethnic and socioeconomically disadvantaged groups) and, therefore, there is a lack of evidence of what works in populations most in need of support.[15,62,83]

Nevertheless, there are good arguments to attempt addressing numerous unhealthy lifestyle factors in a single intervention. Reasons include the potential interaction between lifestyle factors and the greater adverse health outcomes associated with combinations of lifestyle factors, but also because of the inequalities in health outcomes of the populations in whom unhealthy lifestyle factors tend to cluster, namely more socioeconomically deprived populations.

Before going on to describe the links between unhealthy lifestyle factors, socioeconomic deprivation, and adverse health outcomes, the following section introduces the importance of considering social connection as a lifestyle factor.

2.5 Social connection

Social connection is difficult to define and conceptualise due to the almost infinite ways in which humans interact. However, humans have evolved social behaviours with clear survival advantages from regular interactions, close co-operation, and being supported by others, which explain a near universal desire for social conections.[84-87] Further, the evidence of associations between lacking social connections and a range of adverse health outcomes is extensive and growing with the result that social connection is seen as essential for health.[87-89]

There are numerous mechanisms by which social connection could influence health outcomes because of the myriad ways and modes in which humans connect and interact.[85] These purported mechanisms are not a focus of this thesis but a brief overview and examples of mechanisms are given in Chapter 5. Below, a summary of specific epidemiological issues that relate to social connection in this thesis is given before describing why social connection can be considered a lifestyle factor.

2.5.1 Social connection epidemiology

Two key exposures are often examined in the epidemiological literature that relate to social connection, namely, social isolation and loneliness. Social isolation is typically defined as lacking a certain amount or number of social contacts.[90] Whereas loneliness is defined as perceiving a lack of connection.[90] Therefore, social isolation is often considered an objective measure of social connection and loneliness a subjective one.[90] There is a wealth of literature that examines the links between both social isolation and loneliness and adverse health outcomes.[88,89,91-101]

However, there are gaps and inconsistencies in the literature. Firstly, the two terms, social isolation and loneliness, have been used interchangeably despite their differences, which reduces accuracy and clarity around the problems.

Secondly, whilst there is long standing appreciation that the two exposures are often correlated but not the same, there is a lack of research that examines their separate and combined effects.[19] Indeed, metanalyses that provide the effect estimates for each draw on studies that for the vast majority do not mutually adjust for each.[88,89,95,97,101-103] Therefore, our current best estimates for how loneliness is associated with adverse health outcomes could be confounded by an unmeasured effect of social isolation and vice versa. And while the impact of each is purported to be independent (even though meta-analytical evidence for this is lacking) due to their often-observed discrepancy (i.e., individuals defined as socially isolated who do not report feeling lonely and vice versa) the impact of being both socially isolated and simultaneously feeling lonely has been overlooked.[19]

The objective for the study in Chapter 5 was to address those gaps by examining the independent and combined associations between different components of social connection and adverse health outcomes.

2.5.2 Social connection framework

One way of trying to improve and extend the evidence base around social isolation and loneliness is via use of a conceptual framework of social connection. Holt-Lundstad has developed this framework and defined three components of social connection: structural, functional, and quality (Table 2-1).[87]

Table 2-1: Framework of components of social connection
(Taken from Holt-Lundstad, 2018)[87]

Component	Definition	Examples measures
Structural	Existence of and interconnections among different social relationships and roles	Marital status Living alone or not Social networks Social integration Social isolation
Functional	Functions provided by, or perceived to be available because of, social relationships	Received support Perceptions of social support Perceived loneliness
Quality	The positive and negative aspects of social relationships	Marital quality Relationship strain Social inclusion or exclusion

This framework offers a broader approach that encompasses both social isolation and loneliness as well as other measures that could be considered structural or functional measures of social connection that may not be included in other concepts of social isolation and loneliness. For example, an individual could report low levels of social support but may not feel lonely and therefore examining feelings of loneliness as the functional component would overlook perceptions of support. The framework also includes a quality component of relationships and thereby includes a spectrum of positive and negative attributes for different types of social connection and thus provides a more comprehensive assessment of social connection.

I am unaware of alternative frameworks that incorporate such a wide range of different components of social connection and of any published critiques of this framework. However, whilst the framework has the advantage of systematising and organising a wide range of different ways in which humans interact and connect under the umbrella term 'social connection', there may be limitations to the framework.

Firstly, although the framework provides a construct that can assist in the interpretation of the broad evidence base around human connectedness and health, the framework originates from existing and known measures or concepts (e.g., social isolation, loneliness, relationship quality) that have been used in research. Therefore, the framework could be incomplete and would need to be amended as new measures are developed or new aspects of human interaction are defined and conceptualised. For example, new research could identify that different 'modes' or media (e.g., online versus face-to-face) through which humans connect and interact as being associated with adverse health outcomes. If the impact of these modes in which people connect were independent of the structural, functional, and quality components of connections then it would suggest that a new component of social connection should be defined and included within the framework. Therefore, whilst the framework appears to be broad to allow new measures to be incorporated, it may be that use of the framework precludes consideration of other less considered or hitherto undiscovered aspects or components of social connection. Secondly, Holt-

Lundstad et al. suggest that the quality component of social connection represents positive and negative attributes of relationships or connections that are independent of the functional component of social connection. However, it could be argued that quality and functional components appear to be more closely intertwined and less independent than functional and structural components. For example, negative subjective feelings (e.g., perceptions of loneliness) are likely to correlate with poor quality connections (e.g., low relationship satisfaction ratings). And where they do not correlate that could raise questions over the accuracy of the measure of quality. Thirdly, and finally, the advantage of a single framework in organising numerous interacting components also risks reducing or oversimplifying what remains a highly complex, interacting, and dynamic phenomenon.

Nevertheless, this framework was used for the work in Chapter 5 as it could aid in conceptualising and interpreting the independent and combined associations between different components of social connection and adverse health outcomes.

2.5.3 Social connection as a lifestyle factor

Social connection is an increasingly recognised risk factor alongside more traditional lifestyle factors and even considered as a lifestyle factor itself.[16,17] There are several reasons for this.[17]

Firstly, components of social connection (e.g., loneliness or social isolation) have been shown to have similar or larger effect sizes than those of well-known lifestyle factors like smoking or physical inactivity.[18] Secondly, there are numerous aspects of social connection that are behavioural and form daily or regular parts of people's lives (e.g., interacting with friends or family) and therefore would meet the definition of a lifestyle factor used in this thesis. Thirdly, social connection is strongly interrelated with numerous other lifestyle factors, having both direct and indirect effects on the likelihood of having other unhealthy lifestyle factors.[17] Finally, both awareness of the importance of social connection for health and funding to improve social connection have traditionally been lower than those for similarly important risk factors (e.g., smoking, alcohol intake).[18]

Therefore, considering social connection together with other traditional lifestyle factors is seen as critical in improving awareness and support for tackling the growing problem of social disconnection.[19] However, there are few studies that consider the adverse health associated with traditionally recognised lifestyle factors together with measures of social connection.[10] It was for these reasons that measures of social connection were included as a component of a lifestyle score in the analysis that forms Chapter 6.

2.5.4 Social disconnection prevalence

Regardless of how it is measured or conceived, a lack of social connection (e.g., loneliness and social isolation) is seen as a growing public health problem with increasing prevalence globally.[98,99] Chapter 5 uses data from the UK and therefore the prevalence and scale of the problem in the UK is considered briefly here.

The 2021/22 Community Life Survey (a self-report household survey), based on a representative sample of 10,126 adults aged 16+ in England, showed 6% of participants often or always felt lonely.[104] Similarly, analysis of Office for National Statistics data by the Campaign to End Loneliness, suggests that up to 3.24 million people (6%) in the UK were lonely in 2020.[105] However, this figure has increased since then with up to 500,000 more adults feeling lonely in the UK as of 2022.[105] The same analysis shows that this increase is not evenly distributed across demographic groups with the increase in loneliness mainly being in those aged 30-70, men, those of white ethnicity, and those with poor-health.

In 2023, 8.4 million people were living alone in the UK, 8% more than in 2013 and greater than the 6% population increase over that time.[106] Again this increase was not evenly distributed with 93% of the increase in those living alone seen in those aged 65 years or older. With just over half of those living alone (4.3 million) being aged 65 years or older this represents a far larger population than those living in a care home. Census data from 2021 shows there were 278,946 people aged 65 years or older living in a care home in England and Wales; this represents a 4% decrease in the numbers of that age group living in care homes.[107]

Interventions to address social disconnection vary by method and by evidence of efficacy. A review of 60 studies provides weak observational and moderate RCT-based evidence for a range of interventions to reduce loneliness in community-dwelling older adults.[108] Examples of interventions in the review include group-based therapy/exercises/activities, individual in-person interactions, internet-based interventions, and telephone-delivered interventions. Similarly, a scoping review of reviews of loneliness and isolation interventions for older adults suggested the individuality of any subjective experience of isolation meant there would be no one-size-fits-all approach to combating loneliness and social isolation.[109] The authors suggest that a realist approach should be adopted for intervention research in this area in order to establish which intervention works for whom and in what context. Understanding the independent and combined associations between different components of social connection and adverse health outcomes could help inform this intervention research as interventions could be conceptualised and tailored within the social connection framework.

2.6 Socioeconomic status, deprivation, lifestyle, and adverse health outcomes

This section reviews the theoretical concepts of socioeconomic status (SES) and deprivation and how they relate to lifestyle factors and health. It also gives an overview of the epidemiological literature that describes the associations between SES and adverse health outcomes and between socioeconomic deprivation and unhealthy lifestyle factors.

2.6.1 Historical and theoretical perspectives on the association between socioeconomic status, health, and lifestyle

SES is a theoretical construct that stratifies populations based on their characteristics, the environment in which they live and work (e.g., occupation type or class, housing conditions, area of residence), and on their resources (e.g., financial, educational, and material).[110-112]

Much of the underlying theoretical work that has resulted in the development of various measures of SES in relation to health can be attributed to a number of sociologists, philosophers, and clinicians since the industrial revolution such as Engels,[113] Virchow,[114] Durkheim,[115] Weber,[36] Bourdieu,[116] and many others.[117]

These writers have also made significant and lasting contributions to shaping the understanding of how socioeconomic conditions influence human health and concepts of 'lifestyle' and so are particularly relevant for this thesis.[118] Therefore, some of their work is briefly discussed here.

The recognition of wider social and environmental influences on disease and behaviour has much longer historical roots that predate these writers. For example, medical students and physicians more than 2,000 years ago who were familiar with the Hippocratic treatise, *On Airs, Waters, Places*, would have acknowledged the influence of local conditions, such as the weather and water sources on the health of local populations and on:

> 'The mode of life also of the inhabitants that is pleasing to them, whether they are heavy drinkers, taking lunch, and inactive, or athletic, industrious, eating much and drinking little,'[119] (i.e., lifestyle factors).

More recently, the birth of public health and understanding of the influence of social contexts of health has been attributed to those such as Rudolph Virchow (1848), who identified the cause of an epidemic as social and political (i.e., famine, war, and poverty).[114] And to work such as that of Friedreich Engels (1845) who showed mortality rates of those with poor living conditions and working in factories in industrial cities was worse than that of surrounding areas.[113]

Sociologists such as Weber (1922) have sought to provide frameworks to explain how and why society is stratified along socially constructed dimensions or 'structures'.[36] The sociological concept of structure delineates groups of people who share common prospects (including economic and health prospects) and patterns of behaviour. Bourdieu's work (1977) is thought to have extended that of Weber's through additional concepts such as 'habitus', which he defines as:

> 'systems of durable, transposable dispositions, structured structures predisposed to operate as structuring structures, that is, as principles which generate and organize practices and representations that can be objectively adapted to their outcomes without presupposing a conscious aiming at ends or an express mastery of the operations necessary in order to attain them.'[116]

This definition of habitus has been more simply redefined as:

> ‘a cognitive map or set of perceptions in the mind that routinely guides and evaluates a person’s choices and options.’[120]

Bourdieu also described other social structures and concepts such as ‘social capital’, ‘field’, and ‘practice’.[121] However, it is sufficient here to state that Bourdieu sought to explain behaviour within a given social context with habitus being an inclination or disposition for behaviours given societal influences. Bourdieu’s concept of habitus is helpful in relation to SES and lifestyle as it describes how an individual’s expectations of how others will respond to their actions influences that individual’s decisions and behaviour. Habitus determines the expectations and the likelihood of the options available to individuals due to the order or structure in society and to that individual’s position (and their perception of their position) in the social order. Both Weber and Bourdieu try to disentangle agency and structure in relation to behaviour. Agency is the ability of individuals to choose their actions (Weber’s ‘life conduct’) and structure is the socially created rules, norms, and etiquettes (Weber’s ‘life chances’) in which individuals live.

Weber described behaviour as the outcome of a ‘dialectic interplay’ between agency and structure and his work has been interpreted as favouring agency over structure to explain behaviour. Conversely, Bourdieu is considered to have emphasised structural influences over agency.[120]

Other theoretical constructs have arisen specifically in relation to unhealthy lifestyle factors. For example, Blue et al. develop the idea from Weber and Bourdieu (among others) that an unhealthy lifestyle factor, like the excessive consumption of alcohol, is not a single behaviour as conventionally viewed, but a collection of behaviours, or a social practice (e.g., drinking with colleagues after work as opposed to simply alcohol consumption).[122]

The benefit of conceptualising groups of more specific behaviours of lifestyle factors with certain social contexts is to provide a unit of operation or even a target for intervention that straddles the divide between agency and structure. Social practice theory creates new definitions for what the problems are (e.g., as previously stated, drinking after work with colleagues as opposed to a single focus on alcohol consumption). By highlighting the interconnected relationships

between behavioural complexes and groups of individuals, it indicates why interventions to reduce associated harm are required at community and sociocultural-levels rather than the individual.[123] However, social practice theory risks defining numerous related and overlapping practices that lack clear boundaries and focus, which make practical implications difficult to interpret or apply.[123] For example, drinking alcohol at home alone, drinking after work with colleagues, and drinking secretly as young teenagers could be separate practices that require their own interventions. Whereas these practices might all be influenced or improved by focussing on common population-level factors of alcohol consumption such as legal control of alcohol sales.

Cockerham has incorporated prior work of sociologists to create a 'health lifestyle theory'.[124] This theory considers lifestyle factors within a framework that acknowledges the influences of both agency and structure to ultimately determine lifestyle.[120,124] Figure 2-2 shows a simplified version of this 'paradigm'.

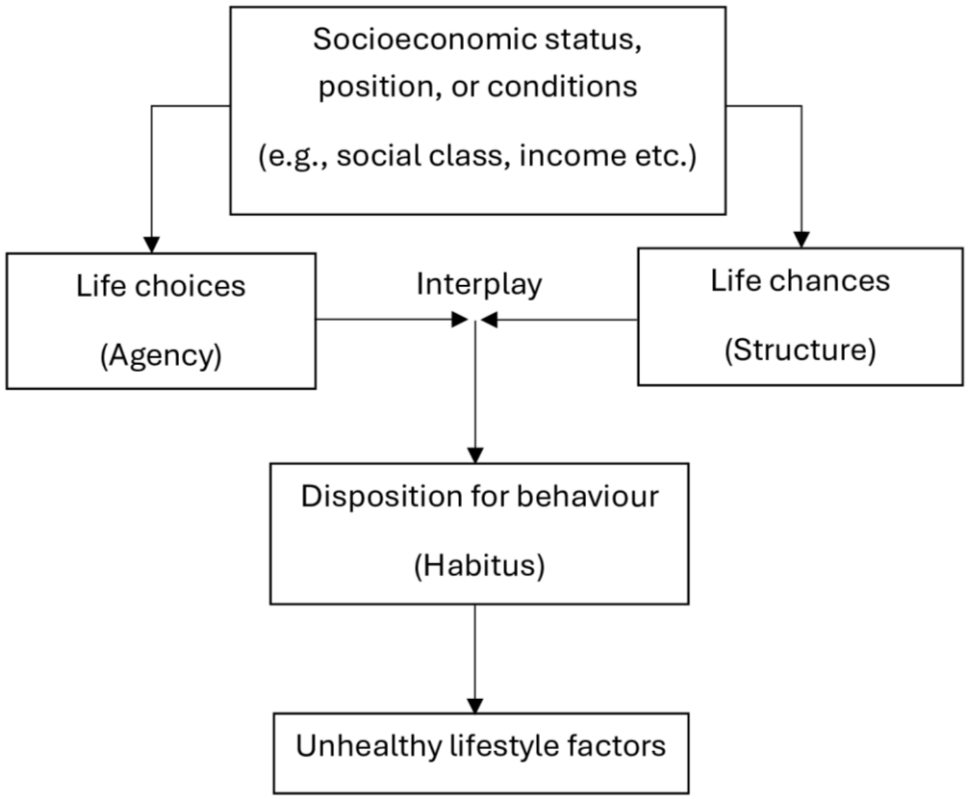


Figure 2-2: Simplified version of Cockerham's health lifestyle paradigm
Adapted from Cockerham (2021).[120] Arrows indicate hypothesised causal relationships.

I briefly describe Figure 2-2 here, using smoking as an example of an unhealthy lifestyle factor. Working from the top to the bottom of the figure, the first set of arrows represent the causal influences of social class on both the choices made by individuals who smoke ('Life choices' or agency) and on the opportunities ('Life chances' or structure) for smoking afforded to individuals. The life choices and chances for smoking then interact ('Interplay') to cause a predilection or 'disposition' (habitus) for smoking, which then, in turn, causes smoking (or not).

While these social theories and explanations for relationships between SES, lifestyle, and health, have not been used formally in analysis of the work in this thesis, they have informed the work more generally including the interpretation and synthesis of its findings.

2.6.2 Socioeconomic status versus socioeconomic position

Confusingly, the terms SES, socioeconomic position (SEP), social class, and social status have all been used interchangeably by researchers and the lay public as there is no widely accepted consensus on these terms. However, each term has different theoretical underpinnings and historical origins. Further, clarity of what is meant by each term is important because where a measure of SES is associated with health the origins and meanings of that measure could have implications for causal mechanisms and theory.[125,126] A brief description is given here for how the origins and concepts differentiate these terms.

Bartley describes social class, at least in sociology and medical sociology, as originating from the work of Karl Marx and Max Weber and involves grouping populations according to their occupation and working conditions (e.g., by ownership of assets and relationships between those who must work and those who manage or supervise their work).[127] In the original sense of the term, there was no hierarchy by 'class' although, prima facie, there would appear to be a clear hierarchy of power with this system of stratification.

Whereas, in lay terms, social class more often means what social status is taken to mean in the medical sociological literature, which involves ranking populations, again according to occupations but linked to prestige and honour

associated with occupations (i.e., similar to the Hindu Caste system). This definition of social status more explicitly involves a hierarchy.

The term SES also implies hierarchy and is strongly linked with developments in sociology and concepts in the US.[128] Occupations of 'higher status' were those deemed important to the functioning of society and therefore needed to attract those with the talents or skills to do that more important work.[128] The difference with this 'structural functionalism' approach to the ranking of occupations and society is that the hierarchy is considered as borne out of people's natural abilities rather than, for example, inherited ownership of assets or relationships to those who supervise their work. Initially, status in the US was measured in small surveys based on ranking occupations according to reputations and prestige, but, as the ranking became harder to judge in larger nationwide surveys, occupations were later ranked according to a combination of average income and education level. Bartley, therefore, writes that measures of SES have their roots in structural functionalism and are actually measures of qualities of individuals which afforded them different levels of socially perceived prestige.[127]

Therefore, use of the term SES in this thesis could be misinterpreted as pertaining to the qualities of individuals based on judgements of status linked to societal functions of occupations. However, meaning of SES in this thesis mainly relates to all population stratifications that are hierarchical, for example, perceptions (e.g., of honour and prestige), ownership of assets or material wealth, educational opportunities, and power. Nevertheless, all types of social characteristics, contexts, and constructs that differentiate groups of people could influence and, therefore, explain differences in both lifestyle factors and health outcomes. Therefore, no system of social stratification was excluded from consideration in this thesis, including those that are less clearly hierarchical.

2.6.3 Socioeconomic status and adverse health outcomes

Based on theoretical constructs that describe and explain how and why society is stratified, there is now a wealth of literature detailing the association between SES and adverse health outcomes, whereby populations experiencing a lower SES

(lower income, lower educational attainment, poorer housing or neighbourhoods etc.) tend to have more adverse outcomes across most outcome measures.[129] However, the broad nature of SES and the subsequent wide range of its definitions and measures mean that the strength of association between SES and adverse health outcomes varies by measure, population, and health outcome.[22,130-133]

Each SES measure has strengths and limitations in terms of estimating inequalities in health and therefore attention is needed when interpreting results using each.[134-138] For example, occupation based measures often exclude those who are unemployed such as students or those in informal work or who are retired and thus may misclassify populations by, for example, grouping 'wealthier' students with 'poorer' unemployed.[125,126] Further, SES measures are context dependent which can also lead to misclassification. For example, changes in an individual's circumstances during their lifetime could mean that a single measure of individual-level income would fail to capture inequalities that are heavily influenced by events and circumstances in early life or childhood. In addition, measures of SES developed in one setting could be less applicable to those in a different setting. For instance, an asset-based index derived in high-income countries could misclassify populations and fail to capture within-population inequalities when applied in low-income country settings.[139] This misclassification could lead to unpredictable effects on SES-outcome associations.[140]

While there is ample evidence for strong associations between lower SES and adverse health outcomes, there is greater variation in this association than tends to be recognised.[131] For example, in a national representative sample of 2,036 adults in Sweden, lower income had stronger associations with adverse health outcomes than measures of occupation and education.[132] While, among women in England of lower, compared to higher SES (as measured by a deprivation index), between 1997-2011, there were 6% more breast cancer deaths, 241% more oropharyngeal cancer deaths, but 34% fewer melanoma deaths.[141] Therefore, while often based on availability, the measure of SES chosen and the health outcome under consideration both influence the estimates of the relationships between exposures and health outcomes and therefore subsequent inferences.

The clear social patterning of adverse health outcomes has resulted in a school of thought around focussing on the wider social context or environment, the so-called social determinants of health (SDH).[142,143] SDH have been categorised in an overarching SDH framework as either structural or intermediary determinants.[144] In this framework, structural determinants include socioeconomic and political context (e.g., macroeconomic policies) and SES (e.g., education, social class). The main groups of intermediary SDH that are considered to explain health inequalities are material, psychosocial, behavioural and health system factors. Material factors are those such as the physical environment (e.g., housing, air pollution, or working conditions). Psychosocial factors relate to people's perceptions of their status as unequal in society leading to stress and worse health. Behavioural factors are the unhealthy lifestyle factors that are a focus of this thesis. However, as alluded to in the overview of theoretical perspectives in relation to SES and lifestyle, these types and subtypes of SDH are interrelated and not easily separable.

2.6.4 Socioeconomic deprivation

Socioeconomic deprivation (hereafter 'deprivation') is one way of considering SES. Deprivation is based on the idea that sections of a population are deprived of certain resources or ways of living that are customary or reasonably expected for a given population.[145,146] Measures of deprivation are therefore relative as they relate to what can be expected or to some kind of population standard.

Numerous measures of deprivation have been used to investigate inequalities in health (Table 2-2). These measures are composites, using multiple components, and are typically area-based often grouping census tracts, political wards, or residential areas together. For example, the Scottish Index of Multiple Deprivation uses data zones of around 700-800 people.[147] Each measure of deprivation varies in its constituent components and each have advantages and disadvantages for investigating adverse health outcomes.[112,131,148,149]

Table 2-2: Example measures of deprivation.

Measure of deprivation	Components
Townsend deprivation index[150]	• Unemployment • Non-car ownership • Non-home ownership • Household overcrowding

Jarman index[151]	• Unemployment • Lone pensioners • Aged under 5 • Unskilled workers • Household overcrowding • Recent change of address in • Ethnic minorities • Single parent
Carstairs deprivation index[152]	• Male unemployment • Overcrowded households • Non-ownership of car or van • Low social class
English Index of multiple deprivation (EIMD)[153]	39 indicators from: • Income • Employment • Education, Skills, and Training • Health and Disability • Crime • Barriers to Housing and Services • Living Environment
Scottish Index of multiple deprivation (SIMD)[147,154]	37 indicators from: • Income • Employment • Health • Education, Skills, and Training • Geographic Access to Services • Crime • Housing

Poverty, another measure of SES, is sometimes used interchangeably with deprivation although it tends to have a greater focus on income compared with the wide range of resources and conditions included in measures of deprivation. Income is often measured on an individual or household level. Other examples of individual-level SES measures include those based on education, occupation type or class, and employment status. Measures of poverty, or income, can be relative (e.g., to a population median) or absolute but even absolute measures of poverty are relative to a standard based on judgements of social norms. For example, the World Bank definition of the line below which people live in extreme poverty is on an income of <US $2.15/day.[155,156]

2.6.5 Deprivation and unhealthy lifestyle factors

Irrespective of SES measure, there is wide ranging evidence that shows how, at least among high-income countries, that there is a strong and consistent association between multiple unhealthy lifestyle factors and lower SES.[15,61,66,67,157-162] Indeed, those who have derived measures of SES have included lifestyle factors themselves as part of measures of deprivation. For example, Townsend, who carried out much of the foundational work to define deprivation, mentions the now commonly considered lifestyle factor of diet in his definition of deprivation:

> 'people can be said to be deprived if they lack the types of diet, clothing, housing, household facilities and fuel and environmental, educational, working and social conditions, activities and facilities which are customary, or at least widely encouraged and approved, in the societies to which they belong'.[146]

The higher clustering of lifestyle factors in more deprived populations is often considered a major part of the explanation for observed inequalities in adverse health outcomes. As a result, lifestyle factors have long been seen as potential targets of health interventions as they have been considered as potentially easier, at least on an individual-level, to modify compared with factors such as education, income, or housing. However, this 'differential exposure' to unhealthy lifestyle factors across SES strata only ever partially accounts for health inequalities. This is perhaps not surprising considering the main factors that contribute to the SDH include material, psychosocial, health-system, and political factors as well as behavioural ones.

Therefore, given the potentially independent effects that these various SDH factors exert there are surprisingly few studies examining the separate and combined effects of SDH factors on adverse health outcomes. For example, while scoping the literature and developing the protocol for the systematic review in this thesis, relatively few studies were found that examined how material and psychosocial deprivation factors influenced the association between multiple unhealthy lifestyle factors and adverse health outcomes (Appendix 1).[31] This represented a potential gap in the literature where it was unknown if there were important differences in risk associated with unhealthy lifestyle factors across SES strata i.e., an SES-based differential of vulnerability

to unhealthy lifestyle factors. This then formed the motivation for trying to synthesise the evidence for how SES might influence associations between combinations of unhealthy lifestyle factors and adverse health outcomes. The resultant systematic review is reported in Chapter 4.

2.7 Differential exposure and vulnerability to unhealthy lifestyle factors

This section first reviews some of the epidemiological evidence on how differential exposure to unhealthy lifestyle factors can explain socioeconomic gradients in adverse health outcomes. It then discusses the concept of differential vulnerability to unhealthy lifestyle factors before describing some of the potential mechanisms that could explain an SES-based differential in vulnerability to unhealthy lifestyle factors.

2.7.1 Differential exposure

Across a range of studies, the higher prevalence of unhealthy lifestyle factors in lower SES groups has been estimated to explain between 6% and 80% of mortality inequalities.[20,21,163,164] A systematic review investigating the contribution of four lifestyle factors (smoking, alcohol consumption, physical activity and diet) towards socioeconomic inequalities in all-cause mortality and cardiometabolic disorders, found that the level of contribution varied by geographic region, health outcome examined, demographic group, and study design.[165] For example, among older adults lifestyle factors appeared to explain less of the gradient in health inequalities: a study of older men in Australia found less than a quarter of the SES disparity in health was explained by smoking, alcohol consumption, physical activity, and body mass index (BMI);[166] and a study of adults in China aged 65 or older, showed that a healthy lifestyle (measured by smoking, alcohol consumption, physical activity, and diet) explained 9% of inequalities in all-cause mortality.[167] In contrast, in a study from Norway, smoking, alcohol, physical activity, and BMI were estimated to account for 40% of the education-linked gradient in health related quality of life.[168]

The potential reasons for these wide variations in how much unhealthy lifestyle factors explain SES variations in adverse health outcomes are likely to be many

and include the strengths and limitations of the SES measures used as explained above, as well as variation in the measurement of lifestyle factors. For example, observational studies often rely on a single baseline measurement of lifestyle exposures whereas repeated measurements of lifestyle factors could provide more accurate estimates of their contributions to socioeconomic inequalities in health.[169]

SES measures tend to be well correlated as there is overlap between their underlying constructs. However, the differences between them may have important implications here where SES measures vary in their associations with lifestyle factors.[159] For example, an area-based measure of SES may have different associations with active commuting than say level of education if there are strong geographical influences on active commuting. This implies that research examining the associations between unhealthy lifestyle factors, SES, and adverse health outcomes should examine a range of SES measures where possible. Variations by SES measure could inform potential mechanisms and therefore interventions to improve lifestyle factors.[125]

However, irrespective of the SES measure used, the evidence suggests that while health inequalities are widening over time in many countries, SES inequalities in unhealthy lifestyle factors are also widening.[170] Therefore, deeper understanding of the relationships between unhealthy lifestyle factors, SES, and adverse health outcomes could inform interventions to reverse those trends. One way of deepening the understanding of those relationships is through examining differential vulnerability.

2.7.2 Differential vulnerability

Differential vulnerability, in the context of unhealthy lifestyle factors and SES, is the concept that, over and above the differential exposure to unhealthy lifestyle factors along SES gradients, there is also a differential effect of those factors along SES gradients that helps to explain health inequalities.[171] This suggests that even if there was no differential exposure (i.e., the same prevalence or level of unhealthy lifestyle factors across SES groups) the effect of the unhealthy lifestyle factors would still be unequal across SES groups, with greater impact on those experiencing lower SES. This would suggest there was an interaction

between unhealthy lifestyle factors and lower SES to increase an individual's vulnerability to the detrimental effects of either unhealthy lifestyle factors, lower SES, or both.[172]

There are some studies that have found differential effects associated with single lifestyle factors. For example, alcohol-related adverse health outcomes were higher in lower, compared to higher, SES groups in studies in Finland, Denmark, and Scotland.[173-175] Similar results have been found for smoking[176-178] and physical inactivity.[163,176] Other studies have found more mixed evidence for differential vulnerability to a range of traditional unhealthy lifestyle factors.[179] However, on scoping the literature it appeared that very few studies had examined whether there was differential vulnerability to a combination of unhealthy lifestyle factors and, specifically, a wide combination of unhealthy lifestyle factors that include both traditional and newer lifestyle factors.

One key implication of differential vulnerability to unhealthy lifestyle factors is that there would be differential benefit from improving lifestyle factors in more deprived populations. This would strengthen the argument for targeting support for healthy living to these more vulnerable groups.[172] It was these ideas that motivated my prior work and, subsequently, this thesis where the aim has been to explore what the risks from multiple unhealthy lifestyle factors are for populations experiencing socioeconomic deprivation and what this means for policy, practice, and interventions to support them.[24]

Therefore, in addition to systematically reviewing the evidence for differential vulnerability to multiple unhealthy lifestyle factors (Chapter 4), analyses in Chapter 6 examine how deprivation (and other sociodemographic factors) affect the association between a wide combination of unhealthy lifestyle factors and adverse health outcomes.

2.7.3 Causal mechanisms and explanations of differential vulnerability

This final section of the wider literature review briefly discusses some of the potential mechanisms that could explain differential vulnerability.

2.7.3.1 Stress

Psychological stress is known to be more prevalent among those affected by socioeconomic deprivation.[180] Increased psychological stress is also associated with adverse health outcomes[181] and has been shown to affect the immune system and pathophysiological processes e.g., atherosclerotic plaque formation,[182] altered insulin metabolism,[183] and inflammation.[184] Therefore, it could be that the combination of both unhealthy lifestyle factors and psychosocial stress could explain disproportionately greater effects of unhealthy lifestyle factors in more deprived populations. However, it has been suggested that stress is unlikely to be an explanatory factor as stress itself is also associated with disproportionate harm in lower SES groups and therefore the differential vulnerability to stress also needs to be explained.[185]

2.7.3.2 Accelerated biological ageing

Older age is a demographic characteristic with among the strongest associations with adverse health outcomes. It has therefore been suggested that vulnerability to any harmful exposure, including unhealthy lifestyle factors, may be due accelerated ageing processes.[186,187] Indeed, measures of biological ageing and physiological decline, such as DNA methylation have consistently been shown to follow SES gradients.[188] A mechanism of accelerated biological ageing could also account for life course approaches as there may be time critical periods in which stressful events have profound impacts on the ageing process and therefore on increasing vulnerability (e.g., in utero or during childhood).[189,190]

2.7.3.3 Unmeasured differential exposure

Another explanation for any observed differential vulnerability to unhealthy lifestyle factors in more deprived populations could be due to two different ways in which differential exposure may be inadequately measured.

Firstly, differential vulnerability could appear spuriously in studies in which the level or severity of the unhealthy lifestyle factor is not fully measured and therefore the gradient of differential exposure is not fully captured. For example, studies of harmful alcohol consumption have found that the very highest alcohol intakes are in the most deprived groups but these levels of

intake may not be captured by common categorical measures where the cut-offs are set at lower levels of alcohol intake.[191] Similarly, in a cross-sectional study of 541 women in the US, while the prevalence of smokers increased with decreasing 'educational group' (47% in the highest education group versus 65% in the lowest education group) so too did the percentage of women who smoked more than 20 cigarettes a day (17% in the highest education group, compared with 40% in the lowest education group).[192] Therefore, studies examining the association between smoking and adverse health outcomes across SES strata could observe differential vulnerability if the smoking rate was not accounted for.

Secondly, differential exposure may be inadequately measured due to a lack of measurement of other unhealthy lifestyle factors. As mentioned above, there could be additive or synergistic interactions between unhealthy lifestyle factors. For example, smoking and high alcohol intake interact synergistically to increase the risk of head and neck cancers.[193] Studies which lack measurements for a range of impactful unhealthy lifestyle factors would fail to capture lifestyle factor interactions, which will be more prevalent among more deprived populations where unhealthy lifestyle factors cluster. This problem could be mitigated against by exploring differential vulnerability to a wider combination of unhealthy lifestyle factors. It was for this reason that a wide range of unhealthy lifestyle factors were included in analyses in Chapter 6.

2.8 Conclusion

In summary, the associations between unhealthy lifestyle factors, socioeconomic deprivation, and adverse health outcomes are clear. Extensive work in these areas include developing theory to guide research and interpret findings. However, gaps in the literature remain as few studies examine the associations between wide combinations of unhealthy lifestyle factors and adverse health outcomes, particularly combinations that include newer or emerging lifestyle factors such as social connection. There are also few studies that examine for SES differential effects from multiple unhealthy lifestyle factors. Further, it is recognised that there has been a failure to examine the contextual social details around unhealthy lifestyle factors. The unmet challenge of combining lifestyle factor epidemiology with sociology is cited as a reason for failing to bridge the

translational gap between behaviour change science and the development of effective interventions.[194] Additionally, the need to address wider social determinants in addition to more individualistic behavioural determinants is under recognised in policy and public discourse.[195] Therefore, studying the ways in which wide combinations of unhealthy lifestyle factors are viewed and conceptualised in the context of deprivation could lead to new insights that inform policy and intervention.

3 Methodology

3.1 Overview

Results chapters of this thesis (Chapters 4-7) are presented in the alternative journal format (i.e., as published or submitted papers) and the specific methods for each of those chapters are contained within the manuscripts. Therefore, this chapter provides an overview and background to the wider methodological considerations of this thesis. This chapter describes the overarching framework for the three work packages (WP) that comprise this mixed-methods programme of work, namely, a systematic review, two quantitative analyses, and a qualitative work package. A flow diagram showing the research questions (RQ) for this thesis, their corresponding WPs, and the sequence in which they were addressed are shown in Figure 3-1. This chapter also describes how each research question was formulated and why the corresponding methods were chosen to answer each question.

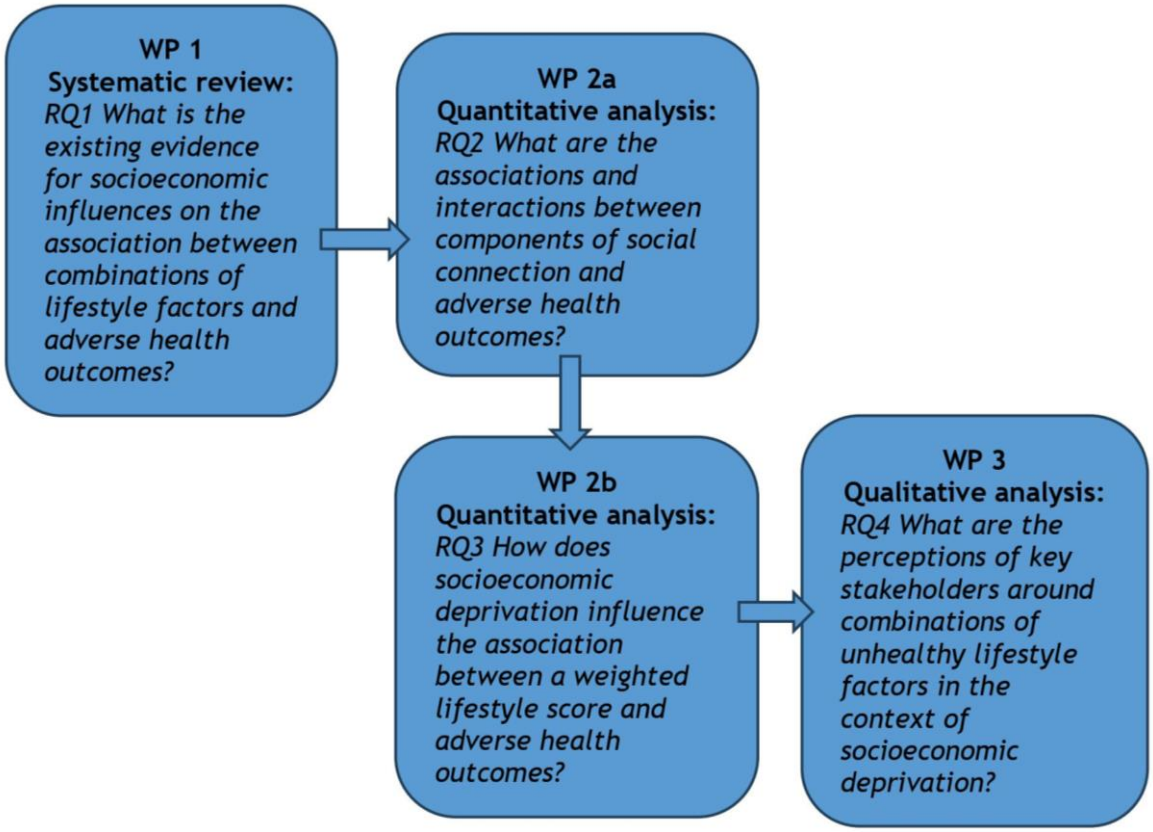


Figure 3-1: Overview of research questions (RQ) and work packages (WP)

Chapter 4 (WP1) reports the systematic review which followed standard methods for a systematic review (Preferred Reporting in Systematic Reviews and Meta-analyses [PRISMA]). Chapters 5 and 6 (WP2a and WP2b) each comprise a statistical analysis of the UK Biobank cohort study and contain details on the statistical methods employed. In this chapter, a wider overview of UK Biobank is given and its strengths and limitations are discussed. Further, the exposure variables (measures of social connection, lifestyle factors, and socioeconomic variables) that were selected and the linked outcome data are described in more detail. Chapter 7 contains the data and analysis for the qualitative work package (WP3). This methods chapter covers the theoretical considerations for integrating the qualitative and quantitative work packages.

3.2 Underlying framework – complex intervention development

From the first conception of this project, I wanted to contribute to our understanding of the risks of unhealthy lifestyle factors in more deprived populations in order that something could be done to reduce those risks. However, in scoping the literature there seemed to be little research around how the adverse health outcomes associated with unhealthy lifestyle factors were affected by deprivation. Rather, most research that examined the adverse health associated with both unhealthy lifestyle factors and deprivation did so with the aim of attempting to explain the higher rates of adverse health outcomes in more deprived groups by examining the proportions of adverse health outcomes that were attributable to the higher prevalence of unhealthy lifestyle factors in those groups. Indeed, there are numerous studies that have: 1) described the socioeconomic distribution of unhealthy lifestyle factors, 2) examined socioeconomic factors and their influence on the prevalence of lifestyle factors, and 3) estimated how lifestyle factors could explain SES gradients in adverse health outcomes. However, there appeared to be few studies that examined the adverse effects of unhealthy lifestyle factors in the context of socioeconomic deprivation. In other words, there was little consideration of the risk to health for those with unhealthy lifestyle factors who were also from deprived backgrounds.

With my background as a general practitioner working in more deprived areas, this felt like a missed opportunity to better understand the risks faced by patients like those I often see in clinical practice. It also felt like a missed opportunity to shine a light on those risks so that resources and interventions could be made proportionate to need and tailored to better support those living in more challenging socioeconomic circumstances.

Therefore, with an ultimate aim of intervening on the problem of unhealthy lifestyle factors in socioeconomically deprived contexts, the Medical Research Council's (MRC) framework for complex intervention development was used as an overarching guide for this thesis, including the development of the research questions and in choosing the appropriate methodologies.[196,197]

The framework for developing complex interventions was felt to be relevant because of the complexity that is likely to be inherent in addressing numerous unhealthy lifestyle factors in the heterogenous and often deeply challenging environment that characterises more socioeconomically deprived areas. As a simple example, many lifestyle factors are influenced by proximity or physical access to local resources (e.g., physical activity and distance to the nearest green or blue space).[198] Further, different communities may vary in their attitudes towards interventions, affecting intervention feasibility and implementation.[199,200] Therefore, interventions designed to support or address lifestyle factors need to be flexible and adaptable to locally available resources, geography, and many other community-level variations.[200,201]

The new and updated (2021) framework of complex intervention development and evaluation was felt to be most appropriate because, compared with preceding guidance,[196,202] it has greater emphasis on and explicitly acknowledges the critical nature of context in intervention development.[197]

The socioeconomic ('contextual') associations with unhealthy lifestyle factors are well recognised and evidenced and therefore interventions aiming to address lifestyle factors in deprived areas need to have contextual considerations at their core. The framework facilitates consideration of the resources required to support any intervention, which is critical in socioeconomically deprived contexts as these are characterised by a lack of resources.[203-205] Moreover,

existing evidence of interventions designed to support healthy living suggest they are less effective in populations experiencing deprivation compared with more affluent groups and are liable to increase health inequalities.[26]

3.2.1 Research perspective

The MRC framework considers four research perspectives from which to design and conduct complex intervention research: efficacy, effectiveness, theory based, and systems.[197] Of these four perspectives, this thesis primarily holds the theory-based perspective in the sense that it 'aims to understand how change is brought about, including the interplay of mechanisms and context'. From this research perspective, this thesis considers the challenges people face trying to live in healthy ways and make healthy behavioural changes across a wide range of lifestyle factors in the context of socioeconomic deprivation.

3.2.2 'Phases and core elements'

The MRC framework defines four phases of complex intervention research: development or identification of the intervention, feasibility, evaluation, and implementation.[197] This thesis generally focusses on the phase of developing or identification of a new intervention by considering the evidence of the problem (i.e., adverse health associated with unhealthy lifestyle factors in the context of socioeconomic deprivation). However, this thesis can also inform two of the other phases of complex intervention research, namely, feasibility, and implementation. For example, developing a weighted lifestyle score to examine whether and how its associations with adverse health outcomes are affected by socioeconomic deprivation (WP2b), can be considered an early phase of intervention development. Whereas the qualitative work package (WP3) examining public and professional views on lifestyle factors in the context of deprivation can provide insights into 'feasibility' and implementation of different lifestyle interventions.

The framework also suggests that all the phases of intervention development contain core elements ('considering context, developing and refining programme theory, engaging stakeholders, identifying key uncertainties, refining the intervention, and economic considerations') that run throughout the broader

complex intervention research process. Here, the core elements that were considered in this thesis are mapped to the relevant work package in Table 3-1 below. However, 'considering context' was a core element and focus throughout this thesis (i.e., examining the risks associated with combinations of unhealthy lifestyle factors whilst considering different socioeconomic contexts and considering the context of deprivation in particular).

Table 3-1: MRC complex intervention development core elements mapped to research questions and work packages.
Corresponding methods used are also given.

Core element	Research question (RQ)	Work package (WP)	Method used
Consider context	All	WP1-WP3	Mixed methods
Identify key uncertainties	RQ1 What is the existing evidence for socioeconomic influences on the association between combinations of lifestyle factors and adverse health outcomes?	WP1	Systematic review
	RQ2 What are the associations and interactions between components of social connection and adverse health outcomes?	WP2	Quantitative - prospective analyses of a population-based cohort
	RQ3 How does socioeconomic status influence the association between a weighted lifestyle score and adverse health outcomes?		
Engage stakeholders	RQ4 What are the perceptions of key stakeholders around combinations of unhealthy lifestyle factors in the context of socioeconomic deprivation?	WP3	Qualitative – focus groups and interviews

3.3 Why mixed methods

It is now well accepted that the different methodological paradigms of quantitative research and qualitative are mutually complementary and

combining them are beneficial when addressing multiple research questions, as in this thesis.[206-208] Simply put, there are limitations to both paradigms and using only one to inform policy and interventions will overlook important evidence from the other.[206] As stated previously, this thesis has a focus on the social context of unhealthy lifestyle factors, which, if only examined and considered through quantitative paradigms with more objective and measurable variables, could discount contextual details and perspectives of individuals and communities that are crucial to understanding the problem and informing intervention and policy.[206] Further, the complexity inherent both in wide combinations of unhealthy lifestyle factors and in the myriad and overlapping conceptualisations of socioeconomic deprivation means that certainty with 'numbers' could be severely limited in understanding people's experiences and interactions. This uncertainty or gap in understanding can be filled with detailed contextual and qualitative data of 'stories' or narratives that can help make sense of the complexities.[207] Mixed methods can therefore harness the combined power of both numbers and stories to address complexity.[209] Indeed, the MRC framework for developing and evaluating interventions supports the use of mixed methods, stating:

> 'For intervention research in healthcare and public health settings to take on more challenging evaluation questions, greater priority should be given to mixed methods, theory based, or systems evaluation that is sensitive to complexity and that emphasises implementation, context, and system fit. This approach could help improve understanding and identify important implications for decision makers, albeit with caveats, assumptions, and limitations. Rather than maintaining the established tendency to prioritise strong research designs that answer some questions with certainty but are unsuited to resolving many important evaluation questions, this more inclusive, deliberative process could place greater value on equivocal findings that nevertheless inform important decisions where evidence is sparse.'[197]

There are different ways in which quantitative and qualitative approaches can be combined, depending on the research questions being posed.[209,210] For example, the perspectives of participants from an intervention trial can be explored using qualitative methods to generate new and deeper understanding about the intervention and its impact on participants, which could then inform the subsequent steps in intervention development. Or findings from a qualitative study that stimulate or refine new hypotheses can then be tested quantitatively.

These are examples of sequential mixed method designs where methodological integration occurs in sequence i.e., a qualitative study is used to interpret or explain findings from a preceding quantitative study or, conversely, a quantitative study is used to test qualitative findings.[209] Mixed methods can also be convergent in design where integration occurs during different stages of the research process (e.g., during data collection or analysis, or both).[209]

The different methodologies in this thesis followed a mainly sequential design with WP1 (systematic review) informing the subsequent WP2 (quantitative studies). WP1 and WP2 then both informed WP3 (qualitative study). However, there was also necessary convergence in the design because, to permit timeous completion, the recruitment process, data collection, and initial analyses for the qualitative study was ongoing during quantitative analyses. Nevertheless, the final set of analyses to be completed were qualitative and informed by the preceding quantitative studies.

The following sections describe why using different methods (systematic review, quantitative analyses, and qualitative methods) was important and enabled the RQs to be addressed.

3.4 Overview of the systematic review work package (WP1)

A central question that has driven much of this thesis has been - does socioeconomic deprivation infer a vulnerability to the negative effects of unhealthy lifestyle factors? For me, this question was borne out of a previous study I led, which found that the association between a combination of nine unhealthy lifestyle factors and mortality was stronger in more deprived groups.[24] This disproportionately stronger association between lifestyle and adverse health in more deprived groups implied there could be a synergy between an unhealthy lifestyle and deprivation. And therefore, over and above the well-recognised higher prevalence of unhealthy lifestyle factors in more deprived populations, it appeared that socioeconomically deprived populations were more susceptible or more vulnerable to the effects of a wide combination of unhealthy lifestyle factors.

The reason why a synergistic interaction between combinations of unhealthy lifestyle factors and deprivation would be important has been described in the previous chapter but, briefly, evidence of synergy would strengthen arguments that policies and interventions should consider both unhealthy lifestyle factors and deprivation together and to do so explicitly.

On scoping the literature in this area it appeared that there were few studies in this area that examined single unhealthy lifestyle factors and even fewer that examined wide combinations of unhealthy lifestyle factors.[211] Further, some studies observed mixed results where none of smoking, alcohol consumption, physical activity levels and BMI had differential effects across SES tertiles.[179] Consequently, systematically reviewing the existing evidence for whether and how socioeconomic status influences the association between combinations of unhealthy lifestyle factors and adverse health outcomes became one of the primary objectives for this thesis. RQ1 was, therefore, formulated as - What is the existing evidence for socioeconomic influences on the association between combinations of lifestyle factors and adverse health outcomes?

Because I expected to find few studies on this topic, I aimed to have broad inclusion criteria. Further, I also expected wide heterogeneity in terms of lifestyle factors and SES measures examined as well as heterogeneity in analytical methods. I therefore took a pragmatic approach to data synthesis, planning to conduct a meta-analysis if appropriate. If not appropriate, I planned to use systematic review without meta-analysis (SWiM) guidelines to structure the synthesis of included studies.[32] SWiM guidelines have an overarching aim of enhancing the transparency of quantitative syntheses. The main items in the guidelines encourage researchers to report:

> ‘how studies are grouped, the standardised metric used for the synthesis, the synthesis method, how data are presented, a summary of the synthesis findings, and limitations of the synthesis.’[32]

3.5 Overview of the quantitative work package (WP2) - UK Biobank data

This section firstly discusses the motivation for quantitative analyses of prospective cohort data. It then describes UK Biobank participant recruitment and participants. It also presents key points of the study protocol including details of the initial baseline assessment at which point the data for the exposure variables examined in this thesis were collected. It provides a description of the baseline variables included in analyses and, where relevant, a brief description of additionally available variables and an explanation for why they were not included in analyses reported here. This section then describes the data linkage and the national registries which provide the outcome variables used in analyses. Finally, a discussion of the strengths and limitations of UK Biobank are presented. Some of this section is partly included in the relevant published and submitted results chapters (5 & 6), and those chapters also contain additional methodological detail that relate to their specific analyses. However, this section aims to provide an overview of the UK Biobank cohort and its procedures for data collection and linkage.

3.5.1 Why quantitative analyses?

The gaps in the epidemiological literature described in Chapter 2 provide the reasons why analysis of large datasets would address both RQ2 (What are the associations and interactions between components of social connection and adverse health outcomes?) and RQ3 (How does socioeconomic status influence the association between a weighted lifestyle score and adverse health outcomes?).

For RQ2 and social connection analyses (Chapter 5) the main literature gaps were:

1) while the exposure variables of social isolation and loneliness are often only moderately correlated and thought to exert independent effects on adverse health outcomes, our best estimates from meta-analyses are informed by studies that rarely mutually adjust for each.

2) studies rarely examine both structural and functional components of social connection in the same dataset and therefore the combined effect of both is under-explored.

For RQ3 and the quantitative analyses examining a wide combination of lifestyle factors (Chapter 6) the main literature gaps were:

1) studies of lifestyle scores (which offer a way of examining the associations between wide combinations of unhealthy lifestyle factors and adverse health outcomes) have typically focussed on traditional lifestyle factors and do not include emerging lifestyle factors such as measures of social connection.

2) the same studies rarely weight each lifestyle factor in the scores to account for the different effect sizes of each unhealthy lifestyle factor.

3) studies of lifestyle scores and associated adverse health outcomes rarely examine for potential SES-based differential vulnerability.

These gaps in the epidemiological literature around the adverse health outcomes associated with certain key exposures (different measures of social connection and unhealthy lifestyle factors and lower SES, respectively) could be addressed by new analyses to assess the strength of associations between exposures and outcomes across specific SES subgroups of the population. The aim was to help in identifying new at-risk groups and thus inform corresponding interventions and policies.

3.5.2 Why use UK Biobank?

The UK Biobank is a population-based biomedical prospective cohort. UK Biobank was chosen to answer these RQ2 and RQ3 because:

1) the wide variety of data collected at baseline would permit examination of an extensive range of exposure variables (i.e., a range of measures of social connection, lifestyle factors, and SES)

2) the large number of research participants in UK Biobank (n = 503,325) meant the statistical power to detect associations would be more likely even when examining associations between adverse health outcomes and smaller subgroups of participants (i.e., participants from different strata of deprivation)

3) the linked outcome data contained within national registries would allow assessment of a range of adverse health outcomes (e.g., all-cause and CVD mortality).

3.5.3 UK Biobank recruitment and study procedures

UK Biobank had a stated aim of being inclusive and, between the years 2006-2010, all adults aged 40-69 years who were registered with the National Health Service (NHS) and living up to 25 miles from one of the 22 study assessment centres were invited to participate.[212] UK Biobank assessment centres were located throughout England, Scotland, and Wales but none were in Northern Ireland. Just over 9 million postal invitations were sent and 503,325 participants were recruited, giving a response rate of around 5.5%.[212]

3.5.4 UK Biobank ethical approval

UK Biobank gained ethical approval from the North West Multi-centre Research Ethics Committee (16/NW/0274). All participants provided written and informed consent for their data to be collected, analysed, and linked to NHS records. The UK Biobank analyses presented here were performed under UK Biobank project application number 14151. Use of UK Biobank data for project 14151 is contingent on a material transfer agreement between the responsible research group (General Practice and Primary Care, University of Glasgow) and UK Biobank.

3.5.5 UK Biobank baseline assessment

Participants completed a touchscreen questionnaire that asked for information around their demographics, lifestyle, living and working environment, and medical history.[213,214] Following the questionnaire, a nurse-led interview was completed which was designed to ascertain further details about participants' country of birth, occupation, and medical history.[215] The interview involved the

nurse reviewing the responses to the touchscreen questionnaire about medical and medication history with the participant in an attempt to improve the accuracy of self-reporting. Participants also provided a range of physical measurements (e.g., height, weight, and blood pressure) and provided biological samples (e.g., blood, urine, and saliva).

Subsamples of the cohort also completed additional questionnaires to extract extra detail and or underwent a range of medical imaging modalities. Data from these additional questionnaires or imaging modalities were not used in this thesis. For example, objective physical activity measurements and additional dietary variables collected in subsamples were not included for analysis in this PhD as they were beyond the scope of this work. Further, it was decided to focus only on self-reported 'behavioural' lifestyle factors rather than objectively measured (e.g., accelerometer) factors or metabolic lifestyle factors (e.g., BMI or blood pressure), which require some form of physiological or anthropometric measurement. While equipment-based objective measures could improve accuracy of exposure assessment they can also act as a barrier to research or interventions, particularly for socioeconomically deprived populations.

To mitigate against regression dilution bias, where measurement error in exposure variables attenuates the associations between exposures and outcomes, subsets of the cohort were invited for repeated baseline assessment.[216] Repeat baseline assessment of 20,000 participants was performed between August 2012 and June 2013 at the UK Biobank co-ordinating centre in Stockport, England.[217] Use of this repeat baseline assessment data was also beyond the scope of this work but provides future opportunities to examine associations between trends or trajectories in social connection measures, lifestyle factors, SES, and adverse health outcomes.

3.5.6 UK Biobank exposure variables

UK Biobank collected a range of lifestyle factors, social connection measures, and socioeconomic variables at baseline. The variables used in this thesis and comparisons with relevant previous work are described below.

3.5.6.1 Social connection measures

There is no consensus on how to conceptualise, define, and, therefore, measure different types of loneliness, social isolation, or social connection and each measure has advantages and disadvantages. This has led to a plethora of measures that aim to capture different components, types, and aspects of how humans interact with inconsistent and overlapping terms and language.[94]

For example, a subjective feeling of loneliness or a perception of social disconnection has been classified into separate types of emotional and social loneliness by Weiss.[218] Maes et al. define emotional loneliness as a feeling of lacking close emotional attachments, affection, or intimacy and define social loneliness as perceiving a lack of people to talk to or from whom to ask for help.[94] As a result of this lack of consensus and complexity in conceptualising different types of loneliness, there are numerous ways in which the subjective feeling of loneliness has been measured in population studies.[101] Interestingly, the most commonly used loneliness scales or questionnaires do not use the words lonely or loneliness (to avoid socially desirable responses) and contain items that relate to psychometric properties or personality traits (e.g., questions around introversion/extraversion) and therefore could be considered indirect measures of loneliness.[219,220]

The social connection measures included in analyses here are described in Chapter 5. To aid discussion around these measures, the UK Biobank survey items that elicited the data for these social connection measures are shown again in Table 3-2. These measures have been used in several previous UK Biobank studies.[221-223] However, unlike those previous studies, they have been considered here using the conceptual lens of a social connection framework detailed by Holt-Lunstad.[18,87] This conceptual framework was described in Chapter 2 but, briefly, it comprises functional, structural, and quality components. This framework has several advantages over other similar concepts (e.g., loneliness and isolation) as discussed in Chapter 5. However, the main advantage is that social connection encompasses both subjective measures (e.g., loneliness) and objective measures (e.g., social isolation) and therefore assists in the interpretation of the independent and combined associations between these types of social connection and adverse health outcomes.

Table 3-2: UK Biobank survey item used as measures of functional and structural components of social connection.

Component	UK Biobank survey item
Functional	How often are you able to confide in someone close to you?
	Do you often feel lonely?
Structural	How often do you visit friends or family or have them visit you?
	Which of the following do you attend once a week or more often? Sports club or gym, Pub or social club, Religious group, Adult education class, Other group activity
	Including yourself, how many people are living together in your household?

Fewer questionnaire items are available to measure components of social connection in UK Biobank than in more comprehensive and specific survey measures (e.g., the Revised University of California, Los Angeles Loneliness Scale (R-UCLA)). Therefore the measures available in UK Biobank may lack detail and nuance around the subjective experience of loneliness or the objective state of isolation that may be important in the association with adverse health outcomes.[219] However, there is often high correlation between shorter item and longer item scales for both loneliness and social isolation. For example, a three-item loneliness scale had a corelation coefficient of 0.82 with the R-UCLA, the most commonly used loneliness scale in population surveys, which has 20 items.[224] Moreover, a longer scale for extracting more nuanced details around the components of social connection would have been prohibitively long in the context of the UK Biobank questionnaire and baseline assessment.

As stated previously, the measures employed here have been used in several previous UK Biobank studies.[221-223] The first of these studies used the measures in Table 3-2 to define loneliness and social isolation scales and subsequent UK Biobank studies have emulated their scales.[221] In that first study, Elovainio et al. suggest that the questions around loneliness (functional component) they use are similar to those in the R-UCLA. Indeed there are items (with Likert scale response options) in the R-UCLA that are similar to the UK Biobank equivalent question relating to being able to confide (e.g., ‘There is no one I can turn to’, ‘There are people I can talk to’, and ‘There are people I can turn to’) and others that could be considered similar to the UK Biobank question of feeling lonely (e.g., ‘I lack companionship’, ‘I do not feel alone’, and ‘I am no longer close to anyone’) but, as stated previously, there are no items in R-UCLA that specifically

mention the words ‘loneliness’ or ‘lonely’.[219] Therefore, the UK Biobank items around the functional component of social connection likely measure slightly different things to well-recognised loneliness scales and further work to examine their validity would be helpful. Elovainio et al. and subsequent UK Biobank studies using the same ‘loneliness scale’ have defined loneliness as scoring negatively to both items i.e., participants had to report both a reduced ability to confide in someone close and that they often feel lonely to be coded as lonely.[221-223] However, given that each question may measure different aspects of the functional component of social connection (e.g., a perception of lacking emotional support versus a perception of loneliness) means that the independent associations of each component question have been overlooked.

Similarly, Elovainio et al., in deriving their social isolation scale comprising the three questions relating to the structural component of social connection in Table 3-2, suggest that their scale is similar to those previously used in UK studies and cite a Whitehall II study.[221,225] This examined measures of social support that were very similar to those in UK Biobank and comprised a social network score based on the frequency of contact with relatives, friends, and colleagues, the frequency of participating in social or religious activities, and the total number of relatives or friends seen at least monthly. However, the Whitehall study also included a dichotomous marital status variable (1=married/cohabiting; 0=never married, separated, divorced or widowed) as a structural measure of social support. However, marital status is not directly available in UK Biobank. All UK Biobank participants (except those who reported living in sheltered accommodation or a care home as well as those who indicated they lived alone) were asked how they were related to the people they lived with and 446,370 (89%) participants reported living with a ‘husband, wife, or partner’. However, this omits data on whether individuals were married, separated, divorced, or widowed thus precluding creation of a comparable marital status variable. For their social isolation scale, Elovainio et al. classified participants as isolated where they scored negatively for at least two from three of the questions around structural component of social connection.[221] However, it is not clear why the threshold of 2/3 items was chosen to classify participants as isolated. Again, as with the loneliness scale, combining the separate questions into a scale meant that the independent associations of each component

question were not examined. This highlights how previous UK Biobank work examining these measures may have overlooked the relative importance of each individual component.

Other closely related psychosocial variables are available in UK Biobank but were not assessed here (Table 3-3). These variables were collected as part of a series of questions which also contained the survey item concerning loneliness. The relevant section of the UK Biobank baseline touchscreen questionnaire was introduced with the following:

> 'Now some questions about your feelings and your mood. Work quickly and do not think about the exact meaning of the question.'

This suggests the study designers considered these items to be closely related. Table 3-3 shows these variables in the order in which they appeared in the touchscreen questionnaire. These variables were not included as they are more readily interpreted as psychosocial features of personality traits or symptoms that could represent the adverse health outcomes of depression and anxiety.

Table 3-3: Related psychosocial variables available in UK Biobank

UK Biobank variable name	Survey item
Mood swings	Does your mood often go up and down?
Miserableness	Do you ever feel 'just miserable' for no reason?
Irritability	Are you an irritable person?
Sensitivity/hurt feelings	Are your feelings easily hurt?
Fed-up feelings	Do you often feel 'fed-up'?
Nervous feelings	Would you call yourself a nervous person?
Worrier/anxious feelings	Are you a worrier?
Tense/'highly strung'	Would you call yourself tense or 'highly strung'?
Worry too long after embarrassment	Do you worry too long after an embarrassing experience?
Suffer from 'nerves'	Do you suffer from 'nerves'?
Loneliness, isolation	Do you often feel lonely?
Guilty feelings	Are you often troubled by feelings of guilt?
Risk taking	Would you describe yourself as someone who takes risks?

Six 'happiness and satisfaction' variables that are closely linked to social connection are also available but were not included in analyses in Chapter 5 (Table 3-4). These survey items were not included in the analyses in this thesis

because they are only available for around 46% (n= 231,287) of all UK Biobank participants as these additional survey items were only introduced part way through fieldwork in April 2009. Future work could examine these family relationship and friendship satisfaction variables as measures of the different quality components of social connection.[96]

Table 3-4: Happiness and satisfaction variables available in UK Biobank

Variable name	**Survey item** (In general, how…
Happiness	… happy are you?
Work/job satisfaction	… satisfied are you with the work that you do?
Health satisfaction	… satisfied are you with your health?
Family relationship satisfaction	… satisfied are you with your family relationships?
Friendships satisfaction	… satisfied are you with your friendships?
Financial situation satisfaction	… satisfied are you with your financial situation?

In summary, there are myriad ways in which to measure different components of social connection with advantages and disadvantages to each. The variables used in this thesis are those that are available for all UK Biobank participants and encapsulate functional and structural components of social connection. These are likely to have reasonable correlation with similar survey measures but are more limited than more comprehensive measures. There are no measures of the quality components of social connection for all UK Biobank participants but future work could examine the relationship satisfaction variables that are available for a subset of participants.

3.5.6.2 Lifestyle factors

Prior to this work I published an analysis of UK Biobank that examined the association between an extended lifestyle score, socioeconomic status, and adverse health outcomes.[24] That lifestyle score comprised nine lifestyle factors and their selection was based on previous analysis of a comparable cohort, the '45 and Up Study', a large (n = 267,079) Australian cohort of middle-aged and older adults.[7] Table 3-5 shows the lifestyle factors and the corresponding survey items used to collect the lifestyle data for the prior studies and for the weighted lifestyle score in Chapter 6. The available response options for these items and corresponding categorisation are described in Chapter 6.

Table 3-5: UK Biobank lifestyle survey items used for weighted lifestyle score compared with previous studies

	Measurement/survey items by study		
Lifestyle factor	**Ding et al. 2015[7]**	**Foster et al. 2018[24]**	**Weighted lifestyle score (Chapter 6)**
Smoking	Have you ever been a regular smoker? AND Are you a regular smoker now?	Do you smoke tobacco now? AND In the past, how often have you smoked tobacco? (2 separate questions)	
Alcohol	About how many alcoholic drinks do you have each week? (1 drink defined as 1 glass of wine, 1 half pint of beer, or 1 shot of spirits).	About how often do you drink alcohol?	In an average week, how many … would you drink?* 1. glasses of red wine 2. glasses of white wine /champagne 3. pints of beer or cider 4. measures of spirits or liqueurs 5. glasses fortified wine (separate questions for each type of drink)
Physical activity	In the last week, how many times did you … [for 10 minutes or more]? 1. walk 2. do any vigorous gardening or heavy work around the yard 3. do any vigorous physical activity [excluding chores/gardening] 4. any other more moderate physical activities [not yet mentioned] AND	In a typical week, on how many days did you … [for 10 minutes or more] ? 1. walk 2. do moderate physical activities 3. vigorous physical activities AND How many minutes did you usually spend … on a typical day? 1. walking 2. doing moderate activities 3. doing vigorous activities (6 questions in total)	

	What do you estimate was the total time that you spent … in the last week? 1. walking 2. doing any vigorous gardening or heavy work around the yard 3. doing vigorous activities 4. doing other moderate activities		
Sedentary time	About how many hours in each 24 hour day do you usually spend doing the following? – sitting	In a typical day, how many hours do you spend watching TV?	
Sleep duration	About how many hours in each 24 hour day do you usually spend doing the following? – sleeping	About how many hours sleep do you get in every 24 hours?	
Diet	Dietary index based on 5 food items (vegetable, fruit, fish, processed meat, and types of milk) based on the 2013 Australian Dietary Guidelines.*	-	
Fruit and vegetable intake	-	About how many of …. would you eat per day? Separate questions for pieces of fresh and dried fruit, tablespoons of salad or cooked/raw vegetables.	
Oily fish intake	-	How often do you eat oily fish? (e.g. sardines, salmon, mackerel, herring)	
Red meat intake	-	How often do you eat…? Separate questions for Beef, lamb or mutton, and pork.	
Processed meat intake	-	How often do you eat processed meats (such as bacon, ham, sausages, meat pies, kebabs, burgers, chicken nuggets)?	
Added salt	-	-	Do you add salt to your food? (Do not include salt used in cooking)

Social participation	-	-	How often do you visit friends or family or have them visit you? **And** Which of the following [leisure/social activities] do you engage in once a week or more often?

***** Additional information provided on the questionnaire: red wine, - six glasses in an average bottle; pints of beer or cider (bitter, lager, stout, ale, Guinness); spirits or liqueurs (whisky, gin, rum, vodka, brandy) - 25 standard measures in a normal sized bottle; fortified wine (sherry, port, vermouth) - 12 glasses in an average bottle). This allowed estimation of units of alcohol per week using the following conversion factors: glasses of red or white wine/champagne * 2.1, pints of beer/cider * 2.5, measures of spirits or liqueurs * 1.2, glasses of fortified wine *1.[226,227]

There are several lifestyle factors variables that are available in UK Biobank that could have been included in analyses here (Table 3-6). Exposure to green space is becoming established as an emerging lifestyle factor but it is not clear how this relates to time spent outdoors, which was an available exposure variable.[228] For example, previous UK Biobank analysis found longer time spent outdoors (>3.5 hours/day) was associated with higher risk of CVD outcomes.[229] The authors suggest this could be explained by negative urban-related factors as most UK Biobank participants were resident in urban areas. Therefore, time spent outdoors was not included.

Table 3-6: Additional lifestyle factor variables available in UK Biobank

Lifestyle factor	Survey item
Time outdoors	In a typical day in summer, how many hours do you spend outdoors?
	In a typical day in winter, how many hours do you spend outdoors?
Sedentary behaviour	In a typical DAY, how many hours do you spend using the computer? (excluding using a computer at work)
	In a typical DAY, how many hours do you spend driving?
Dietary factor	How often do you eat cheese?
	What type of milk do you mainly use?
	What type of spread do you mainly use?
	What type of bread do you mainly eat?

	How many bowls of cereal do you eat a week?
	What type of cereal do you mainly eat?
	How many cups of tea do you drink each day?
	How many cups of coffee do you drink each day?
	What type of coffee do you usually drink?
	How do you like your hot drinks? (Such as coffee or tea)
	How many glasses of water do you drink each day?
	Have you made any major changes to your diet in the last 5 years?
	Does your diet vary much from week to week?
Sexual risk behaviours	What was your age when you first had sexual intercourse?
	About how many sexual partners have you had in your lifetime?
	Have you ever had sexual intercourse with someone of the same sex?
	How many sexual partners of the same sex have you had in your lifetime?

Other sedentary behaviours beyond TV viewing time (i.e., duration of leisure-based computer use and driving) were not included as these sedentary behaviours tend not to have as strong associations with adverse health outcomes as TV viewing. Whereas TV time was included as this sedentary behaviour tends to predominate in UK, both in terms of overall time and in strength of associations with adverse health outcomes and was based on prior work I did.[230,231] The dietary factors listed in Table 3-6 might feature as part of composite dietary indices or measures (e.g., 'How often do you eat cheese', 'What type of milk do you mainly use?, and 'What type of spread do you mainly use?' could comprise a dairy intake index) which could be a proxy for or form part of a poor diet and therefore be important to include in analyses. However, on their own, there is markedly less evidence for their independent associations with adverse health outcomes compared with the dietary factors that were included in analyses (i.e., frequency of intake of red and processed meat, fruit and vegetables, oily fish, and added salt). In other words, whether the dietary factors listed in Table 3-6 (frequency of cheese consumption, type of milk, type of spread, type of bread, number of bowls of cereal, cups of tea/day cups of coffee/day, type of coffee, hot drink temperature preference, glasses of water per day etc.) are independent dietary risk factors is not well recognised and,

therefore, these factors were not included in analyses. The sexual risk behaviours were not included because it was felt they did not fit the definition of lifestyle factor used in this thesis. However, these somewhat subjective decisions for including certain lifestyle factor variables or not are open to criticism and future work could incorporate some of these additional factors as evidence changes.

There are also variables not shown in Table 3-6 that could also be defined as lifestyle factors but are less commonly defined as such (e.g., questions on driving over the speed limit [risk taking behaviour], and frequency of use of a mobile phone to make calls) and so were excluded here. Additionally, there were survey items related to sleep apnoea (e.g., likelihood of dozing during the day or snoring) and, although these would be correlated with other sleep variables (e.g., sleep duration), they do not assess sleep duration directly and so were not included.

Concerning alcohol consumption, there were also several variables not included in analyses here (e.g., When you drink alcohol is it usually with meals?, About how often do you drink alcohol?). Prior UK Biobank work has examined these various dimensions of alcohol consumption and found a complex pattern of associations with adverse health outcomes.[232] Therefore, for simplicity, and given there is no safe level of alcohol consumption (i.e., no level at which any potential protective cardiovascular effect of alcohol outweighs the associated cancer risks) the choice was made to use only the alcohol related survey items in Table 3-5, which allowed calculation of weekly alcohol units consumed.[233]

The wider evidence for the associations between each of the lifestyle factor variables chosen for inclusion in this thesis and adverse health outcomes are briefly given in the prior work on which their selection is based and in Chapter 6.[7,24,230]

3.5.7 UK Biobank outcome variables

All participants consented for their data to be linked to their NHS records and to national mortality and morbidity registers.[234] Different registries provide data to UK Biobank for different adverse health outcomes. Data on deaths are sent to UK

Biobank from NHS England for participants based in England and Wales and from the NHS Central Register for participants based in Scotland. This data includes the date of death and the International Classification of Disease version 10 (ICD-10) codes for the primary and contributory causes of death given on death certificates.[235]

The adverse health outcomes assessed in UK Biobank analyses in this thesis were all-cause mortality and CVD mortality. These two outcomes were selected as they are known to have strong associations with social connection, lifestyle, and socioeconomic status.[24,88,236] Further, mortality outcomes are considered 'hard' outcomes and while there is possibility of misclassification for the cause of death there is likely to be much less error in reporting a death event.

CVD mortality was defined as any death with ICD-10 codes I05–I99, Z86.7, G45, and G46 given as the primary cause of death. These ICD-10 codes were selected to define CVD deaths by way of review and discussion of the ICD-10 manual by two primary care clinicians (myself and one of my supervisors, FSM). It was felt these codes likely represent chronic CVD diseases and, from ICD-10 section 'IX - Diseases of the circulatory system', only acute rheumatic fever (ICD-10 codes I00–I02) was excluded. Chronic diseases are more relevant here due to the likely chronic nature of pathophysiological mechanisms that link social connection, lifestyle, or socioeconomic status to CVD.

Dates and causes of hospital admissions are also provided to UK Biobank by Health Episode Statistics for participants in England and Wales and by the Scottish Morbidity Records for participants in Scotland.[237] Further, primary care data is also provided by a range of primary care data service providers for around 45% of all participants (n = approximately 230,000).[238] The primary care data includes information on diagnostic codes, symptoms, laboratory results, prescription data, and administrative codes (e.g., referrals to secondary care). Use of all these data were outside the scope of the analyses in Chapters 5 and 6 due to the high number of analyses already required for the two mortality outcomes. However, these additional data provide future opportunities to examine the associations between social connection, lifestyle, socioeconomic factors, and other adverse health outcomes. For example, in previous UK

Biobank work, I examined the associations between lifestyle, socioeconomic status, and CVD incidence.[24]

3.5.8 Statistical thinking and results presentation

The details of the analytical methodology for each UK Biobank analyses are provided within each chapter (Chapters 5 and 6). However, here I provide further details on the 'statistical thinking' in how the results of the analyses are presented.

The main results from the quantitative analyses of UK Biobank in this thesis rely on the outputs of Cox Proportional hazard models. As well as calculating a hazard ratio to estimate the association between exposure variables and outcomes, outputs from Cox models often also include a hypothesis testing statistic i.e., a p-value. P-values are derived from Z-values, which are calculated by dividing the estimated coefficients by their standard errors.[47] Z-values follow a normal distribution under the null hypothesis, which states that the true coefficient is 0 (or that the true hazard ratio is 1 i.e., that the exposure variable is not associated with the outcome). A p-value therefore represents an estimate of the probability of observing a Z-value as extreme as the one calculated assuming that the null hypothesis is true. An alternative definition for p-values is given by Greenland et al., 'The p-value is then the probability that the chosen test statistic would have been at least as large as its observed value if every model assumption were correct, including the test hypothesis.'[239]

P-values are a function of effect sizes and sample size.[240] Where a sample size is large then p-values can be small even when the effect size is small and of negligible clinical importance. Therefore, while p-values contain additional information regarding hypothesis testing (the probability of a Z-value as extreme as the one calculated assuming that the null hypothesis is true), the utility of providing p-values in the context of research based on large sample sizes such as UK Biobank when testing multiple relationships has been questioned.[240,241] Therefore, the results presented from Cox proportional hazard models include only the hazard ratios and confidence intervals in order to focus on the magnitude of associations and the uncertainty of these estimates.

3.5.9 UK Biobank strengths

The comprehensive data provided by participants at baseline provide a rich source of potential exposure variables. This permits the examination of a range of social connection measures, lifestyle factors, socioeconomic variables, and adverse health outcomes whilst controlling for several potential confounding variables (e.g., sex, age, ethnicity, blood pressure, and BMI). For social connection, this allows examination of the independent and combined associations across different components of social connection within the same dataset, which is lacking in previous research.[19] While for analyses of lifestyle factors, the rich data provides opportunities to examine a broad range of lifestyle factors.

The large sample size of UK Biobank increases the precision in the estimates (narrower confidence intervals) of the associations between exposures and outcomes. The linked outcome data provided by national mortality registers reduces the chance that participants are lost to follow up thereby minimising bias of estimates arising from differential loss to follow up.[242]

3.5.10 UK Biobank limitations

The UK Biobank response rate of 5.5% is low compared with other population-based cohorts. However, there has been a general trend of declining response rates to participate in epidemiological studies over time.[243] Therefore, considering the large scale and detailed level of data collection, these response rates may be as high as can be expected without dramatic increases in costs or new and innovative recruitment strategies.[243] Although a response rate of 5% is relatively low, UK Biobank study designers have argued, as others have done previously,[244] that only sufficiently large numbers of participants with varying levels of exposure variables are required to detect generalisable associations.[33,245,246] In other words, while the cohort was not representative and had a low response rate, prevalence estimates would remain inaccurate but important associations between exposures and health outcomes could still be identified so long as the absolute number of participants recruited was sufficient to include participants with a broad range of exposure levels.[245] For this reason, UK Biobank study designers opted not to recontact individuals who were

undecided whether to participate. This decision been given as a possible reason for the low response rate for UK Biobank but it has been cited as a cost saving measure that permitted the large scale and efficiency of UK Biobank.[245]

Although, the sample size is large, UK Biobank participants are not representative of the UK wider population as there are relatively few participants of black and minority ethnic origin and participants tend to be healthier and wealthier than the UK average.[247] This selection bias could result in collider bias and lead to the identification of spurious associations where none exist or influence the estimates of association.[248,249] Bias can be defined as 'a type of error that may affect the results of a study because of weaknesses in its design, analysis or reporting.'[250] When an outcome and an exposure independently cause a third variable, that variable is termed a 'collider'.[251] Controlling for a collider in statistical models results in collider bias, which can lead to spurious associations between exposure and outcome. Further there are varying levels of missingness in UK Biobank which can also affect estimates of associations. The percentage with missing data and how missing data were handled are given in the relevant manuscripts.

3.6 Overview of qualitative work package (WP3)

A qualitative method was chosen for exploring key stakeholder perspectives of lifestyle factors in the context of socioeconomic deprivation. Aspects of the methods for WP3 are contained in the Chapter 7 in the form of a paper submitted for publication. This chapter presents some of the key methodological considerations that pertain to qualitative research that are not included in that manuscript. This chapter first briefly describes the epistemological and ontological paradigms that contrast this work package (WP3) with those of the systematic review (WP1) and quantitative analyses (WP2). It then discusses the rationale for a qualitative method. Methodological and ethical considerations for the recruitment process and selection of key stakeholder participants are presented as well as additional detail around the process of data collection. It then presents some additional detail of the approach to analysis that is not in Chapter 7. Finally, a description is given of the approach to the final synthesis and discussion (Chapter 8) for this thesis that drew on findings across the mixed methods.

3.6.1 Epistemology and ontology

Epistemology is the study of knowledge - of what can be known and how it can be known.[252] Ontology is the study of existence or being and concerned with what exists or is real.[252] The epistemological position of social and health research can be crudely dichotomised as either positivist or interpretivist.[253] Positivism is the philosophical view that knowledge can be deduced and is objective and value free. Interpretivism maintains that knowledge can only be interpreted or induced from the perspective of individuals and that knowledge is therefore subjective and value laden. Different epistemologies exist between these two positions such as post-positivism, critical realism, and critical-rationalism.[254]

After considering different philosophical viewpoints by reading and reflecting on the literature cited in this thesis, my own epistemological and ontological viewpoint seems to align most closely with critical rationalism as this incorporates fallibilism, critical realism, and revisionism. My interpretation of critical rationalism is that there is an objective reality that includes socially constructed phenomena but humans are fallible in their attempts to understand reality and therefore our beliefs and assumptions require adaptation and 'improvements' following critique.[254,255]

This viewpoint seems to sit somewhere between positivism and interpretivism, which has allowed me to consider lifestyle factors and socioeconomic factors both as objective entities that impact on individuals and communities but also as socially constructed phenomena that influence and are borne out of peoples' interpretation of the world. This philosophical standpoint also represents a pragmatic one that can assist in synthesising data from both quantitative and qualitative research paradigms that are typically more positivist and interpretivist, respectively.

Epistemology and ontology are often discussed in qualitative research but less often discussed in quantitative research.[256] However, declaring the epistemological position or ontological assumptions are important for making coherent claims of qualitative data and analysis.[257] Further, understanding the underlying epistemology and ontology is perhaps more important for qualitative

than quantitative research as this can guard against biases. Whereas there are statistical and technical methods to mitigate biases in quantitative research.

However, considering underlying philosophical assumptions are important in both quantitative and qualitative research. For example, ontological assumptions held by researchers can determine how they define reality and influence the approach to research such as how research questions are formulated.[252] Ontological assumptions in relation to this thesis can be discussed in terms of whether the key exposures (e.g., lifestyle factors and socioeconomic factors) are viewed as objective entities that exist separately and outside people (an 'objectivist' view) or as social constructions that are created through the perceptions and actions of people (a 'constructivist' view).

With the aim of answering RQ4 (What are the perceptions of key stakeholders around combinations of lifestyle factors in the context of socioeconomic deprivation?) and therefore of understanding the perspectives of individuals and context a more subjective or interpretive approach was most appropriate. This informed the decision to employ qualitative research methodology for this work package.

3.6.2 Recruitment and participants

Details around specific methods of recruitment and the participants are given in Chapter 7. A broader discussion of the reasons behind the recruitment strategy are discussed here. Within my supervisory team, we discussed who we felt constitutes key stakeholders in relation to lifestyle factors in the context of socioeconomic deprivation. We had the pragmatic consideration of time and resources for the project as well as fulfilling criteria for methodological rigour (i.e., diversity in recruitment and views collected). Planning for this project also took place during social restrictions of the COVID-19 pandemic which influenced the decisions around method of data collection (e.g., remote versus face-to-face meetings with research participants). I wanted to gather as broad a collection of views as possible, including the public as well as relevant professionals.

For members of the public, I limited recruitment of participants to adults residing in Scotland. This was mainly a pragmatic decision as this would limit the

focus of recruitment to a Scottish pool of participants but also because responsibility for health and social care is devolved to Scotland and therefore views gathered could be better aligned for the Scottish context.

For views of professionals, I decided to focus on primary care professionals providing care for those living in more deprived areas, again aiming to recruit those based in Scotland. The aim for this WP was to understand key stakeholders' perspectives to fulfil a wider objective of informing intervention and policy around lifestyle in deprived contexts. As 56% of NHS contacts are undertaken in primary care and primary care is viewed as a trusted source of information and support within communities, future intervention and policy around lifestyle factors in deprived contexts will have important implications for, and likely involve, primary care.[258] Further, with the relatively new role of community links workers (CLW) embedded in primary care in Scotland, whose remit closely aligns with the focus of this thesis, namely, to support those in deprived contexts 'to live well', views of primary care links workers were thought to be critical.[259] Finally, my background as a general practitioner (GP) and working within a primary care research group meant that the wider primary care connections and resources available to me and my supervisory team would facilitate recruitment to this WP.

However, in addition to primary care professionals, I also wanted to gather views of professionals who focus on population health. These professionals were considered key stakeholders due to the wider aim of the research to inform policy and interventions that might operate at population-level. I considered their views critical to understanding the population and societal-level influences on both lifestyle factors and socioeconomic deprivation. Therefore, I aimed to gather views of those with day-to-day working experience of population level factors that directly concern or were closely related to lifestyle factors and socioeconomic deprivation (i.e., public health consultants and policy makers). Again, for pragmatic reasons, I focussed on those who fit these criteria and worked in Scotland.

3.6.3 Ethics

There was no NHS facility or data involved in this research and therefore ethical approval for this study was sought from the University of Glasgow's College of Medical Veterinary and Life Sciences Ethics Committee (application number 200210156). The application to the ethics committee along with copies of the participant information sheets are shown in Appendix 2. Consent forms for focus group and interview participants are given in Appendix 3 and Appendix 4, respectively.

There were minimal ethical considerations as the risks to participants were low. The primary risks were around protection of personal information and anonymity. It was therefore made clear in the participant information sheets and consent forms that participation in the study was fully voluntary, that any data provided would remain anonymous, and that participants could withdraw from the study at any time. However, it was also made clear that once participant views were collected, transcribed, anonymised, and analysed it would be impossible to remove their contributions at that stage. The pool of eligible professionals for this study is relatively small in Scotland and, therefore, an additional protection for anonymity was discussed with professionals – that any quotes used in subsequent reports could be checked with the relevant participant prior to publication.

Data is stored on encrypted university servers and only me and the immediate supervisory team (COD and FSM) have access to the data. Once all transcriptions and quotations for use were checked then all personal identifiers were deleted from the research data. All participants gave written and verbal consent. To assist in recruitment, participants were offered reimbursement for travel and a shopping voucher (public participants) or a small fee (professional participants) in line with NIHR guidance.[260]

3.6.4 Data collection

Data were collected via focus groups with members of the public and semi-structured interviews with professional participants. Topic guides for the focus groups and interview schedules for the semi-structured interviews were

developed iteratively with my supervisory team and were informed by their knowledge and interests. The research team included two academic GPs and a primary care scientist with a background in health inequalities. Development of the topic guides and schedules also drew on WP1 and WP2, wider literature review, as well as prior quantitative analysis examining the effect of socioeconomic status on the association between an extended lifestyle sore and adverse health outcomes.[24] The guides and schedules were also informed by discussion with the National Health Service Research Scotland Primary Care Patient and Public Involvement (NRS PPI) Group who were consulted as part of work for the funding application for this thesis. The topic guides and interview schedules are given in Appendix 5 and Appendix 6.

3.6.4.1 Data collection via focus groups with the public

Three focus groups were conducted online using either Zoom or MS Teams and were recorded using the facility available in the software. The audio recordings were transcribed using a professional transcription service. After initial review of the transcripts, it was felt that additional data could provide new interpretations when it was felt that perspectives from female and older participants were missing and therefore an additional focus group was arranged. At this stage of the project COVID-19 restrictions were reducing and therefore a face-to-face focus group was arranged. This focus group took place on university premises. Each focus groups lasted approximately 60-80 minutes.

3.6.4.2 Data collection via interviews with professionals

All interviews of professionals were conducted online via video conferencing software. Interviews lasted approximately 60 minutes and were recorded.

3.6.5 Analysis

Reflexive thematic analysis (RTA) was chosen as the method by which to analyse stakeholders' perspectives. Compared with other forms of thematic analysis, such as codebook approaches, RTA is a form of thematic analysis that facilitates the subjectivity inherent in creating knowledge.[261,262] Further the codes and themes are created via the researcher's reflexive engagement with the data. RTA openly acknowledges the researchers prior experience in influencing the

results and even views this as an advantage over other methods as opposed to a risk in terms of biasing results. The overall aim of RTA, therefore, has less to do with reproducibility (as two researchers using RTA are expected to analyse and interpret the same data differently on account of their different prior experience, skills, and resources) and more to do with generating rich interpretations of the data.[263] However, more than one researcher may work together in a reflexive and collaborative manner in order to achieve richer interpretations rather than aiming for consensus.[263] A full description of the analytical process is contained in Chapter 7.

3.7 Final discussion and synthesis

The final synthesis and contextualisation of all findings from this thesis within existing literature is contained in Chapter 8. While the integration of quantitative and qualitative methodologies occurred sequentially, the final synthesis attempts to provide an overview of the combined mixed methods findings.[264] The methods for this synthesis were largely informal and similar to what is described as a narrative summary.[264] However, with attempts to critically interpret findings across methodologies, the synthesis is also informed by concepts from critical interpretative syntheses, which, whilst designed for reviews of mixed methods was still informative here.[265,266]

As with all interpretation, while there are attempts to be transparent, the reflexive nature of the synthesis means it is not possible to be fully transparent even when all attempts to be transparent are made (e.g., unconscious bias can lead to the subtle emphasis of one result from one methodology over another from a different methodology). Methods to support transparency including presenting the results to the study advisory and PPI groups, presenting the work at conferences as well as discussing the findings and interpretation with my supervisory team. However, the obverse of this risk to reproducibility is flexibility, which facilitates the combining of disparate types of evidence. Further, rather than purely summarising or aggregating the data from this thesis, the synthesis aimed to generate new ideas and concepts that could inform policy and interventions that relate to multiple unhealthy lifestyle factors in the context of deprivation. Nevertheless, and stated here to increase transparency, the final synthesis should be read with the understanding that a level of

subjectivity and reflexivity has been applied to the selection and contextualisation of findings and to their interpretation.

3.8 Chapter summary

This chapter has presented the main methodological and theoretical considerations for this thesis. It contains the reasons behind the methodological choices and data selected for analysis and discusses the strength and weaknesses. This thesis uses mixed methodology, including a systematic review without meta-analysis (WP1), two quantitative analyses of a large cohort (WP2), and qualitative analysis of transcribed data from focus groups and interviews with members of the public and a range of relevant professionals.

The methods employed here have facilitated:

1. highlighting the gaps in the evidence around disproportionate harm from lifestyle factors in more deprived populations and synthesised the evidence that suggests an additive effect of deprivation on lifestyle associated harms.

2. examination of the independent and combined associations of the functional and structural components of social connection and all-cause and CVD mortality.

3. the creation of a novel weighted lifestyle score that incorporates measures of social connection and socioeconomic deprivation, which could be developed for clinical use.

4. analysis of the detailed perspectives of key stakeholders that will inform lifestyle intervention and policy development in areas affected by socioeconomic deprivation.

4 The influence of socioeconomic status on the association between unhealthy lifestyle factors and adverse health outcomes: a systematic review

4.1 Chapter summary

This chapter is formed of a systematic review of studies that examine the associations between combinations of lifestyle factors and adverse health outcomes. This addresses the first research objective – to describe and synthesise the evidence for the effect of socioeconomic deprivation on the association between combinations of unhealthy lifestyle factors and adverse health outcomes.

There is an additional sentence regarding the tool used to assess study quality, otherwise the text and figures in this chapter are as per: Foster HME, Polz P, Gill JMR et al. The influence of socioeconomic status on the association between unhealthy lifestyle factors and adverse health outcomes: a systematic review [version 2; peer review: 2 approved, 1 approved with reservations]. Wellcome Open Res 2023, 8:55 (https://doi.org/10.12688/wellcomeopenres.18708.2).

Associated Supporting information ('Extended data') is given in Appendix 7 and is also available to download via figshare: https://doi.org/10.6084/m9.figshare.24649755.v1.

The corresponding protocol for this review is given in (Appendix 1) as published: Foster H, Polz P, Mair F, et al. Understanding the influence of socioeconomic status on the association between combinations of lifestyle factors and adverse health outcomes: a systematic review protocol. BMJ Open 2021; 11:e042212. (https://doi.org/10.1136/bmjopen-2020-042212)

4.2 Abstract

Background: Combinations of lifestyle factors (LFs) and socioeconomic status (SES) are independently associated with cardiovascular disease (CVD), cancer, and mortality. Less advantaged SES groups may be disproportionately vulnerable to unhealthy LFs but interactions between LFs and SES remain poorly understood. This review aimed to synthesise the available evidence for whether and how SES modifies associations between combinations of LFs and adverse health outcomes.

Methods: Systematic review of studies that examine associations between combinations of >3 LFs (e.g., smoking/physical activity/diet) and health outcomes and report data on SES (e.g., income/education/poverty-index) influences on associations. Databases (PubMed/EMBASE/CINAHL), references, forward citations, and grey literature were searched from inception to December 2021. Eligibility criteria were analyses of prospective adult cohorts that examined all-cause mortality or CVD/cancer mortality/incidence.

Results: Six studies (n=42,467–399,537; 46.5–56.8 years old; 54.6–59.3% women) of five cohorts were included. All examined all-cause mortality; three assessed CVD/cancer outcomes. Four studies observed multiplicative interactions between LFs and SES, but in opposing directions. Two studies tested for additive interactions; interactions were observed in one cohort (UK Biobank) and not in another (National Health and Nutrition Examination Survey (NHANES)). All-cause mortality HRs (95% confidence intervals) for unhealthy LFs (versus healthy LFs) from the most advantaged SES groups ranged from 0.68 (0.32–1.45) to 4.17 (2.27–7.69). Equivalent estimates from the least advantaged ranged from 1.30 (1.13–1.50) to 4.00 (2.22–7.14). In 19 analyses (including sensitivity analyses) of joint associations between LFs, SES, and all-cause mortality, highest all-cause mortality was observed in the unhealthiest LF-least advantaged suggesting an additive effect.

Conclusions: Limited and heterogenous literature suggests that the influence of SES on associations between combinations of unhealthy LFs and adverse health could be additive but remains unclear. Additional prospective analyses would help clarify whether SES modifies associations between combinations of unhealthy LFs and health outcomes.

Registration: Protocol is registered with PROSPERO (CRD42020172588;25 June 2020).

4.3 Introduction

Unhealthy lifestyle factors (LFs) (*e.g.*, smoking, alcohol, poor diet, low physical activity (PA)) are key modifiable risk factors for non-communicable diseases (NCDs) and mortality[267]. While single LFs have, by themselves, strong associations with NCDs and mortality, combinations of unhealthy LFs have stronger associations. Meta-analyses show that, compared with healthy LFs, combinations of at least three unhealthy LFs are associated with more than twice the risk of all-cause, cardiovascular disease (CVD), and cancer mortality, and CVD incidence [72,73]. Examining adverse health outcomes associated with combinations of LFs can help to capture 'real life' risks more completely as unhealthy LFs tend to cluster together - individuals with one unhealthy LF often have more than one [15,62]. And the impacts of one unhealthy LF may interact (additively or multiplicatively) with other unhealthy LFs [268-270].

In addition to examining the associations between combinations of LFs and adverse health outcomes (*e.g.*, all-cause, CVD, and cancer mortality, and CVD incidence), examining the effect of socioeconomic status (SES) on those associations can deepen understanding of the distribution of these lifestyle-related adverse health outcomes among populations. As with most health outcomes, all-cause, CVD, and cancer mortality, and CVD incidence all follow clear and long-recognised SES-health gradients where individuals of less advantaged SES (*e.g.*, those with lower educational attainment, lower income, or who live in areas of higher deprivation) tend to have higher rates of both morbidity and mortality [271,272]. SES is a theoretical construct that differentiates sections of society by their means and access to resources (*e.g.*, financial, educational, material) and by the ways in which they live (*e.g.*, occupation type or class, housing type/conditions, neighbourhood/post code area) [272] . The broad scope that SES encompasses means 1) there are numerous ways in which SES can be operationalised or measured [130,132]; and 2) there are numerous aspects of SES that could be expected to influence and have strong associations with both LFs and lifestyle-related adverse health outcomes [22,133]. For example, there is higher prevalence of unhealthy LFs in less advantaged SES groups and clustering of multiple unhealthy LFs in such population groups is often cited as an explanation for observed lifestyle-

related adverse health inequalities [15]. However, 'differential exposure' to unhealthy LFs only partially explains lifestyle-related health inequalities; higher prevalence of unhealthy LFs is estimated to account for 6–80% of SES related mortality inequalities [15,20,21,163,164].

Beyond differential exposure, further explanations for lifestyle-related health inequalities may involve interactions between LFs and SES; so-called 'differential vulnerability' [163], where SES strengthens the association between lifestyle and adverse health outcomes. A study of over 300,000 UK Biobank (UKB) participants observed multiplicative interactions between a combination of unhealthy LFs and SES, where less advantaged SES groups had disproportionately higher lifestyle-related all-cause and CVD mortality [24]. Similar interactions between lifestyle and SES have been observed for single LFs: smoking, alcohol, and PA [175,176,273]. A multiplicative interaction between LFs and SES supports a vulnerability hypothesis, where less advantaged groups are disproportionately vulnerable to the adverse effects of unhealthy LFs [163,176]. Whereas additive interactions, where the effects of a combination of unhealthy LFs and SES are added rather than multiplied [274], can also highlight vulnerable groups and inform policy or interventions [172]. Mechanisms that explain differential lifestyle vulnerability or that explain how and why SES affects associations between lifestyle and adverse health outcomes are unclear but could include interactions with other factors associated with less advantaged SES (*e.g.*, stress, reduced access to health care) or accelerated biological ageing *via* greater cumulative risks over the life-course (*e.g.*, poorer childhood health or increased adverse childhood experiences) [185,186,275].

4.3.1 Aims

Understanding whether SES influences the association between combinations of unhealthy LFs and adverse health outcomes could help reduce excess risk in less advantaged populations by deepening understanding of how complex lifestyle risks vary across society and by identifying higher risk LF combinations. This could inform health policy, guide the development of interventions targeting more vulnerable groups, and support health care professionals managing multiple risk factors in their patient population. This

systematic review aims to identify, describe, and synthesise the evidence for whether SES modifies associations between combinations of unhealthy LFs and adverse health outcomes (all-cause mortality, incidence and mortality from CVD or cancer). This review addresses the following research questions: Does SES modify the association between combinations of unhealthy LFs and adverse health outcomes? And if so, how?

An important linguistic caveat: 'lifestyle' can imply choices made freely by individuals, leading to potential stigma. However, resource scarcity and the wider socioeconomic environment experienced by those in less advantaged SES groups clearly influences choices, for example, by making healthier choices less likely [276,277]. Moreover, lifestyle choices in the context of poverty or material deprivation may represent 'optimal' choices given wider socioeconomic influences that shape decision making and abstract future planning [278,279]. Nevertheless, the word lifestyle remains recognised in the context of modifiable behaviours and is therefore used here.

4.4 Methods

4.4.1 Search strategy and study selection

This review followed a protocol and was conducted in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [280-282]. The protocol is registered with a database of prospectively registered systematic reviews (PROSPERO CRD42020172588; 25 June 2020) [31,283].

Search strategies were developed with a specialist university librarian and adapted for three databases: PubMed (RRID:SCR_004846), EMBASE (RRID:SCR_001650), and EBSCO CINAHL (RRID:SCR_022707) (S1-3 Tables, Appendix 7) [284] . The search strategy of a previous systematic review of combinations of LFs and type 2 diabetes served as a template and was adapted to include SES related terms [71] . As per that previous review, this current review focusses on combinations of LFs, and therefore search terms relating to LFs included general terms like 'lifestyle' or 'health behaviour' rather than terms for individual LFs like 'smoking' or 'alcohol'. Search terms also included terms for combinations of LFs (*e.g.*, 'combined', 'multiple',

'score'). Searches from inception (PubMed-1966; EMBASE-1947; CINAHL-1984) to 17th December 2021 were supplemented by searches of references, forward citations, and grey literature [31] .

4.4.2 Eligibility criteria and screening

Inclusion criteria were developed using an adapted PICOS (population, intervention, comparator, outcome, study design) framework, with 'I' (intervention) replaced with 'E' (exposure) [285] . Inclusion criteria:

1) Population: any general adult population (age ≥18 years). Studies of participants with an index condition were excluded.
2) Exposure - examination of two main exposures:
    i. combination of ≥3 LFs: studies that also included metabolic/intermediate factors (*e.g.*, blood pressure/body mass index (BMI)) as part of their combination of LFs were included so long as the combination also included ≥3 'behavioural' LFs (*e.g.*, smoking/PA/diet).
    ii. SES: any SES measure (e.g., income/education/poverty-index).
3) Comparator: data for the influence of SES on associations between combinations of unhealthy LFs and adverse health.
4) Outcomes: at least one from: all-cause mortality, incidence and mortality from CVD or cancer.
5) Study design: prospective observational cohort. All types of analysis were included, and no study was excluded based on analysis method.

Exclusion criteria: not in English; abstracts/conference presentations only; ineligible design (*e.g.*, review/case-control/cross-sectional/qualitative). Studies were uploaded to 'DistillerSR' software (Version 2.38. DistillerSR Inc.; 2022. Accessed December 2021-February 2022; alternative software, Rayyan) and duplicates removed. Two reviewers (PP and HF/CO'D) screened titles and abstracts independently. Conflicts were resolved by discussion or included for full-text screening. Two reviewers (PP and HF) screened full-texts independently; conflicts were resolved by discussion with a third reviewer (CO'D).

4.4.3 Data extraction

Two reviewers (HF and PP/CO'D) extracted data independently using a piloted proforma (S4 Table, Appendix 7) [284] . After peer review, the proforma was adapted to include the distribution of type and number of unhealthy LFs among participants [31] . Quality was measured using the Newcastle-Ottawa Scale for cohort studies (NOS), a common measure of quality in observational or nonrandomised studies [286] . Other measures of quality exist (e.g., Risk Of Bias In Non-randomized Studies - of Exposure (ROBINS-E) tool) [287], each with advantages and disadvantages, but NOS is straightforward to use, specific for observational studies, and adaptable [286,288]. The NOS was adapted to include assessments of confounder adjustment, sensitivity analysis, and missing data methodology (S5 Table, Appendix 7) [31,284]. To compare study results, the following data from SES stratified analyses for each outcome was used to form our 'main comparator': 1) risk estimates for participants with the unhealthiest LF combination (using healthiest LF combination as reference) in the most advantaged SES group (*e.g.*, highest education, highest ranking occupation) were compared with 2) equivalent estimates (unhealthiest *versus* reference healthiest LF combination) in the least advantaged SES group (*e.g.*, lowest education, lowest ranking occupation). Studies frequently used more than two categories/quantiles of LF combinations, however only the estimates for the healthiest and unhealthiest categories were extracted. For example, for a study with a lifestyle score based on eight LFs, which study authors classified into five categories (scores 0–3, 4, 5, 6, and 7–8), the estimates for scores 0–3 and 7–8 were extracted. Estimates from SES stratified analyses were used for the main comparator because some studies did not report analyses examining combined influence of LF and SES using a single reference group (*i.e.*, analyses comparing all groups to the group with the healthiest combination of LFs and in the most advantaged SES group). However, results for these analyses were also extracted as they provide information on the combined influence of SES and lifestyle. To make direct comparisons, estimates from studies where the unhealthiest group was the reference were transformed to make the 'healthiest' group the reference. This transformation was achieved by dividing: 1) all hazard ratios (HRs) by the HR of the healthiest category (the healthiest category HR then becomes 1.00), 2) all lower confidence intervals (CIs) by the lower CI of the healthiest

category, and 3) all upper CIs by the upper CI of the healthiest category. This then requires swapping the upper and lower CIs because transformed lower CIs become upper CIs.

Meta-analysis was not appropriate due to the heterogeneity of included studies. Instead, results were reported and synthesised according to Synthesis Without Meta-analysis (SWiM) guidelines [32]. In accordance with transparent reporting of the synthesis methodology, this review adhered to the following approach - study results were grouped by outcome and compared by: 1) main models evaluating influence of SES; 2) model adjustment; 3) additional models, including sensitivity analyses; 4) tests for interactions; and 5) results for our main comparator.

4.5 Results

Results of the searches and screening are shown in a PRISMA flowchart (Figure 4-1).

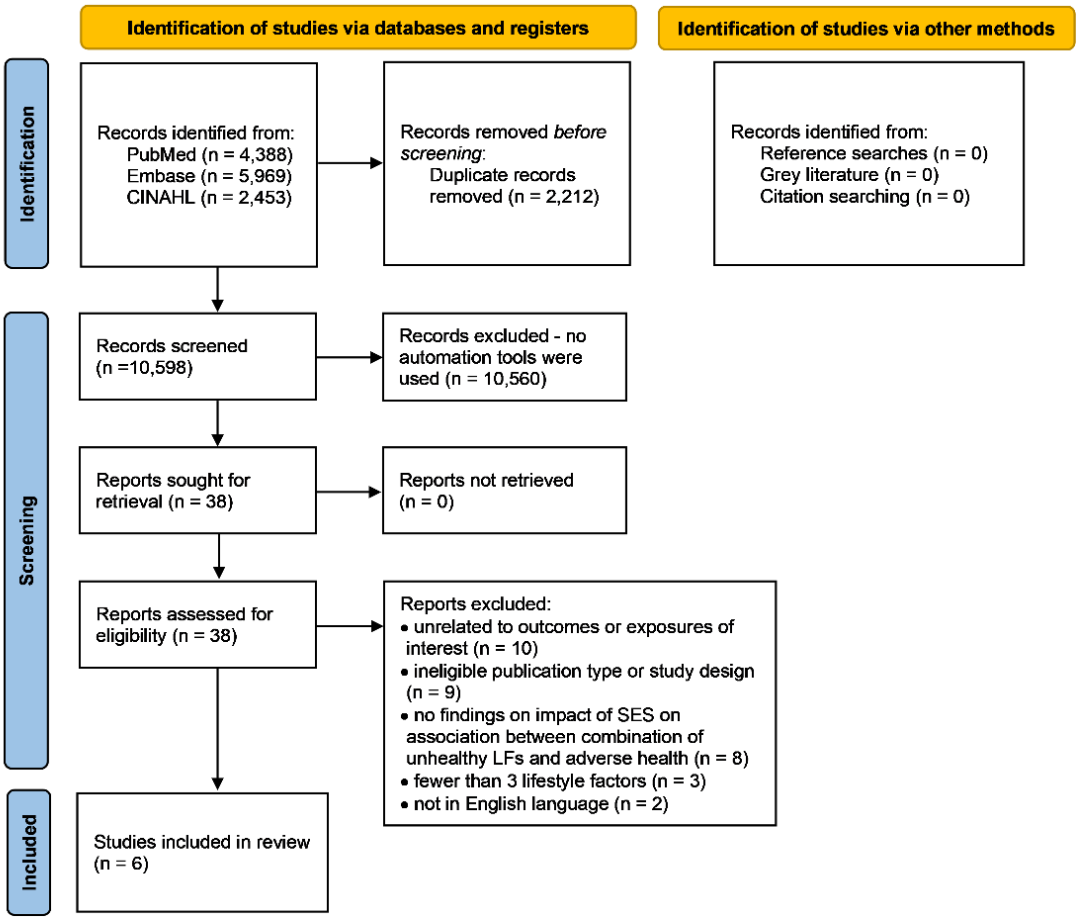


Figure 4-1: PRISMA flow chart of searches and screening results.

4.5.1 Study populations

Six studies of five cohorts were included in this review (Table 4-1) [24,289-293]. Two studies analysed the same USA-based cohort, The Southern Community Cohort Study (SCCS), but each study examined different LFs and SES exposures and therefore both were included [289,291]. Similarly, two studies analysed UKB and examined different exposure variables and outcomes [24,293]. The remaining cohorts analysed were The Japan Collaborative Cohort Study (JCCS) [290] , the National Health Interview Survey (NHIS) [292] , and US National Health and Nutrition Examination Survey (NHANES) [293] . SCCS was designed to investigate ethnic inequalities in healthcare and 86% of participants were recruited from community health centres; JCCS, UKB, NHIS and NHANES are general population cohorts with NHIS and NHANES designed to be nationally representative [294,295]. Participants per study ranged from 42,467–399,537; mean age ranged from 46.5–56.8 years; and the proportion of women from 54.6–59.3%. Ethnic composition of populations analysed varied: SCCS cohort was 67% African American [289,291]; JCCS ethnicity was not reported, but likely predominantly Japanese [290] ; UKB was 95% White British [24,293]; NHIS ethnicity was not reported [292] ; NHANES was 73.6% White [293] . Average follow-up time ranged from 4.3–19.3 years. In assessing the influence of SES on associations between combinations of unhealthy LFs and health outcomes, all studies examined all-cause mortality. In addition, three studies examined CVD mortality [24,290,293]; two examined CVD incidence [24,293]; one examined heart disease mortality and incidence of myocardial infarction and stroke [293]; two examined mortality from coronary heart disease (CHD), stroke, and cancer [290,293]; and one examined mortality from 'non-CVD and non-cancer' causes[290] . The non-CVD and non-cancer results are not reported here as they are outside the scope of this review.

Table 4-1: Characteristics of included studies.

Author, year	Cohort Country Type N	Age (years) Women (%) Ethnicity	Follow up length	Lifestyle factors and definitions of unhealthy (source or justification for unhealthy definition) Categories for analysis *	SES measure Categories for analysis	Outcome
Andersen *et al.*, 2016 [289]	Southern Community Cohort Study USA Prospective cohort designed to assess ethnic disparities in health outcomes, 86% recruited from community health centres 79,101 participants	Median age 51 (IQR 13) 59.3% women 67% African American	Max. 9 years (average not reported)	Self-reported at baseline Based on guidelines - i) current or former smoking (WHO) ii) alcohol intake >1 drink/d for women; >2 drinks/d for men (2010 Dietary Guidelines for Americans) iii) PA <150 min/wk moderate or <75 min/wk vigorous aerobic, or equivalent combination (2008 PA Guidelines for Americans) iv) sedentary time within 3 longest quartiles *i.e.*, >5.75 h/d (Avoid inactivity and limit sedentary behaviours; American Cancer Society) v) diet quality (HEI) in 3 lowest	Income self-reported at baseline Dichotomised as </≥ $15,000 p.a.	i) All-cause mortality *via* linkage to national registry

Author, year	Cohort Country Type N	Age (years) Women (%) Ethnicity	Follow up length	Lifestyle factors and definitions of unhealthy (source or justification for unhealthy definition) Categories for analysis *	SES measure Categories for analysis	Outcome
				quartiles *i.e.*, <66.7 (USDA's Center for Nutrition Policy and Promotion and previous publication) Five categories according to the number of lifestyle factor guidelines met (0, 1, 2, 3, ≥4).		
Eguchi *et al.*, 2017 [290]	The Japan Collaborative Cohort Study Japan Prospective cohort of general population 42,647 participants	Mean age 56.8 (SD not reported) 56.8% women Likely to be mainly Japanese ethnicity	Median 19.3 years (IQR 11.6–20.8)	Self-reported at baseline i) current smoking ii) alcohol intake > 2 'gou'/d (>46 g ethanol/d) iii) PA: <0.5 h/d or <5 h/wk walking and/or in sports iv) sleep duration: <5.5 or >7.4 h/d v) BMI: <21 or >25 vi) fruit intake: <1x/d	Education level as age at last formal education self-reported at baseline Dichotomised as </≥ 16 years old	i) All-cause mortality ii) CVD mortality iii) CHD mortality iv) Stroke mortality v) Cancer mortality

Author, year	Cohort Country Type N	Age (years) Women (%) Ethnicity	Follow up length	Lifestyle factors and definitions of unhealthy (source or justification for unhealthy definition) Categories for analysis *	SES measure Categories for analysis	Outcome
				vii) fish intake: <1x/d viii) milk intake: <almost daily; Five categories: according to healthy lifestyle score with one point for each lifestyle factor threshold met (0–3, 4, 5, 6, 7–8)		All *via* death certificate review
Andersen *et al.*, 2018 [291]	Southern Community Cohort Study USA Prospective cohort designed to assess ethnic disparities in health outcomes, 86% recruited from community health centres 77,896 participants	Median age 51 (IQR 13) 57.1% women 66.1% African American	Median 8 years (IQR not reported)	Self-reported at baseline Based on guidelines - i) alcohol intake >1 drink/d for women, >2 drinks/d for men (2010 Dietary Guidelines for Americans) ii) PA <150 min/wk moderate or <75 min/wk vigorous aerobic, or equivalent combination (2008 PA Guidelines for Americans) iii) sedentary time within 3	Neighborhood deprivation index (NDI): 2000 U.S. Census data linked to participant's residential address incorporating education, employment, housing, occupation, and poverty Quartiles	i) All-cause mortality *via* linkage to national registry

Author, year	Cohort Country Type N	Age (years) Women (%) Ethnicity	Follow up length	Lifestyle factors and definitions of unhealthy (source or justification for unhealthy definition) Categories for analysis *	SES measure Categories for analysis	Outcome
				longest quartiles *i.e.*, >6.5 h/d (Avoid inactivity and limit sedentary behaviours; American Cancer Society) iv) diet quality (HEI) in 3 lowest quartiles *i.e.*, <65.5 (USDA's Center for Nutrition Policy and Promotion) Four categories: according to number of guidelines met (0, 1, 2, 3–4)		
Foster *et al.*, 2018 [24]	UKB UK Prospective cohort of general population 328,594	Mean age 55.6 (SD 8.1) 54.6% women 95% White	Mean 4.9 years (SD 0.83, range 3.3–7.9) for all-cause and CVD mortality, 4.1 years (0.81 SD; range 2.4–7.0) for CVD incidence	Self-reported at baseline; Based on UK guidelines where available: i) current smoking ii) alcohol intake daily or almost daily iii) PA <150 min/wk moderate or	Townsend deprivation index: national census data incorporating car ownership, household overcrowding, owner occupation, and unemployment aggregated for and linked to participant postcode of residence Quintiles	i) All-cause mortality ii) CVD mortality iii) CVD incidence All *via* linkage to

Author, year	Cohort Country Type N	Age (years) Women (%) Ethnicity	Follow up length	Lifestyle factors and definitions of unhealthy (source or justification for unhealthy definition) Categories for analysis *	SES measure Categories for analysis	Outcome
				<75 min/wk vigorous iv) TV viewing time ≥4 h/d v) sleep duration <7 or >9 h/d vi) fruits and vegetables <400g/d vii) oily fish <1 portion/wk viii) red meat >3 portions/wk ix) processed meat >1 portion/wk Three categories: according to lifestyle risk score with one point for each unhealthy definition met (0–2, 3–5, 6–9)	Secondary SES measures for sensitivity analyses: i) Household income (£p.a.) self-reported at recruitment Five categories: >100,000; 52,000-100,000; 30,000-51,999; 18,000-29,999; <18,000 ii) Educational attainment self-reported at recruitment Five categories: College/University degree; A levels or equivalent; O levels or equivalent; CSEs or equivalent; none of the above	national registries
Choi *et al.*, 2022 [292]	National Health Interview Survey USA	Age ≥30 (average not reported)	Mean 12.7 years	Self-reported at baseline: i) current smoker and ex-smokers who quit <20 years ago	Household income as a ratio of family income to federal poverty level	i) All-cause mortality via linkage to national registry

Author, year	Cohort Country Type N	Age (years) Women (%) Ethnicity	Follow up length	Lifestyle factors and definitions of unhealthy (source or justification for unhealthy definition) Categories for analysis *	SES measure Categories for analysis	Outcome
	Prospective cohort of general population aged ≥30 years 189,087	Proportion female not reported Ethnicity not reported		ii) weekly alcohol intake >14 drinks for men, >7 drinks for women (2016 NIAAA guidelines), or >5 drinks/d at least monthly iii) PA <150 min/wk moderate or <75 min/wk vigorous and/or strengthening activities <2 d/wk (2008 Physical Activity Guidelines for Adults) iv) BMI <18.5 or ≥35 Five categories according to the number of lifestyle factors: 0, 1, 2, 3, or 4	Dichotomised as < or ≥200% of federal poverty level	
Zhang *et al.*, 2021 [293]	US National Health and Nutrition Examination Survey (NHANES) USA Prospective cohort	NHANES: Mean age 46.5 51.3% women	NHANES: Mean 11.2 years UKB: Mean 11.0 years for all-cause	Self-reported at baseline; i) smoked >100 cigarettes in lifetime ii) daily alcohol intake >1 drink for women, 2 drinks for men (National guidelines for USA and	NHANES: i) family poverty to income ratio: low (≤1); middle (1-4); and high (≥4) ii) educational attainment: less than high school diploma;	i) All-cause mortality ii) CVD mortality iii) CVD incidence

Author, year	Cohort Country Type N	Age (years) Women (%) Ethnicity	Follow up length	Lifestyle factors and definitions of unhealthy (source or justification for unhealthy definition) Categories for analysis *	SES measure Categories for analysis	Outcome
	of general population 44,462 & UKB UK Prospective cohort of general population 399,537	73.6% White UKB: Mean age 56.1 52.5% women 95.6% White	mortality, 8.8 years for CVD incidence	UK) iii) leisure time physical activity at level of lower two thirds of study participants iv) diet quality (HEI) at level of lower 2/5 th of participants for US NHANES (2015-20 Dietary Guidelines for Americans & 1992 food guide from US Department of Agriculture); meeting 5/10 diet recommendations for UKB (evidence-based recommendations) Three categories according to number of lifestyle factors (score): 0–1, 2, 3–4 Sensitivity analyses included: a weighted lifestyle score to account for differing magnitude of associations between each LF	high school graduate or equivalent; and college or above iii) occupation (US socioeconomic index): upper (index ≥50); lower (index <50); and unemployment iv) health insurance: private; public only; none Variables i)-iv) were self-reported at recruitment and combined *via* latent class analysis to generate 3 latent classes/categories of low, medium, and high SES UKB: i) income (£p.a.): >100,000; 52,000-100,000; 30,000-51,999; 18,000-29,999;	iv) heart disease mortality, NHANES only v) coronary heart disease mortality, UKB only vi) stroke mortality, UKB only vii) myocardial infarction incidence, UKB only vii) stroke incidence, UKB only

Author, year	Cohort Country Type N	Age (years) Women (%) Ethnicity	Follow up length	Lifestyle factors and definitions of unhealthy (source or justification for unhealthy definition) Categories for analysis *	SES measure Categories for analysis	Outcome
				and outcomes; and a combination of LFs that included BMI outwith 18.5–24.9.	<18,000 ii) educational attainment: College/University degree; A levels or equivalent; O levels or equivalent; CSEs or equivalent; NVQ, HND, HNC, or equivalent; other professional qualifications; none of the above iii) employment: employed (including self-employed, retired, unpaid/voluntary work, full/part time students); unemployed Variables i)-iii) were self-reported at recruitment and combined *via* latent class analysis to generate 3 latent classes/categories of low, medium, and high SES	All *via* linkage to national registries

Author, year	Cohort Country Type N	Age (years) Women (%) Ethnicity	Follow up length	Lifestyle factors and definitions of unhealthy (source or justification for unhealthy definition) Categories for analysis *	SES measure Categories for analysis	Outcome
					Secondary SES measures in sensitivity analyses included: each SES factor individually; Townsend index (UKB only)	

N, number of participants included in analysis; *Categories for analysis shows the number of categories used by study authors to analyse the associations between the combination of lifestyle factors and health outcome (*e.g.*, a study of five lifestyle factors, with possible scores of 0 to 5, could be analysed using the score categories of 0, 1, 2, 3, and ≥4; *i.e.*, with scores 4 and 5 grouped together); Outcomes, adverse health outcome used to assess interaction between lifestyle and SES (some studies reported additional health outcomes but these were not used to assess interaction); SES, socioeconomic status; IQR, Interquartile range; PA, physical activity; BMI, body mass index (kg/m ^{2}); WHO, World Health Organisation; d, day; wk, week; h, hours; min, minutes; $, US dollars; pa per annum; £, British pounds; TV, television; HEI, Healthy Eating Index, which measures adherence to the Dietary Guidelines for Americans. HEI is based on 12 dietary components: total fruits; whole fruits; total vegetables; greens and beans; whole and refined grains; dairy; total protein foods; seafood and plant proteins; fatty acids; sodium; and calories from solid fats, alcohol, and added sugars (range 0–100; higher values indicate healthier diet); CVD, cardiovascular disease; CHD, coronary heart disease; UKB, UK Biobank; NIAAA, National Institute on Alcohol Abuse and Alcoholism; A level, General Certificate of Education Advanced Level; O-level, General Certificate of Education Ordinary Level; CSE, Certificate of Secondary Education; NVQ, National Vocational Qualification; HND/HNC, Higher National Diploma/Certificate.

4.5.2 Combinations of unhealthy lifestyle factors

The number of LFs comprising the combination in each study ranged from four to nine and included: smoking, alcohol, PA, sedentary time, television (TV) viewing time, various individual dietary factors, a dietary index, and sleep duration (Table 4-1). Two studies included BMI in main analyses and one study included BMI in a sensitivity analysis [290,292,293]. Alcohol and PA were included in all studies and dietary factors were missing from only one study [292]. Smoking was included in five studies but excluded from relevant analyses in the remaining study [291] . All LF data was collected *via* baseline questionnaire or interview.

4.5.3 Definition or classification of unhealthy for individual lifestyle factors

In each study individual LFs were dichotomised as healthy/unhealthy with one point per factor summed to create an unweighted score. Two studies also created weighted scores using the strength of association between individual LFs and outcomes [289,293] . However, only one of these examined the effect of SES on a weighted score for which results were extracted here [293] . Three studies summed healthy LFs to create 'healthy' scores [289,290,293], while the remaining three studies created 'unhealthy' scores (results were harmonised to show increasing risk with increasingly unhealthy lifestyle) [24,291,292].

The definition of unhealthy for each individual LF included in the LF combinations varied (Table 4-1). Unhealthy smoking status was defined as current smoking [24,290] , current/any former smoking [289] , current/quitting <20 years ago [292] , and smoking more than 100 cigarettes in a lifetime [293]. Unhealthy alcohol intake was defined as: >1 drink/day for women or >2 drinks/day for men [289,291,293], >5 drinks/day monthly [292], >46 g alcohol/day [290], and 'daily/almost daily intake', respectively [24]. Unhealthy PA levels were defined as <150 minutes/week moderate or <75 minutes/week vigorous PA in four studies [24,289,291,292], as strengthening activities on <2 days/week [292], as not achieving either ≥0.5 hours/day walking or ≥5 hours/week walking/playing sports [290] , and as having leisure time PA levels in the lower two thirds of study participants [293] . Unhealthy sedentary time, considered in two studies, was defined as the three quartiles with longest sedentary time

(*i.e.*, >5.75 and >6.5 hours/day), respectively [289,291]. Unhealthy TV viewing time, examined in one study, was defined as ≥4 hours/day [24]. Unhealthy sleep duration, examined in two studies, was classified as <5.5/>7.4 hours/day [290] and <7/>9 hours/day [24] , respectively.

Dietary factors examined varied considerably. Three studies of two USA-based cohorts used a national dietary index (comprising fruit, vegetables, grains, proteins, fatty acids, sodium, and calories from fats, alcohol, and added sugars), defining unhealthy as either belonging to the three lowest quartiles [289,291] or two lowest quintiles [293] . The Japanese cohort study included three dietary components, defining unhealthy as: fruit <once/day; fish <once/day; and milk <almost daily [290] . One of the studies examining the UK-based UKB included four components, classifying unhealthy as: fruit and vegetables <400 g/day; oily fish <1 portion/week; red-meat >3 portion/week; and processed-meat >1 portion/week [24]. Whereas the other study of UKB classified unhealthy as meeting at least five of 10 recommendations [293] .

Justification for the classification of 'unhealthy' varied. One study cited WHO guidelines for the classification of unhealthy smoking [289]. Four studies of USA-cohorts used US national guidelines to define unhealthy alcohol intake and diet [289,291,293]. And of those, two also used US guidelines to define PA and sedentary time [289,291]. One study adapted a previous lifestyle score [7] , using UK guidelines or standards from the original score [24]. One study did not report the basis for their definitions of unhealthy for eight LFs including a BMI outwith 21–25 [290] . The other study that examined BMI in their main analyses based the definition of unhealthy (<18.5 or ≥35) on prior analysis of the data [292]. Unhealthy BMI (outwith 18.5–24.9) was based on previous research in the third study that included BMI in a sensitivity analysis [293] .

Most studies had approximately normal distributions of the total number of unhealthy LFs among participants (S6 Table, Appendix 7) [284]. One study of UKB, with nine LFs, had relatively few participants with six to nine unhealthy LFs [24]. The other study of UKB, with four LFs, had more participants with unhealthy LFs [293] . The proportion of study participants with specific unhealthy LFs also varied. For example, the proportion of study participants with unhealthy smoking status ranged from 9.6% to 64%; some of this

discrepancy is likely due to differences in the definition of unhealthy (*i.e.*, current *versus* current/former smoking).

4.5.4 Socioeconomic status

SES measures varied by study (Table 4-1). For main analyses, two studies used area-based deprivation indices: Neighborhood deprivation index (NDI) and Townsend deprivation index (TDI) [24,291]. Data for both indices were obtained *via* national censuses from or near baseline. NDI comprises five 'domains': education, employment, housing, occupation, and poverty [291]. Whereas TDI comprises data on car ownership, household overcrowding, owner occupation, and unemployment [24]. Two studies used self-reported individual-level measures of income at recruitment [289,292] and one of these operationalised income as a ratio of family income to the USA federal poverty level [292]. One study used age at last formal education obtained *via* baseline self-report for the main analyses [290]. Finally, one study of two cohorts used latent class analysis to generate an overall SES variable from four SES measures (income, occupation, education, and health insurance) in analysis of one cohort and three SES measures (income, education, and employment status) in analysis of the second cohort [293]. In sensitivity analyses, two studies examined alternative SES measures [24]. One study swapped area-based TDI for annual household income and, separately, individual-level educational attainment [24]. The second study performed multiple sensitivity analyses of alternative SES measures by replacing a latent class SES variable with 1) each SES measure (income, occupation, education, health insurance, and employment status) used to generate the latent class; 2) an SES score based on each single SES measure; 3) and TDI [293].

4.5.5 Categories for analysis

Categorisation of the two main exposures (combination of LFs and SES) used in analyses varied (Table 4-1). Categories for combinations of LFs ranged from three to five and were not always related to the number of LFs included and often influenced by the number of participants with unhealthy LFs. For example, one study examined nine LFs and split participants into three categories: 'healthy' (score 0–2), 'moderately healthy' (score 3–5), and

'unhealthy' (score 6–9) [24]; whereas another study included eight LFs and split participants into five categories [290].

For SES measures, the following categories were used: income dichotomised as </≥ $15,000 US dollars per annum [289] ; age at last formal education dichotomised as </≥ 16 years [290] ; quartiles of NDI [291]; quintiles of TDI [24]; ratio of family income to federal poverty level dichotomised as < or ≥200% of federal poverty level [292]; three latent classes of low, medium and high SES [293].

4.5.6 Analysis procedures

Each study conducted descriptive analyses, examining independent associations between combinations of LFs and outcomes and between SES and outcomes. All studies used Cox-proportional hazard models in their main analyses to estimate HRs and 95% confidence intervals (CIs) for outcomes for each LF combination category, stratified by SES (Table 4-1). Three studies additionally stratified these analyses; one by ethnicity and sex together (African American/White and female/male) [289] , three by sex alone [290,291,293], one by ethnicity (White/Non-white) [293], and one by age (≥60/<60 years) [293]. One study that stratified by sex alone, also performed a separate analysis on the total population (not stratified by sex) [290]. Two studies did not additionally stratify by sociodemographics [24,292]. The number of confounder variables chosen by studies ranged from five to 14 (Table 4-2). All studies adjusted for either age, age plus age squared, or used age as the time-varying covariate. Studies varied in their additional analyses to investigate the influence of SES and included: single reference group analyses to investigate the joint associations of combinations of unhealthy LFs, SES, and outcomes [24,290,291,293]; Kaplan-Meier survival curves for combinations of unhealthy LFs stratified by SES [290]; tests for multiplicative interactions between combinations of unhealthy LFs and SES [24,289-293]; and tests for additive interactions (Table 4-2)[24,293].

Table 4-2: Methods/results for influence of SES on association between combinations of unhealthy LFs and outcomes.

Study	Methods	Covariates (n)	Interaction tests between combinations of unhealthy LFs and SES	Main interaction results ($P_{interaction\ or\ RERI}$)	Result summary
Andersen 2016 [289]	1) Cox-proportional hazard models for all-cause mortality for combination of unhealthy LF categories 2) Models stratified by low/high income in sub-group analysis	Enrolment source, education, marital status, neighbourhood deprivation, and BMI (5)	Likelihood ratio tests, comparing main effects models with and without cross-product terms	All-cause mortality: 0.002 (African American men); 0.89 (African American women); 0.04 (White men); 0.49 (White women)	Significant multiplicative interaction for African American and White men only: highest HRs for combination of unhealthy LFs and high income Only stratified (sex, ethnicity) results available
Eguchi 2017 [290]	1) Cox-proportional hazard models for outcomes for combinations of unhealthy LF categories, stratified by low/high education (analyses for total population and separate analyses further stratified by sex) 2) Cox-proportional hazard models for	Age, sex, history of hypertension, history of diabetes, perceived mental stress and regular employment (6)	Cross-product of dichotomous education level and healthy lifestyle score (continuous) in models for total CVD and all-cause mortality outcomes only	All-cause mortality: 0.11 CVD mortality: 0.23 (both for total population only)	1) No evidence of multiplicative interaction, with similar HRs for combinations of unhealthy LFs and both high and low SES 2) Single reference group analysis provides evidence for additive interaction for all-cause and CVD mortality: higher HRs in least healthy combination of LFs and lowest education groups 3) Survival curves suggest additive interaction: steeper curve (highest mortality) for combination of unhealthy LFs in low education

Study	Methods	Covariates (n)	Interaction tests between combinations of unhealthy LFs and SES	Main interaction results ($P_{\text{interaction or RERI}}$)	Result summary
	combinations of unhealthy LF categories and education level using single reference group (all-cause and CVD mortality only) 3) Kaplan-Meier survival curves for combinations of unhealthy LF categories, stratified by low/high education (all-cause and CVD mortality only) 4) Sensitivity analysis examining two modified LF combinations				group 4) Sensitivity analysis: i) extended definition of healthy sleep and ii) dichotomous diet score (five components) in addition to extended sleep definition - consistent with findings from main analysis
Andersen 2018 [291]	1) Cox-proportional hazard models for all-cause mortality for combinations of unhealthy LF	Enrolment source, ethnicity, education, income, marital status, and insurance status (6)	Likelihood ratio tests, comparing main effects models with and without cross-product terms	All-cause mortality: 0.28 (men); 0.99 (women)	1) No evidence of multiplicative interaction with similar HRs for combinations of unhealthy LFs in both high and low SES 2) Single reference group analysis

Study	Methods	Covariates (n)	Interaction tests between combinations of unhealthy LFs and SES	Main interaction results ($P_{\text{interaction or RERI}}$)	Result summary
	categories, stratified by NDI quartiles 2) Cox-proportional hazard models for all-cause mortality for combinations of unhealthy LF categories, stratified by NDI quartiles using single reference (also stratified by sex)				provides evidence for additive interaction in men and women for all-cause mortality: highest HRs in the least healthy combination of LFs and lowest SES (highest NDI) group
Foster 2018 [24]	1) Cox-proportional hazard models for outcomes for combinations of unhealthy LF categories, stratified by SES quintiles (TDI, income, and education examined separately) 2) Cox-proportional hazard models for joint associations of combinations of	Age, sex, ethnicity, month of assessment, hypertension, systolic blood pressure, medication for hypercholesterolaemia or hypertension, and BMI (8)	1) Interaction term between combinations of unhealthy LFs and SES variables in models 2) Interaction sensitivity analyses (deprivation index): a) additional models with interaction term and i) dichotomous and ii)	<u>Deprivation index</u> All-cause and CVD- mortality: <0.0001 CVD incidence: 0.11 <u>Income</u> All-cause mortality: 0.001 CVD mortality: <0.0001 CVD incidence:	1) Significant multiplicative interaction between combination of unhealthy LFs and deprivation/education for all-cause and CVD mortality but not for CVD incidence. Significant multiplicative interaction between combinations of unhealthy LFs and income for all outcomes 2) Single reference analysis showed highest HRs for all-cause and CVD mortality in the least healthy combination of LFs and lowest SES

Study	Methods	Covariates (n)	Interaction tests between combinations of unhealthy LFs and SES	Main interaction results ($P_{\text{interaction or RERI}}$)	Result summary
	unhealthy LF categories and SES measures (single reference group)		continuous combination of unhealthy LF variable b) Estimation of three measures of 'biological interaction': RERI, AP, and synergy index	0.009 <u>Education</u> All-cause mortality: 0.002 CVD mortality: 0.047 CVD incidence: 0.051 (all for total population only)	groups 3) Interaction sensitivity results consistent with main findings with significant interaction across three measures of additive interaction
Choi 2022 [292]	1) Cox-proportional hazard models for all-cause mortality for number of unhealthy LFs, stratified by high/low income group	Age, age squared, sex, education, race/ethnicity, acculturation, income assistance, health insurance, and marital status (9)	Unclear, but likely an interaction term between combinations of unhealthy LFs and income in models	<u>Primary outcomes</u> All-cause mortality: <0.05	Significant multiplicative interaction between combinations of unhealthy LFs and income for all-cause mortality. Mortality risk associated with each additional unhealthy LF was higher in higher income group.
Zhang 2021 [293]	1) Cox-proportional hazard models for outcomes for combinations of unhealthy LF categories, stratified	Age, sex, marital status (NHANES only), assessment centre (UKB only), self-reported race/ethnicity, acculturation score, BMI, hypertension, diabetes, CVD,	1) Interaction term between combinations of unhealthy LFs and SES variables in models	All-cause mortality: 0.85; RERI =0 (NHANES), <0.001; RERI >0 (UKB)	1) NHANES: no significant multiplicative (product term for interaction) or additive interaction (RERI) between combination of unhealthy LFs and SES for all-cause or heart disease mortality

Study	Methods	Covariates (n)	Interaction tests between combinations of unhealthy LFs and SES	Main interaction results ($P_{\text{interaction or RERI}}$)	Result summary
	by SES category 2) Cox-proportional hazard models for outcomes joint associations of combinations of unhealthy LF categories and SES measures (single reference group) 3) Sensitivity analyses for models stratified by SES category by examining subgroups: male/female, white/non-white ethnicity, age </≥60 years 4) Sensitivity analyses of joint associations substituting individual-level latent class SES for: a) Each SES component	cancer, lung disease (UKB only). (10-14)	2) Estimation of RERI	CVD mortality: 0.002; RERI >0 (UKB) CVD incidence: 0.016; RERI >0 (UKB) <u>Secondary outcomes</u> Heart disease mortality: 0.29; RERI =0 (NHANES) Coronary heart disease mortality: 0.008; RERI >0 (UKB) Stroke mortality:	UKB: both significant multiplicative and additive interactions between combination of unhealthy LFs and SES for all-cause mortality, CVD mortality, CVD incidence, coronary heart disease mortality, and stroke mortality but not for myocardial infarction incidence or stroke incidence 2) Results for product term for interaction and RERI similar across sensitivity analyses (individual-level SES, individual/area-level SES mutual adjustment) 3) In both cohorts, joint association analysis showed highest HRs in the least healthy combination of LFs and lowest SES groups for all outcomes and across all sensitivity analyses 4) Subgroup analyses showed significant multiplicative and additive interactions between combination of unhealthy LFs and

Study	Methods	Covariates (n)	Interaction tests between combinations of unhealthy LFs and SES	Main interaction results ($P_{interaction\ or\ RERI}$)	Result summary
	used to generate latent class, separately b) Townsend index (area-level) with adjustment for latent class SES (UKB only) and *vice versa*			0.002; RERI >0 (UKB) Myocardial infarction incidence: 0.050; RERI >0 (UKB) Stroke incidence: 0.032; RERI >0 (UKB)	SES for most subgroups (sex/ethnicity/age) and primary outcomes in UKB but not in NHANES 5) Subgroup analyses of the joint associations of combination of unhealthy LFs and SES showed higher HRs in men *vs.* women and in younger *vs.* older adults for all-cause mortality in both cohorts, and in younger *vs.* older adults for CVD mortality in UKB

[i] LFs, lifestyle factors; SES, socioeconomic status; $P_{interaction}$, p-value for interaction between combinations of unhealthy LFs and SES; RERI, relative excess risk due to interaction; HR, hazard ratio; NDI, Neighborhood deprivation index; TDI, Townsend deprivation index; BMI, body mass index; CVD, cardiovascular disease; 'Biological interaction', the degree of interaction between risk factors in terms of deviation from additivity in adverse health outcome rates [296] ; AP, attributable proportion; UKB, UK Biobank; NHANES, US National Health and Nutrition Examination Survey.

4.5.7 Study quality

Results for study quality as measured by the adapted NOS ranged from 5–9 (max. 9; S7 Table, Appendix 7) [284]. Only two studies examined more than one SES measure [24,293] and only three studies attempted to reduce the chance of reverse causality by demonstrating participants were free from disease at the start of the study[24,290,293].

4.5.8 The influence of socioeconomic status on lifestyle-associated health

Using the main comparator as an assessment of the influence of SES on the association between combinations of unhealthy LFs and outcomes, results across studies were mixed and varied by outcome (Figure 4-2 and Figure 4-3). A synthesis of results, including the main comparator, is structured by outcome below.

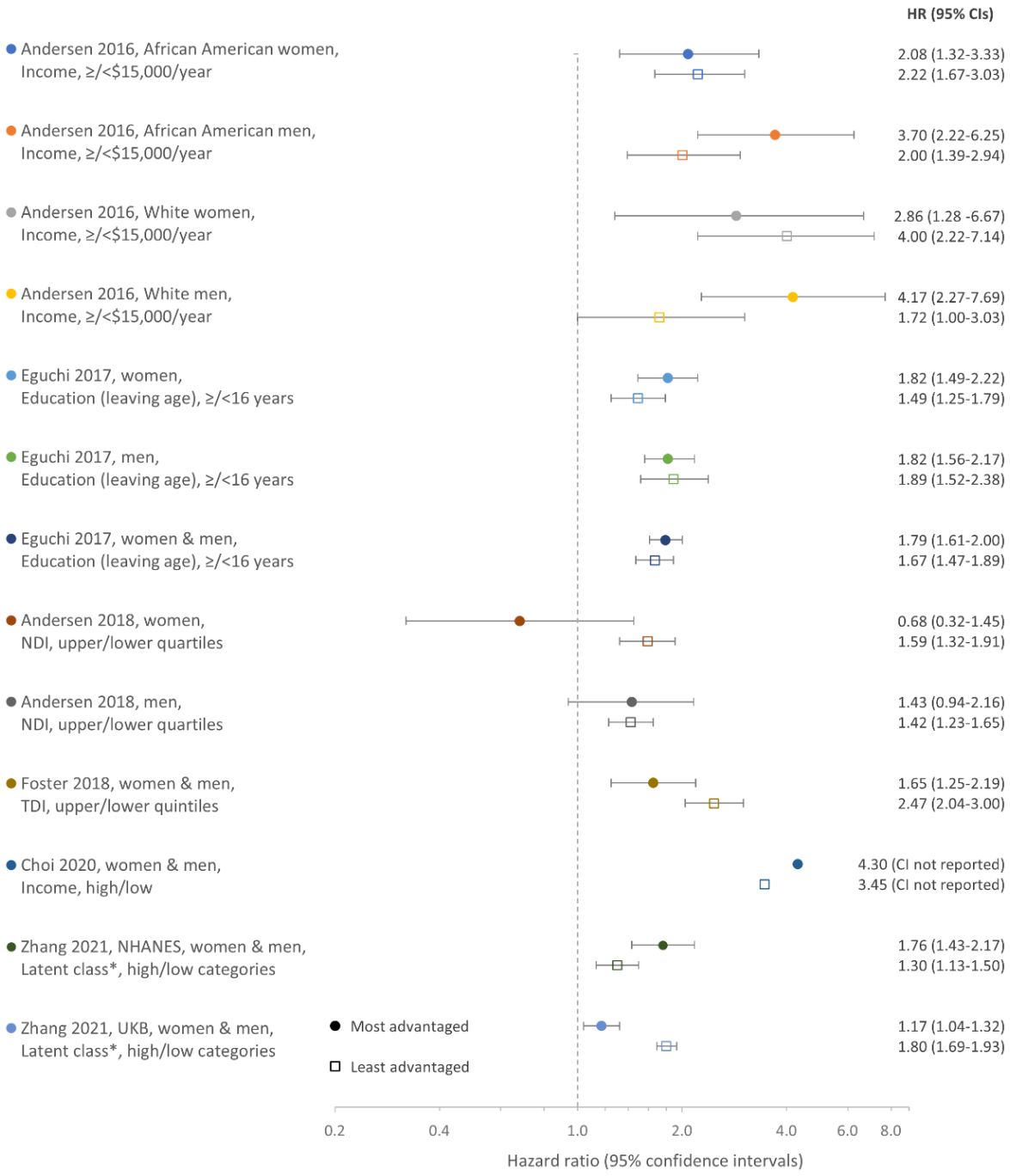


Figure 4-2: Hazard ratios for the association between combinations of unhealthy LFs and all-cause mortality in the most and least advantaged SES groups by study and population. Comparison of HRs from SES stratified analyses for the associations between combinations of unhealthy LFs and all-cause mortality in the most and least advantaged SES groups (main comparator). Combinations of healthy LFs in the same SES strata (most/least advantaged) are the reference group. Legend indicates the study, population, SES measure, and definition for the most/least advantaged SES groups.*Latent class analysis based on income, education, occupation/employment, and (for NHANES only) health insurance.

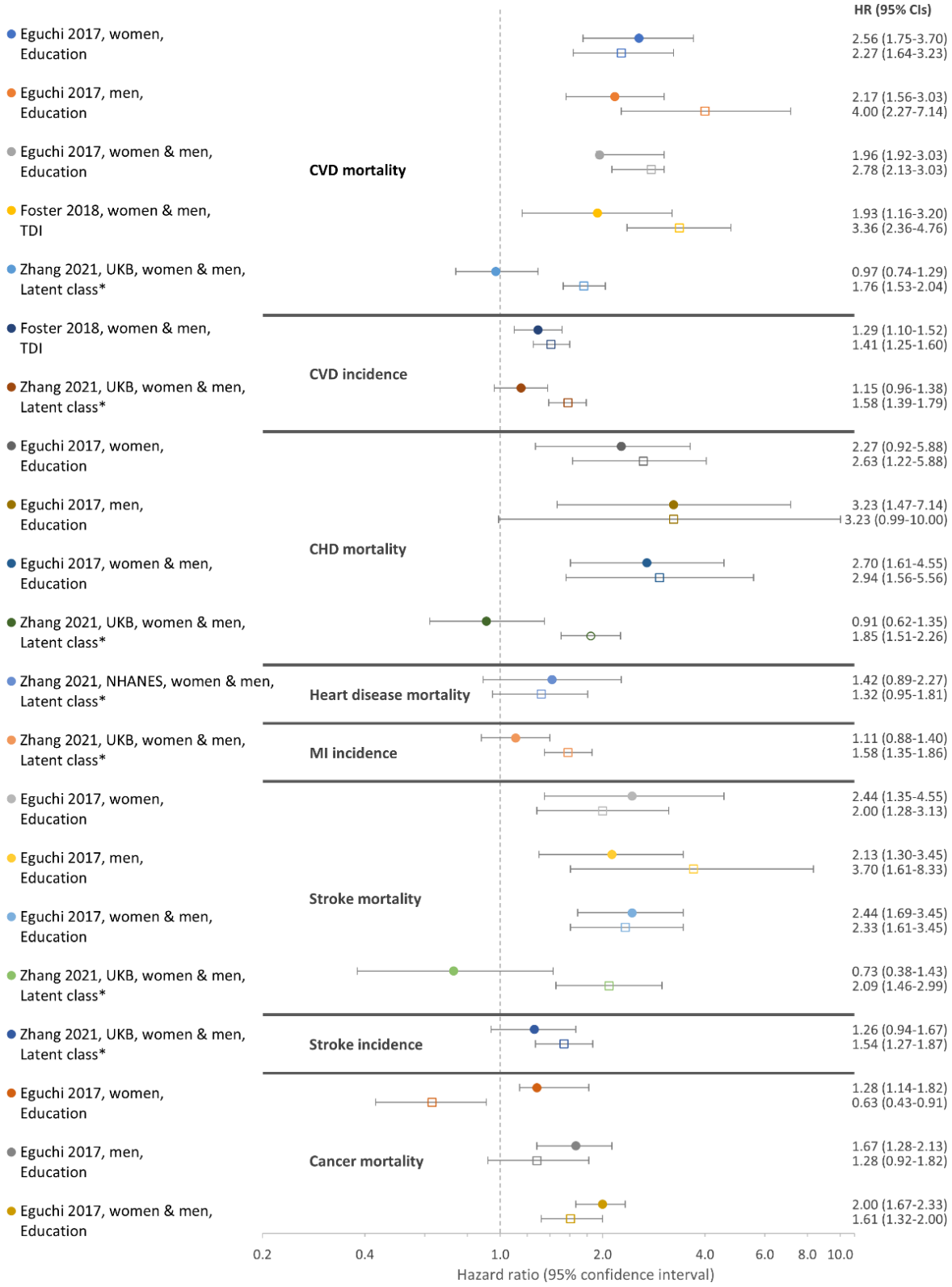


Figure 4-3: Hazard ratios for the association between combinations of unhealthy LFs and adverse health outcomes in the most and least advantaged SES groups by study and population.

Comparison of HRs from SES stratified analyses for the associations between combinations of unhealthy LFs and adverse health outcomes in the most and least advantaged SES groups (main comparator). Combinations of healthy LFs in the same SES strata (most/least advantaged) are the reference group. Legend indicates the study, population, and SES measure.

4.5.8.1 All-cause mortality.

Estimates from 13 main analyses were available for the main comparator for all-cause mortality as some analyses were additionally stratified by sex or by both sex and ethnicity (Figure 4-2 and Table S8A, Appendix 7). All studies observed that, compared with healthy LFs, combinations of unhealthy LFs were generally associated with higher all-cause mortality. However, the difference between the higher all-cause mortality associated with a combination of unhealthy *versus* that associated with healthy LFs was greater in the most advantaged SES group in seven analyses, but greater in the least advantaged group in the remaining six analyses (Figure 4-2 and Table S8A, Appendix 7). However, there was considerable overlap of CIs from most and least advantaged SES groups and the difference between some estimates from most and least advantaged groups were similar. HRs (95%CIs) from the most advantaged groups ranged from 0.68 (0.32–1.45) to 4.17 (2.27–7.69); equivalent estimates from the least advantaged groups ranged from 1.30 (1.13–1.50) to 4.00 (2.22–7.14).

Two analyses were additionally stratified by sex alone and, for women, the difference in all-cause mortality associated with unhealthy *versus* healthy LFs was greater in the least advantaged group in one study [291] but greater in the most advantaged group in the other study [290] (and *vice versa* for men). The study that additionally stratified by both sex and ethnicity observed the difference in all-cause mortality associated with combinations of unhealthy *versus* healthy LFs was consistent for sex across two ethnic groups: greater in the least advantaged SES group for women of both African American and White ethnicity, but greater in the most advantaged SES group for men of both ethnicities [289]. One study stratified by sex for sensitivity analysis and observed similar all-cause mortality associated with combinations of unhealthy *versus* healthy LFs for both sexes in the most *versus* least advantaged groups [293]. However, the same study examined two cohorts and found that although the difference in all-cause mortality associated with combinations of unhealthy *versus* healthy LFs was small for men and women, it was greater in the most advantaged group in one cohort (NHANES) and in the least advantaged group in the other cohort (UKB) [293]. Sensitivity analysis results from one study of two cohorts that additionally stratified by ethnicity alone (White/Non-White) were mixed [293]. In the same study, sensitivity analysis stratified by age alone (≥60/<60 years old) suggested that all-cause

mortality associated with combinations of unhealthy LFs was relatively higher for those <60 years old in the least advantaged groups in both cohorts [293]. Five main analyses from four cohorts examined the total population (not further stratified by sociodemographic variables) and the difference in all-cause mortality associated with unhealthy *versus* healthy LFs was greater in the most advantaged SES group in three cohorts (JCCS, NHIS, NHANES) [290,292,293] but greater in the least advantaged group in another cohort (UKB) [24,293]. Similarly mixed results were found with the sensitivity analyses (Table S8, Appendix 7) [284].

Results of tests for multiplicative interactions were also mixed (Table 4-2). A significant multiplicative interaction between the combination of LFs and SES was observed in four studies, but in opposing directions [24,289,292,293]. A significant multiplicative interaction was observed, with greater all-cause mortality associated with combinations of unhealthy LFs in the most advantaged group in the entire cohort of one study [292] but only in men in another study [289]. Whereas a significant multiplicative interaction was seen in two studies of UKB, where the difference in all-cause mortality associated with unhealthy *versus* healthy LFs was greater in the least advantaged group [24,293]. The multiplicative interaction observed in UKB was observed consistently across a set of interaction sensitivity analyses (Table 4-2) [24,293]. Two studies tested for and found significant additive interactions in the same cohort (UKB) [24,293] but one of these studies did not observe significant additive interactions in similar analysis of a second cohort (NHANES) [293] . Four studies of three cohorts examined combined associations of combinations of unhealthy LFs and less advantaged SES in eight analyses by comparing all groups to a single reference: the healthiest LF-most advantaged group [24,290,291,293]. In these analyses, HRs (95%CIs) for all-cause mortality for the least healthy-least advantaged group ranged from 1.43 (1.11–1.84) to 3.53 (3.01–4.14) (Table S9, Appendix 7) [284]. The highest all-cause mortality was observed in the least healthy-least advantaged groups in seven of eight of these analyses, suggesting an additive interaction between unhealthy LFs and less advantaged SES. For sensitivity, two studies examined additional measures of SES separately in single reference group analyses and consistently observed the highest all-cause mortality in the least healthy-least advantaged groups irrespective of SES measure [24,293]. Further evidence for an additive interaction came from the

steeper Kaplan-Meier curves for an unhealthy combination of LFs in the least advantaged *versus* most advantaged group in one study [290].

4.5.8.2 CVD mortality.

Three studies examined CVD mortality in two cohorts. Compared with healthy LFs, combinations of unhealthy LFs were consistently associated with higher CVD mortality [24,290,293]. In analyses stratified by SES alone, all three studies observed the difference in CVD mortality associated with unhealthy *versus* healthy LFs was greater in the least advantaged SES group: HRs (95%CIs) in the least advantaged groups were 2.78 (2.13–3.03) [290] , 3.36 (2.36–4.76)[24], and 1.76 (1.53–2.04) [293], respectively. Equivalent estimates in the most advantaged groups were 1.96 (1.92–3.03), 1.93 (1.16–3.20), and 0.97 (0.74–1.29) (Figure 4-3 and Table S8D, Appendix 7). One of these studies also stratified analyses by sex and found the difference in CVD mortality associated with unhealthy *versus* healthy LFs was greater in the most advantaged group for men but in the least advantaged group for women [290] . However, the unhealthy *versus* healthy LFs CVD mortality for women was similar in the most and least advantaged groups. Similar results were found in sensitivity analyses (Table S8, Appendix 7)[284] . For LF-SES interactions for CVD mortality, one study provided evidence of an additive interaction through both single reference group analyses and steeper Kaplan-Meier survival curves [290]. In this study's single reference group analysis, the highest CVD mortality was associated with those in the least healthy-least advantaged group (Table S10, Appendix 7)[284,290]. However, the same study found no significant multiplicative interaction for CVD mortality (Table 4-2). By contrast, two other studies, both examining CVD mortality in UKB, reported a significant multiplicative interaction and in the single reference analysis, the least healthy-least advantaged group had markedly higher hazards than the least healthy-most advantaged group: 4.59 (3.33–6.32) *vs.* 2.01 (1.21–3.33)[24] and 2.65 (2.09–3.38) and 1.06 (0.80–1.39)[293] , respectively (Table S10, Appendix 7)[284]. Both studies observed significant multiplicative interactions for CVD mortality consistently irrespective of SES measure and across interaction sensitivity analyses [24,293].

4.5.8.3 Other outcomes.

Estimates for CVD incidence were provided by two studies of UKB and, in SES stratified analyses, compared with healthy LFs, combinations of unhealthy LFs were associated with higher CVD incidence [24,293]. The difference in CVD incidence associated with combinations of unhealthy *versus* healthy LFs was greater in the least advantaged groups in both studies (Figure 4-3 and Table S8E, Appendix 7). In combined single reference analysis, HRs (95%CIs) for the least healthy-most advantaged *versus* least healthy-least advantaged groups were: 1.30 (1.10–1.53) *versus* 1.75 (1.55–1.97) [24] and 1.18 (0.99–1.41) *versus* 2.09 (1.78–2.46) [293], respectively (Table S10, Appendix 7)[284]. Results from tests for SES-LF interactions for CVD incidence were mixed. Significant additive and multiplicative interactions were observed in one study (examining four LFs and latent class SES)[293] but not the other (examining nine LFs and area-based TDI)[24].

Two studies examined additional outcomes [290,293]. One of these performed SES-stratified analyses but did not report single reference group analyses or tests for interaction for these outcomes [290]. In this study's SES-stratified analyses, the difference in hazards associated with combinations of unhealthy *versus* healthy LFs for the total population was greater in the most advantaged group for mortality from stroke, and cancer but greater in the least advantaged group for CHD mortality (Figure 4-3 and Table S8F, Appendix 7). Equivalent estimates from analyses additionally stratified by sex were similar, although, in men, the difference in hazards for stroke mortality was greater in the least advantaged group (Figure 4-3 and Table S8F, Appendix 7). In SES-stratified analyses in the second study that examined additional outcomes in two cohorts, the difference in hazards associated with combinations of unhealthy *versus* healthy LFs for the total population was greater in the most advantaged group for mortality from 'heart disease' in NHANES but greater in the least advantaged group for coronary heart disease and stroke, cancer, and stroke and myocardial infarction incidence in UKB [293].

4.6 Discussion

Our review shows that the influence of SES on the association between a combination of unhealthy LFs and adverse health outcomes is unclear. There are

several reasons for this. Firstly, few studies investigate this problem; only six studies met our eligibility criteria. Secondly, studies that do investigate this problem are heterogenous, varying by: cohort characteristics; lifestyle, SES, and covariate variables; outcomes assessed; and methodology by which SES influence was examined. Thirdly, where broadly similar estimates were compared directly (*i.e.*, *via* our main comparator), results were mixed: the difference in hazards associated with combinations of unhealthy *versus* healthy LFs was greater in the most advantaged SES group for some studies or cohorts and outcomes but in the least advantaged group for others. Fourthly, results for tests for multiplicative interactions between combinations of LFs and SES were conflicting. For example, for all-cause mortality, two studies found no evidence of multiplicative interaction [290,291]; two studies reported significant multiplicative interactions but observed a moderating influence of SES in opposing directions [24,289]; while a fifth study, of two cohorts, found significant multiplicative interactions in one cohort but not the other [293]. Finally, the quality of included studies varied, with only one scoring the highest possible quality score, so available study estimates may be biased. For example, a limitation to all studies examining SES measures is the difficulty of recruiting participants from less advantaged backgrounds, which introduces selection bias [297,298].

The heterogeneity and nature of the LF exposure variables examined by the included studies warrants further discussion. Firstly, each LF was measured or surveyed differently (e.g., diet assessed via a national dietary index comprising numerous survey items compared with diet assessed by a few specific food items; SES assess). Secondly, studies varied in their definitions of unhealthy (dichotomisation) for the same LF (e.g., unhealthy alcohol intake defined as >14/>7 weekly drinks for men/women vs. drinking daily or almost daily). Thirdly, dichotomising the LF exposure prior to creating the LF score or combination fails to capture the more complex dose-response and non-linear associations LFs have with adverse health-outcomes[299,300]. Fourthly, risk estimates associated with combinations of different LFs are difficult to compare where combinations from different studies lack shared LFs (*e.g.*, combination 1: smoking, alcohol, and physical inactivity *vs.* combination 2: sedentary time, unhealthy diet, and sleep duration). And fifthly, each LF will have differential contributions to the level of risk associated with the overall combination (*e.g.*, smoking is likely to drive the

largest share of risk associated with CVD mortality) [301] , thus making comparisons of estimates associated with unweighted combinations of different LFs hard to interpret. However, because unhealthy LFs are known to cluster among individuals, participants who report the unhealthiest LF combinations might be comparable even when different LFs are examined [15,62,268-270,302]. For example, among participants who report the unhealthiest combination of unhealthy LFs in 'study 1', there will be some participants with similar numbers and types of unhealthy LFs as those in 'study 2' who also report the unhealthiest combination even if study 2 examines fewer LFs because of clustering. This similarity or comparability is more likely where studies share more LF components (e.g., study 1: alcohol, unhealthy diet, and physical inactivity vs. study 2: alcohol, unhealthy diet, physical inactivity, and smoking). Although, in studies that examine more LFs or have more categories for LF analysis, those classified as the unhealthiest may represent a more extreme group. Nevertheless, the aim of this review was to identify and appraise all studies that examined the effect of SES on the association between any LF combination and adverse health. Restricting the searches of this review to identify only those studies with the same or similar combinations of LFs would have yielded even fewer results and limited the synthesis. To explore the effect of specific LF combinations and of SES on the associations with adverse health outcomes, future research could attempt to identify the riskiest LF combinations, whether and how the riskiest combinations vary by SES, and whether and how SES affects the associations between specific combinations and adverse health outcomes. Further, there is currently a lack of guidance on how to live in healthy ways that considers LF combinations and there is no consensus definition for 'unhealthy lifestyle' overall. Numerous single LF specific guidelines exist, but these are often too complex to digest for most people and they often fail to account for interactions with other LFs or social contexts [303-305]. Therefore, as part of precision medicine, future research could explore the non-linear associations and interactions for a wide range of LFs to define 'unhealthy' levels for LFs within specific (or personalised) combinations and across SES the spectrum. These efforts could provide new targets for intervention and inform policies attempting to address unhealthy LFs in the least advantaged sections of society [306] .

The range of SES measures used across studies highlights the myriad ways in which SES can be measured and ranked [130]. Although there is likely to be a high degree of correlation across SES measures, the impacts of different SES measures on the association between combinations of LFs and adverse health could be different [131]. For example, an individual-level measure (*e.g.*, age at last formal education) could have a weaker modifying effect on the association between combinations of LFs and adverse health than an area-based deprivation index if wider socioeconomic factors included or captured by the index (directly or indirectly) have a greater effect on the association. For instance, proximity and access to healthy food or green spaces for PA could be more strongly associated with area-based SES indices than with individual-level SES measures [307]. Having few studies using the same SES measure limited the ability to draw conclusions on how the SES measure influences SES effects. For example, of the six studies that examined all-cause mortality, two use income, one uses education, two use area-based indices, and one uses a combination of income, education level, occupation/employment, and health insurance in a latent class analysis. Future reviews, with a greater number of included studies, could stratify and synthesise results by SES measure to investigate this further. Irrespective of SES heterogeneity, if an effect of SES was identified that was consistent across a broad range of SES measures this would strengthen the evidence for a general SES effect. Whereas if SES effects were consistently associated with one type of SES measurement (*e.g.*, income) and not others (*e.g.*, area-based indices) this could generate hypotheses and inform research that aims to explain underlying mechanisms of SES effects [131]. The aim of this review was to identify all available evidence and therefore studies were not excluded on the basis of LF and SES exposure variables despite the expected difficulties in comparability.

Notwithstanding study heterogeneity and the lack of data, the studies' assessments of the influence of SES on the association between a combination of unhealthy LFs and adverse health outcomes point broadly towards an additive influence of SES. Examining the combined effect of SES and combinations of unhealthy LFs by way of a single reference group (the healthiest LF-most advantaged group), four studies of five cohorts provide evidence for an additive interaction for multiple outcomes [24,290,291,293]. Two of these studies, both

examining UKB, also observed significant results from formal tests for additive interactions as well as significant multiplicative interactions in same direction [24,293]. Together, this evidence does not strongly support a vulnerability hypothesis but it does provide some evidence against the so-called Blaxter hypothesis [308] . The Blaxter hypothesis suggests that detrimental effects of unhealthy lifestyles are masked by other adverse factors also associated with less advantaged SES (*e.g.*, insecure income, poor housing, more frequent adverse childhood experiences). If this hypothesis were correct, in analyses stratified by SES and in least advantaged SES groups, associations between combinations of LFs and adverse health would be similar whether the LFs were healthy or unhealthy (*i.e.*, a combination of unhealthy LFs would have little influence on a population with an already high risk due to other factors). However, in all studies, compared to those with healthy LFs, there were higher hazards for adverse health outcomes in those with a combination of unhealthy LFs irrespective of SES level. One study observed a multiplicative interaction (in men only), where the difference in hazards associated with a combination of unhealthy *versus* healthy LFs was greater in the most advantaged SES group, which could support the Blaxter hypothesis [289]. However, the authors did not report a single reference group analysis, which could help clarify the combined associations. Overall, the impression of an additive interaction between least advantaged SES and combinations of unhealthy LFs seen in four studies of five cohorts and a multiplicative interaction in the same direction in two studies suggests that the detrimental effects of combinations of unhealthy LFs are not masked by other harmful factors associated with less advantaged SES but are at least in addition to, and potentially synergistic with, those factors. This finding, if borne out in future research, would indicate that less advantaged SES populations have the highest absolute risks associated with combinations of unhealthy LFs and would, therefore, support a strategy of focussing lifestyle resources on less advantaged SES populations where need is greatest.

4.6.1 Strengths and limitations

This review is strengthened by a rigorous pre-specified protocol [283]; a comprehensive search strategy including database, reference, citation, and grey literature searches [31]; and by reviewers working independently. Further, data synthesis follows SWiM guidelines and is fully transparent [32]. However, this

review is limited by the small number of studies included and by the high level of heterogeneity between studies, which precluded meta-analysis. Therefore, the conclusions drawn here about whether and how SES influences the association between combinations of unhealthy LFs and adverse health may be altered by future research. Importantly, differential vulnerability to combinations of unhealthy LFs could be due to differential exposure that is not captured *via* questionnaires. For example, excess alcohol in less advantaged SES populations may be more extreme than excess alcohol in more advantaged groups [191] . Similarly, residual confounding, with unaccounted for differences between more and less advantaged populations, could also explain observed differential vulnerability. Our search terms were extensive and the databases we searched likely contained the vast majority of eligible articles [309]. Searching additional databases such as Scopus and Web of Science, which are noted for their use as citation indexes rather than primary sources [310], may have revealed additional eligible studies but this was beyond the time and human resources available for this project. Updating the searches could also identify newer studies, including studies that look at different health outcomes, such as specific cancers and studies from low- or middle-income countries [311,312]. It is unlikely that the addition of these studies would change the result around the heterogeneity of data but it could reveal more consistency in terms of SES effects. Our eligibility criteria may have been too restrictive resulting in few studies and retrospective studies may have yielded additional evidence. Generally, however, retrospective observational studies may have more biases than well-designed prospective ones. Future prospective studies, where data are updated during follow-up, could reduce potential misclassification bias by capturing participants' lifestyle changes. While the adverse health outcomes included here account for the vast majority of mortality and NCD burden [313], others, such as dementia and renal disease, are growing in prevalence and have similar lifestyle risk factors [314,315]. Our decision to extract the 'healthiest' vs. 'unhealthiest' in both the most and least advantaged groups as our main comparator may have limited our synthesis. Examining the effect of SES on associations between the extremes of lifestyle and adverse health outcomes may miss how SES might affect the relationships in more nuanced ways as it relies on the assumption that SES effects will be seen at the extremes of lifestyle. However, we also extracted results for interactions between SES and LFs, which

provided further evidence for whether and how SES effects associations between LFs and adverse health outcomes. As more studies examine associations between combinations of LFs and adverse health outcomes in more detail (e.g., by examining non-linear associations and using continuous rather than categorical or ordinal variables for LF combinations) [299,300], future reviews could examine how SES effects the shape of relationships between LFs and adverse health outcomes. The aim of this review was to identify and synthesise the evidence for SES modification of associations between LF combinations and adverse health outcomes, not to explain any identified effect modification. However, strong evidence for SES effect modification of such associations could prompt attempts to uncover underlying mechanisms, such as cumulative risks or accelerated biological ageing [185,186,275].

4.7 Conclusions

This is the first systematic review to examine if and how SES modifies associations between combinations of unhealthy LFs and adverse health outcomes. Prospective studies that examine this problem are few and heterogenous. The influence of SES on lifestyle-associated adverse health could be additive but remains unclear. New research using multiple datasets, a range of lifestyle and SES measures, and a comprehensive list of adverse health outcomes would improve understanding of SES influence on lifestyle risks and thereby inform lifestyle-related policy and interventions.

5 Social connection and mortality in UK Biobank: a prospective cohort analysis.

5.1 Chapter summary

This chapter provides an analysis of the associations and interactions between measures of social connection and adverse health outcome. This chapter addresses research objective 2), which was to examine the associations and interactions between measures of social connection and adverse health outcomes to help determine which measures of social connection could be included in a broad measurement of lifestyle i.e., a lifestyle score.

Other than some brief additional text regarding test for interaction and the final two paragraphs of the '5.6.3 Comparison with wider literature' section, which offer additional explanations for why the association with mortality was observed to be stronger for structural than functional isolation, the text and figures are presented as published in: Foster, H.M.E., Gill, J.M.R., Mair, F.S. et al. Social connection and mortality in UK Biobank: a prospective cohort analysis. BMC Med 21, 384 (2023).

The related additional files are given in Appendix 8 and Appendix 9 or are available via https://doi.org/10.1186/s12916-023-03055-7.

5.2 Abstract

Background: Components of social connection are associated with mortality, but research examining their independent and combined effects in the same dataset is lacking. This study aimed to examine the independent and combined associations between functional and structural components of social connection and mortality.

Methods: Analysis of 458,146 participants with full data from the UK Biobank cohort linked to mortality registers. Social connection assessed using two functional (frequency of ability to confide in someone close and often feeling lonely) and three structural (frequency of friends/family visits, weekly group activities, and living alone) component measures. Cox proportional hazard models used to examine associations with all-cause and cardiovascular disease (CVD) mortality.

Results: Over median 12.6 years (IQR 11.9 – 13.3) follow-up, 33,135 (7.2%) participants died, including 5,112 (1.1%) CVD deaths. All social connection measures were independently associated with both outcomes. Friends/family visit frequencies < monthly were associated with higher risk of mortality indicating a threshold effect. There were interactions between living alone and friends/family visits and between living alone and weekly group activity. For example, compared with daily friends/family visits-not living alone, there was higher all-cause mortality for daily visits-living alone (HR 1.19 [95% CI 1.12-1.26]), for never having visits-not living alone (1.33 [1.22-1.46]), and for never having visits-living alone (1.77 [1.61-1.95]). Never having friends/family visits whilst living alone potentially counteracted benefits from other components as mortality risks were highest for those reporting both never having visits and living alone regardless of weekly group activity or functional components. When all measures were combined into overall functional and structural components there was an interaction between components: compared with participants defined as not isolated by both components, those considered isolated by both components had higher CVD mortality (HR 1.63 [1.51-1.76]) than each

component alone (functional isolation 1.17 [1.06-1.29]; structural isolation 1.27 [1.18-1.36]).

Conclusions: This work suggests 1) a potential threshold effect for friends/family visits, 2) that those who live alone with additional concurrent markers of structural isolation may represent a high risk population, 3) that beneficial associations for some types of social connection might not be felt when other types of social connection are absent, and 4) considering both functional and structural components of social connection may help to identify the most isolated in society.

5.3 Background

Social connection is a complex phenomenon that encompasses numerous emotional, physical, and behavioural aspects of human interaction. Social connection can be classified into inter-related conceptual components, including functional (e.g., subjective feelings of loneliness) and structural (e.g., objective frequency of social contacts) components.[90,316] Deficits of either component are associated with higher risk of all-cause mortality and cardiovascular disease (CVD).[95,97,100,222,317] The mechanisms by which components of social connection are associated with mortality are unclear and may vary by component, or by the measure used, but are thought to be mediated via direct (e.g., altered blood pressure, poorer immune function, neurodevelopmental impairment)[85,318,319] and indirect effects (e.g., via poorer mental health or wellbeing, lower physical activity or higher tobacco and alcohol consumption).[92,320-323] Further explanations involve reverse causality, whereby long term health conditions or disabilities can impair people's ability to form or sustain relationships.[324,325] Nevertheless, the prevalence of a lack of social connection (9.2%-14.4% of the global population are estimated to feel lonely and 25% of adults worldwide may be socially isolated)[98,99] and the associated mortality justify attempts to understand how each component impacts on mortality in order to develop targeted interventions (Table 5-1).

Table 5-1: Components of social connection
(Taken from Holt-Lundstad, 2018).[316]

Component	Definition	Example measures
Functional	Functions provided by, or perceived to be available because of, social relationships	Received support Perceptions of social support Perceived loneliness
Structural	Existence of and interconnections among different social relationships and roles	Marital status Living alone or not Social networks Social integration Social isolation

Often, only moderate or weak correlations are observed between different components, which may reflect the dependent and independent relationships between them.[326-329] For example, individuals with a shrinking social network might feel lonelier as a result, while others with a growing social network could also feel increasingly lonely if the quality of those relationships is poor. Adding to the challenge of understanding how different components of social connection are associated with adverse health outcomes is the numerous heterogenous ways by which studies have operationalised and measured different aspects of each component.[89,92,94,103,330] Prior studies have often focussed on a single item measure, for example, showing that a 'sense of loneliness' (functional) or living alone (structural) are independently associated with higher risk of all-cause mortality.[331,332] Alternatively, some studies have used composite scales or indices but still with a focus on a single component of social connection (e.g., Revised UCLA Loneliness Scale measuring the functional component or the Berkman-Syme Social Network Index measuring the structural component).[333,334] A meta-analysis of prospective studies examining the association between both subjective (functional) or objective (structural) isolation and all-cause mortality found the average effect sizes to be similar (26-32% increased likelihood of mortality) for each type of isolation.[317] However, the effect sizes represent aggregate effects of different measures with no consideration of the strength of association of individual measures on health outcomes. Further, meta-analyses that have quantified associations between measures of social connection and mortality, have highlighted the lack of studies that include measures of both functional and structural components or that examine for potential synergistic interactions between them.[97,100,317] Indeed a lack of research examining different components of social connection in the same dataset to disentangle their independent, additive, and multiplicative effects was highlighted in a recent U.S. Surgeon General's Advisory.[19] These are missed opportunities, as more detailed understanding of the health impact of different components of social connection and their interactions could help guide policy and interventions designed to increase and enhance social connectedness and improve related health outcomes.

Previous studies often refer to functional and structural components of social connection as loneliness and social isolation, respectively. However, the social connection framework offers advantages for conceptualising and studying both the separate and combined effects of different social measures. First, the terminology of social connection is more neutral than 'loneliness' or 'isolation' and thereby implies a spectrum of either beneficial or detrimental associations. Second, studies often lack methodological or theoretical underpinnings and use loneliness and social isolation as catch all phrases, which while supposedly 'widely understood' may be interpreted differently depending on the researcher. Third, social connection offers a broad framework that encapsulates both loneliness and social isolation alongside but separately from other measures of social connection. For example, a subjective feeling of loneliness can be considered a measure of the functional component but so too can perceptions of social support, which may not always be perceived as loneliness. Therefore, the terminology of the social connection framework remains flexible and inclusive whilst avoiding some assumptions around loneliness or social isolation and, as a result, could help when interpreting estimates of the separate and combined effects of various social measures.

The first aim of this study was to understand the strength of association between independent measures of functional and structural social connection and all-cause and CVD mortality. The second aim was to understand if and how these measures interact with one another in combined associations with adverse health outcomes. Our study was guided by the following research questions (RQ):

RQ1. What is the strength of association between two functional measures of social connection – frequency of ability to confide and perceived loneliness – and all-cause and CVD mortality, and is there an interaction between the measures?

RQ2. What is the strength of association between three structural measures of social connection – frequency of friends and family visits, weekly leisure/social activities, and living alone – and all-cause and CVD mortality, and is there an interaction between these measures?

RQ3. What is the pattern of the combined association between measures of a) functional and b) structural components of social connection and all-cause and CVD mortality?

RQ4. Is there an interaction between functional and structural components of social connection for all-cause and CVD mortality?

5.4 Methods

5.4.1 Study design and participants

We analysed baseline data from the UK Biobank study which recruited 502,536 participants via postal invitation between 2006-2010. Participants attended one of 22 assessment centres in England, Scotland, or Wales to complete a questionnaire, nurse-led interview, and have physical measurements taken.[246] More details of UK Biobank procedures and assessments can be found online (biobank.ndph.ox.ac.uk/ukb/) and in the study protocol.[335] We excluded those without full data on all variables used in analyses (n=44,390 [8.8%]) as detailed below (participant flowchart – Additional file 1: Fig. S1). Participants who reported 'do not know', or 'prefer not to answer' for any variable were considered missing.

5.4.2 Outcome ascertainment

UK Biobank participants consented to data linkage to national mortality registers. We examined two adverse health outcomes: all-cause and CVD mortality. Any International Classification of Diseases (10th Revision) code from I05-I99, Z86.7, G45, and G46 given as the primary cause of death were chosen to define CVD deaths after discussion by two primary care clinicians (HMEF and FSM). These codes likely represent chronic CVD diseases, including cerebrovascular disease, with only acute rheumatic fever (I00-I02) excluded. Dates and causes of death are contained within death certificates provided by linkage to the National Health Service (NHS) Information Centre (England and Wales) and the NHS Central Register (Scotland). Censoring dates varied by country of baseline assessment (England and Wales, 30 September 2021; Scotland, 31 October 2021).

5.4.3 Functional and structural component measures

Two functional and three structural component measures of social connection used in previous studies were examined in this study (Table 5-2).[221-223,329] For frequency of friend and family visits, the categories of 'never or almost never'

and ‘no friends or family outside household’ were collapsed into a single category, ‘never’. This was justified on the basis that these responses are similar and there being few participants with no friends or family outside the household (n= 1,031). For simplicity, categories for ordinal variables were renamed as: ‘daily’, ‘2-4 times a week’, ‘weekly’, ‘monthly’, ‘once every 3 months’, and ‘never’.

Table 5-2: Functional and structural component measures and categories

Component	Measure	Categories
Functional	Frequency of ability to confide in someone close	daily, 2-4 times a week, weekly, monthly, once every 3 months, and never
	Often feeling lonely	yes, no
Structural	Frequency of friends and family visits	daily, 2-4 times a week, weekly, monthly, once every 3 months, and never
	Weekly group activity	yes, no
	Living alone	yes, no

5.4.4 Covariate data

Baseline self-reported sex (female, male), ethnicity (White, Mixed, Asian or Asian British, Black or Black British, Chinese, or Other ethnic group), smoking status (current, never/former), alcohol intake (> vs ≤ 35 [females] and > vs. ≤ 50 [males] weekly units of alcohol – previously identified cut offs high risk drinking in England and UK Biobank),[232,336] self-reported physical activity levels (< vs. ≥450 MET [metabolic equivalent of task] minutes per week as per UK physical activity guidelines)[337,338] were used as potential explanatory variables. A count of baseline self-reported long-term conditions confirmed at nurse-led interview was based on a list of 43 long-term conditions.[339] Month of assessment was included as a covariate as self-reported measures of social connection may vary by season.[340] Socioeconomic position was measured using the area-based measure of deprivation, Townsend index (comprising car ownership, household overcrowding, owner occupation, and unemployment) and was based on preceding census data and postcode of residence at recruitment and analysed as a continuous variable.[341] Body mass index (BMI) was calculated by trained personnel at baseline assessment and used as a continuous measure (kg/m^2).

5.4.5 Statistical analysis

We compared those participants with complete data to those with missing data using descriptive statistics. For our main analyses, we used time to event analysis (Cox proportional hazard models) to examine the associations between exposures and mortality outcomes for those participants with full data only. Follow up time was calculated as the time difference between date of assessment and either censor date or date of death, whichever occurred first. Table S1 (Appendix 8) shows the analyses performed and the corresponding research question each analysis addresses. Measures of social connection and the covariates included in our models may be highly correlated and lead to multicollinearity and model instability.[342] Therefore to detect potential multicollinearity, we calculated generalised variance inflation factors (GVIF) for all variables included in our Cox models using a linear regression model with follow-up time as the outcome.[343]

5.4.5.1 Functional component analyses

First, we examined the association between each functional component measure (frequency of ability to confide in someone close and often feeling lonely) and adverse health outcomes separately, adjusting for the known and likely confounders: sex, ethnicity, Townsend index, and month of assessment, smoking status, alcohol intake, physical activity, BMI, long-term condition count, frequency of friend and family visits, weekly group activity, living alone, and mutually for frequency of ability to confide/often feeling lonely (analyses 1 & 2, Table S1, Appendix 8). Next, we examined the combined association of both functional component measures (with a single reference group of almost daily ability to confide in someone close and not often feeling lonely) and their interactions for adverse health outcomes (analysis 3, Table S1, Appendix 8). Full information is often lacking in study reports that would allow readers to assess the size and significance of interaction on both multiplicative and additive scales.[344] To provide sufficient data for interpretation, we explored interactions on both multiplicative and additive scales by calculating estimates for multiplicative interaction, relative excess risk of interaction (RERI), attributable portion (AP), and a synergy index (SI).[344] RERI, AP, and SI each provide

assessment of additive interaction of two binary exposures variables. RERI is often considered the standard measure for interaction on the additive scale. While others argue for the SI, which is the ratio of the combined and individual effects of two exposures, as this measure does not require both exposures to be risk factors for the outcome (whereas the other measures do). The AP provides different information and estimates the proportion of disease among those with both exposures that is attributable to their interaction. A RERI or AP > 0 means positive interaction while < 0 means negative interaction or less than additivity (and 0 means no interaction). An SI > 1 indicates an interaction that is greater than additivity and an SI < 1 indicates negative interaction or less than additivity. The interaction tests require four exposure groups which meant dichotomising the ordinal variable of frequency of ability to confide in someone close. To inform dichotomisation, we used results from the independent and mutually adjusted association between the ordinal variable and adverse health outcomes. Therefore, the dichotomous confide variable was coded as (≥ once every 3 months vs never). To examine interactions between functional and structural components we created a new dichotomous 'functional isolation' variable. Functional isolation was defined using the independent and mutually adjusted associations with adverse health outcomes of each functional component measure and therefore coded as either never able to confide or (yes) often feeling lonely. We examined the associations between this new variable and adverse health outcomes (analysis 4, Table S1, Appendix 8).

5.4.5.2 Structural component analyses

Next, we examined the association between each of the structural component measures (frequency of friends and family visits, weekly group activity, and living alone) and adverse health outcomes separately, with models adjusted as above but with mutual adjustment for each structural measure and the new functional isolation variable (analyses 5-7, Table S1, Appendix 8). Then, to examine joint associations and interactions between the structural component measures we examined associations and interactions between 1) frequency of friends and family visits and engagement in weekly group activity, 2) frequency of friends and family visits and living alone, and 3) weekly group activity and

living alone (analyses 8-9 & 11, Table S1, Appendix 8). To examine interactions, we dichotomised the ordinal variable of frequency of friends fand family visits as ≥ monthly/< monthly based on its independent and mutually adjusted associations with adverse health outcomes. Where there was evidence for interaction, we also examined stratified associations (analyses 10 & 12, Table S1, Appendix 8). We then combined the three structural component measures into a new dichotomous 'structural isolation' variable, coding structural isolation as having less than monthly friends and family visits or no weekly group activity or living alone. We examined the association between this new variable and adverse health outcomes (analysis 13, Table S1, Appendix 8).

5.4.5.3 Functional and structural components together

To examine the combined effect of functional and structural measures together and to assess the impact on any dose response relationship, we examined associations between 1) frequency of ability to confide in someone close, often feeling lonely, structural isolation, and adverse health outcomes (analysis 14, Table S1, Appendix 8); and 2) frequency of friends and family visits, weekly group activity, living alone, functional isolation, and adverse health outcomes (analysis 15, Table S1, Appendix 8). Finally, we examined the combined associations and interaction between the two new overall functional and structural isolation variables and adverse health outcomes (analysis 16, Table S1, Appendix 8).

5.4.5.4 Sensitivity analyses

Accounting for participants' prior health status is critical for estimating the associations between social connection and adverse health.[317] To reduce the chance that findings could be explained by reverse causality (e.g., where poor baseline health status would explain both reduced social connectedness and higher mortality), we repeated all analyses after excluding all those who reported having CVD (diabetes, coronary heart disease, atrial fibrillation, chronic heart failure, chronic kidney disease, hypertension, stroke/transient ischaemic attack, or peripheral vascular disease) or cancer at baseline as well as

those who died within two years of recruitment (analysis 17, Table S1, Appendix 8).

All analyses were conducted using R statistical software version 4.2.0.

5.5 Results

5.5.1 Descriptive statistics

44,390 (8.8%) participants with missing data were excluded. Compared with those with complete data, participants with missing data were more likely to be male, older, from minority ethnic backgrounds, have been assessed in spring or summer months (April-September), be from more deprived areas, be current smokers, have low physical activity levels, have a higher BMI, and have more long-term conditions (Table S2, Appendix 8). After excluding those without full data, 458,146 (91.2%) UK Biobank participants were included in the main analyses. The mean age of participants was 56.5 years (standard deviation [SD] 8.1; range 38-73), 54.7% were women, and 95.5% were of white ethnicity (Table 5-3). Generally, compared to all participants, those reporting any measure of reduced social connection were more likely to: be from a minority ethnic background, be more deprived, engage in more unhealthy behaviours (smoking, high alcohol intake, and low physical activity levels), have a higher BMI, and have more long-term conditions. Of those who reported each measure of reduced social connection there was variation in the percentage who were female: often feeling lonely (62.9% women); not engaging in weekly group activities (55.1% women); living alone (58.5% women); never able to confide in someone close (40.9% women); and friend and family visits less than monthly (42.0% women). GVIF values, calculated to detect multicollinearity, ranged from 1.00 to 1.16 and were well below the proposed threshold of 10 (Table S3, Appendix 8).[342] This reduced the concern of multicollinearity and strengthened the argument for including all the social connection measures as separate variables in the models.

Table 5-3: Descriptive characteristics of study participants by measures of functional and structural components of social connection

	Functional component measures		Structural component measures			Total
	Never able to confide in someone close	Often feels lonely	Friends and family visits less than monthly	Does not engage in weekly group activities	Lives alone	
N	66,638	83,915	37,580	137,801	84,472	458,146
Female	27,285 (40.9%)	52,818 (62.9%)	15,775 (42.0%)	75,912 (55.1%)	49,409 (58.5%)	250,761 (54.7%)
Age	57.3 (7.9)	55.5 (8.0)	55.4 (7.9)	56.0 (8.0)	57.8 (7.9)	56.5 (8.1)
Ethnicity						
White	63,067 (94.6%)	78,476 (93.5%)	34,328 (91.3%)	131,333 (95.3%)	80,670 (95.5%)	437,462 (95.5%)
Mixed	424 (0.6%)	653 (0.8%)	305 (0.8%)	890 (0.6%)	622 (0.7%)	2,646 (0.6%)
Asian*	1,166 (1.7%)	1,866 (2.2%)	872 (2.3%)	2,412 (1.8%)	717 (0.8%)	6,931 (1.5%)
Black†	1,181 (1.8%)	1,748 (2.1%)	1,198 (3.2%)	1,625 (1.2%)	1,615 (1.9%)	6,499 (1.4%)
Chinese	233 (0.3%)	202 (0.2%)	248 (0.7%)	452 (0.3%)	161 (0.2%)	1,148 (0.3%)
'Other'	567 (0.9%)	970 (1.2%)	629 (1.7%)	1,089 (0.8%)	687 (0.8%)	3,460 (0.8%)
Month of assessment						
January	4,831 (7.2%)	5,693 (6.8%)	2,600 (6.9%)	9,858 (7.2%)	5,834 (6.9%)	32,468 (7.1%)
February	5,655 (8.5%)	6,880 (8.2%)	3,009 (8.0%)	11,285 (8.2%)	7,086 (8.4%)	37,992 (8.3%)
March	6,638 (10.0%)	8,329 (9.9%)	3,657 (9.7%)	13,711 (9.9%)	8,520 (10.1%)	45,314 (9.9%)
April	5,842 (8.8%)	7,305 (8.7%)	2,986 (7.9%)	11,843 (8.6%)	7,189 (8.5%)	39,690 (8.7%)
May	6,792 (10.2%)	8,879 (10.6%)	3,826 (10.2%)	14,199 (10.3%)	8,886 (10.5%)	46,858 (10.2%)
June	6,697 (10.0%)	8,752 (10.4%)	3,850 (10.2%)	14,259 (10.3%)	8,951 (10.6%)	46,677 (10.2%)
July	5,661 (8.5%)	7,359 (8.8%)	3,329 (8.9%)	12,063 (8.8%)	7,230 (8.6%)	38,956 (8.5%)
August	4,927 (7.4%)	6,373 (7.6%)	2,788 (7.4%)	10,592 (7.7%)	6,479 (7.7%)	34,372 (7.5%)
September	4,601 (6.9%)	5,961 (7.1%)	2,666 (7.1%)	9,688 (7.0%)	5,826 (6.9%)	32,942 (7.2%)
October	5,501 (8.3%)	7,016 (8.4%)	3,277 (8.7%)	11,378 (8.3%)	6,822 (8.1%)	38,783 (8.5%)
November	5,576 (8.4%)	6,861 (8.2%)	3,294 (8.8%)	11,220 (8.1%)	6,976 (8.3%)	38,202 (8.3%)
December	3,917 (5.9%)	4,507 (5.4%)	2,298 (6.1%)	7,705 (5.6%)	4,673 (5.5%)	25,892 (5.7%)

Winter assessment‡	32,118 (48.2%)	39,286 (46.8%)	18,135 (48.3%)	65,157 (47.3%)	39,911 (47.2%)	218,651 (47.7%)
Townsend index	-0.99 (3.27)	-0.63 (3.34)	-0.77 (3.36)	-1.14 (3.14)	-0.03 (3.43)	-1.39 (3.04)
Current smoker	9,035 (13.6%)	12,673 (15.1%)	4,983 (13.3%)	18,510 (13.4%)	13,083 (15.5%)	47,234 (10.3%)
High alcohol intake	6,747 (10.1%)	7,625 (9.1%)	3,900 (10.4%)	9,238 (6.7%)	8,400 (9.9%)	41,125 (9.0%)
Low physical activity	15,755 (23.6%)	20,149 (24.0%)	8,778 (23.4%)	37,802 (27.4%)	16,946 (20.1%)	89,942 (19.6%)
BMI, kg/m^2	28.0 (4.96)	28.0 (5.39)	27.5 (5.05)	27.7 (5.09)	27.6 (5.20)	27.4 (4.78)
Number of Long-term conditions	1.32 (1.31)	1.51 (1.41)	1.24 (1.27)	1.28 (1.30)	1.40 (1.35)	1.20 (1.23)

Figures given are N (column %) or mean (SD); *Asian or Asian British; †Black or Black British. ‡October-March; Higher Townsend index equates to higher levels of deprivation; High alcohol intake, > 35 (females) and > 50 (males) weekly units of alcohol; Low physical activity, <450 MET minutes per week; BMI, body mass index.

5.5.2 Association with adverse health outcomes

After a median follow-up of 12.6 years (IQR 11.9 – 13.3) there were 33,135 (7.2%) deaths, of which 5,112 (1.1%) were CVD deaths.

5.5.3 Functional component measures – independent associations

Models of the association between the frequency of the ability to confide in someone close and outcomes showed that participants who reported never being able to confide were associated with higher all-cause and CVD mortality compared with the reference group of those who reported being able to confide daily: HR 1.07 (95% CI 1.03-1.10) and 1.17 (1.09-1.26), respectively (Table 5-4 & Figure 5-1). Indeed, for both outcomes, there were no substantial differences in effect sizes across all categories of frequency in ability to confide in someone close apart from never able to confide. Models of the association between often feeling lonely and outcomes showed that compared to those who reported not often feeling lonely, those often feeling lonely were also associated with higher all-cause and CVD mortality: HR 1.06 (1.03-1.09) and 1.08 (1.00-1.16) (Table 5-4 & Figure 5-1).

Table 5-4: Models of association between functional component measures and all-cause and CVD mortality.

Outcome	Measure	N	Deaths (%)	HR	LCI	UCI
All-cause mortality	***Frequency of ability to confide in someone close***					
	Daily	246,851	16,588 (6.7%)	1 (ref)	-	-
	2-4 times a week	44,267	2,787 (6.3%)	0.99	0.95	1.03
	Weekly	50,320	3,556 (7.1%)	1.00	0.96	1.04
	Monthly	24,403	1,766 (7.2%)	1.01	0.96	1.06
	Once every 3 months	25,667	1,893 (7.4%)	0.99	0.94	1.03
	Never	66,638	6,545 (9.8%)	1.07	1.03	1.10
	Often feels lonely					
	No	374,231	26,182 (7.0%)	1 (ref)	-	-
	Yes	83,915	6,953 (8.3%)	1.06	1.03	1.09
	Functional isolation*					
	No	329,312	21,831 (6.6%)	1 (ref)	-	-
	Yes	128,834	11,304 (8.8%)	1.08	1.06	1.11
CVD mortality	***Frequency of ability to confide in someone close***					
	Daily	246,851	2,425 (1.0%)	1 (ref)	-	-
	2-4 times a week	44,267	380 (0.9%)	0.96	0.86	1.07
	Weekly	50,320	504 (1.0%)	0.99	0.90	1.09
	Monthly	24,403	272 (1.1%)	1.06	0.93	1.20
	Once every 3 months	25,667	300 (1.2%)	1.05	0.93	1.18
	Never	66,638	1,231 (1.8%)	1.17	1.09	1.26
	Often feels lonely					
	No	374,231	3,932 (1.1%)	1 (ref)	-	-
	Yes	83,915	1,180 (1.4%)	1.08	1.00	1.16
	Functional isolation*					
	No	329,312	3,140 (1.0%)	1 (ref)	-	-
	Yes	128,834	1,972 (1.5%)	1.16	1.09	1.23

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, frequency of friend and family visits, weekly group activity, living alone, and mutually for frequency of ability to confide in someone close and often feels lonely. *Functional isolation defined as never able to confide in someone close or often feels lonely for which models were adjusted as above but without adjusting for frequency of ability to confide in someone close or often feels lonely. HR, hazard ratio; LCI, lower confidence interval; UCI, upper confidence interval.

5.5.4 Functional component measures – combined associations and interactions

Models examining combined associations (Table S4, Appendix 8) and interactions (Table S5, Appendix 8) between frequency of ability to confide and often feeling lonely for adverse health outcomes did not provide clear evidence for interaction on either multiplicative or additive scales. Based on the pattern of their independent and mutually adjusted associations across both outcomes (Table 5-4), we combined both measures into a new dichotomous functional isolation variable, with isolation coded as reporting either never able to confide, often feeling lonely, or both. Compared to those with no functional isolation (self-reporting able to confide at least every 3 months and not often lonely), participants with functional isolation were associated with higher all-cause and CVD mortality: HR 1.08 (1.06-1.11) and 1.16 (1.09-1.23) (Table 5-4 & Figure 5-1).

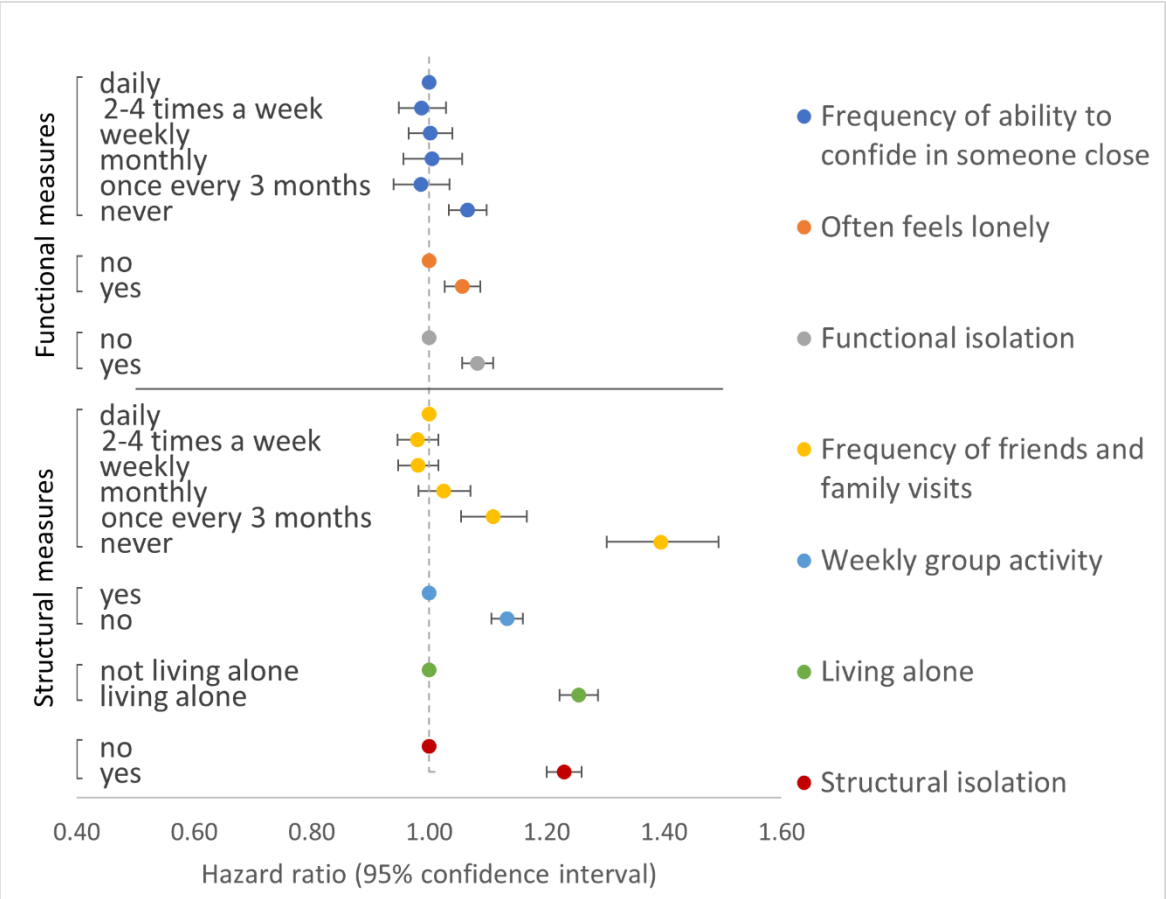


Figure 5-1: Models of association between functional and structural component measures and all-cause mortality.
Models were adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, and mutually for each of the functional and structural component measures. Functional isolation defined as either never able to confide or often feeling lonely. Structural isolation defined as having <monthly friends and family visits or not engaging in weekly group activity or living alone.

5.5.5 Structural component measures – independent associations

Fully adjusted models of associations between the frequency of friends and family visits and all-cause mortality showed that participants who reported visits with friends and family less often than once a month were associated with substantially higher risk of all-cause mortality: HRs (95% CI) for once every 3 months and never were 1.11 (1.05-1.17) and 1.39 (1.30-1.49), respectively (Table 5-5 & Figure 5-1). The same pattern was observed for CVD mortality but with stronger associations and wider confidence intervals (Table 5-5). Compared with those who reported engaging in weekly group activity, those who reported not engaging in weekly group activity had higher all-cause and CVD mortality: HRs (95% CIs) were 1.13 (1.11-1.16) and 1.10 (1.04-1.17), respectively (Table 5-5

& Figure 5-1). Equivalent estimates for those who reported living alone, compared with those who lived with at least one other, were 1.25 (1.22-1.29) and 1.48 (1.38-1.57) (Table 5-5 & Figure 5-1).

Table 5-5: Models of association between social isolation measures and all-cause and CVD mortality.

Outcome	Measure	N	Deaths (%)	HR	LCI	UCI
All-cause mortality	***Frequency of friends and family visits***					
	Daily	53,581	4,548 (8.5%)	1 (ref)		
	2-4 times a week	141,881	10,491 (7.4%)	0.98	0.95	1.02
	Weekly	163,720	10,693 (6.5%)	0.98	0.95	1.02
	Monthly	61,384	4,021 (6.6%)	1.02	0.98	1.07
	Once every 3 months	30,026	2,327 (7.7%)	1.11	1.05	1.17
	Never	7,554	1,055 (14.0%)	1.39	1.30	1.49
	Engages in weekly group activity					
	Yes	320,345	22,047 (6.9%)	1 (ref)		
	No	137,801	11,088 (8.0%)	1.13	1.11	1.16
	Lives alone					
	No	373,674	24,228 (6.5%)	1 (ref)		
	Yes	84,472	8,907 (10.5%)	1.25	1.22	1.29
	Structural isolation*					
	No	242,570	14,952 (6.2%)	1 (ref)	-	-
	Yes	215,576	18,183 (8.4%)	1.23	1.20	1.26
CVD mortality	**Frequency of friends and family visits**					
	Daily	53,581	694 (1.3%)	1 (ref)		
	2-4 times a week	141,881	1,524 (1.1%)	0.95	0.86	1.04
	Weekly	163,720	1,614 (1.0%)	0.95	0.87	1.04
	Monthly	61,384	627 (1.0%)	0.99	0.89	1.11
	Once every 3 months	30,026	418 (1.4%)	1.16	1.03	1.32
	Never	7,554	235 (3.1%)	1.53	1.32	1.78
	Engages in weekly group activity					
	No	320,345	3,367 (1.1%)	1 (ref)		
	Yes	137,801	1,745 (1.3%)	1.10	1.04	1.17
	Lives alone					
	No	373,674	3,547 (0.9%)	1 (ref)		
	Yes	84,472	1,565 (1.9%)	1.48	1.38	1.57
	Structural isolation*					
	No	242,570	2,139 (0.9%)	1 (ref)	-	-
	Yes	215,576	2,973 (1.4%)	1.35	1.28	1.43

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, new dichotomous loneliness variable – never or almost never able to confide in someone close OR often feeling lonely, and mutually for frequency of friend and family visits, weekly group activity, and living alone. *Structural isolation defined as friends and family visits <monthly or no weekly group activity or living alone for which models were adjusted as above but without adjusting for frequency of friends and family visits, weekly group activity, or living alone. HR, hazard ratio; LCI, lower confidence interval; UCI, upper confidence interval.

5.5.6 Structural component measures – combined associations and interactions

5.5.6.1 Frequency of friends and family visits and weekly group activity

Models of combined associations between frequency of friends and family visits and weekly group activity (reference group of daily friends and family visits and engaging in weekly group activity) showed higher all-cause mortality associated with never having friends and family visits irrespective of whether participants reported engaging in weekly group activity (HR 1.50 [1.37-1.64]) or not (HR 1.49 [1.36-1.65]) (Figure 5-2 & Table S6, Appendix 8). A similar pattern was present for CVD mortality (Figure 5-2 & Table S6, Appendix 8). There was a lack of evidence for an interaction between the two exposures of friend and family visit frequency (≥monthly versus <monthly) and weekly group activity for both all-cause and CVD mortality (Table S7, Appendix 8).

Figure 5-2: Models of combined associations between frequency of friends and family visits, weekly group activity or living alone, and all-cause (ACM) or CVD mortality (CVDM).
Models of combined associations between frequency of friends and family visits, weekly group activity or living alone, and all-cause (ACM) or CVD mortality (CVDM). Models adjusted for sex, ethnicity, Townsend index, month of assessment, smoking, alcohol intake, physical activity level, body mass index, long-term condition count, and mutually for weekly group activity, living alone, and functional isolation.

5.5.6.2 Frequency of friends and family visits and living alone

Combined associations between frequency of friends and family visits and living alone (reference group of daily friends and family visits and not living alone) showed those who reported living alone had markedly stronger associations with each of the adverse health outcomes at every level of friend and family visit frequency (Figure 5-2 & Table S8, Appendix 8). For example, compared with daily friends and family visits and not living alone, all-cause mortality HRs for those who reported never having friends and family visits were 1.33 (1.22-1.46) in those not living alone and 1.77 (1.61-1.95) in those living alone. Tests for interaction provided some evidence for a multiplicative interaction between

friend and family visit frequency and living alone for all-cause (HR 1.11 [1.03-1.20]) but less so for CVD mortality (HR 1.07 [0.90-1.27]) (Table S9, Appendix 8). However, tests were suggestive of an additive interaction for CVD mortality (RERI 0.27 [-0.01, 0.57]; AP 0.13 [-0.02, 0.25]; SI 1.35 [0.99, 1.86]). This was consistent with the markedly higher HRs for CVD mortality in those never having friends and family visits who also lived alone (HR 2.23 [1.82-2.73]) compared with those never having friends and family visits but not living alone (HR 1.49 [1.21-1.84]) (Table S8, Appendix 8). In view of the evidence for interaction, stratified models were performed. Examining participants who reported not living alone and living alone separately, showed that the relative association with all-cause mortality of never having friends and family visits, compared to daily visits, was very similar in those living alone HR (1.40 [1.26-1.55]) as when the same comparison was made in those not living alone (HRs 1.36 [1.24-1.50]) (Table S10, Appendix 8). The same pattern was seen for CVD mortality (Table S10, Appendix 8). This is consistent with a stronger independent association with adverse health outcomes for never having friends and family visits compared with living alone (Table 5-5 & Figure 5-1).

5.5.6.3 Weekly group activity and living alone

Combined associations between weekly group activity and living alone (reference group of [yes] engaging in weekly group activity and not living alone) showed those who reported living alone had markedly stronger associations with each of the adverse health outcomes whether they engaged in weekly group activity or not (Table S11, Appendix 8). For example, compared with those who reported engaging in weekly group activity and not living alone, all-cause mortality HRs for those who reported no weekly group activity were 1.11 (1.08-1.14) in those not living alone and 1.46 (1.40-1.52) in those living alone. Tests for interaction provided some evidence for a multiplicative interaction between weekly group activity and living alone for all-cause (HR 1.07 [1.02-1.13]) but less so for CVD mortality (HR 1.05 [0.93-1.19]) (Table S12, Appendix 8). However, tests were more suggestive of an additive interaction for CVD mortality (RERI 0.12 [-0.06, 0.30]; AP 0.07 [-0.04, 0.17]; SI 1.23 [0.90, 1.67]). Models stratified by living alone showed that, compared with those who reported engaging weekly

group activity, the association with all-cause mortality for those not engaging in weekly group activity was higher among those living alone (HR 1.19 [1.14-1.25]) than when the same comparison was made among those not living alone (HRs 1.11 [1.08-1.14]) (Table S13, Appendix 8). The same pattern was seen for CVD mortality (Table S13, Appendix 8).

5.5.6.4 Frequency of friends and family visits, weekly group activity, and living alone combined

Based on the pattern of their independent associations with both adverse health outcomes, we combined all three structural component measures into an overall dichotomous structural isolation variable, with isolation coded as <monthly friends or family visits, or not engaging in weekly group activity, or living alone. Compared to those without, participants with structural isolation were associated with higher all-cause and CVD mortality: HR 1.23 (1.20-1.26) and 1.35 (1.28-1.43) (Table 5-5 & Figure 5-1).

5.5.7 Functional and structural component measures – combined associations and interactions

5.5.7.1 Frequency of ability confide, often feeling lonely, and structural isolation

Examining combined associations between the two functional component measures, structural isolation, and all-cause mortality showed that, when structural isolation was present, reporting never being able to confide was associated with similarly higher all-cause mortality regardless of often feeling lonely (HR 1.41 [1.34-1.49]) or not (HR 1.38 [1.32-1.44]) (Figure 5-3 & Table S14, Appendix 8). However, when structural isolation was absent, there was a greater difference in all-cause mortality associated with reporting never able to confide between those reporting often feeling lonely (HR 1.16 [1.07-1.26]) versus those reporting not often lonely (HR 1.07 [1.02-1.12]). A similar pattern was present for CVD mortality but with stronger associations and wider confidence intervals (Fig. S2 & Table S14, Appendix 8).

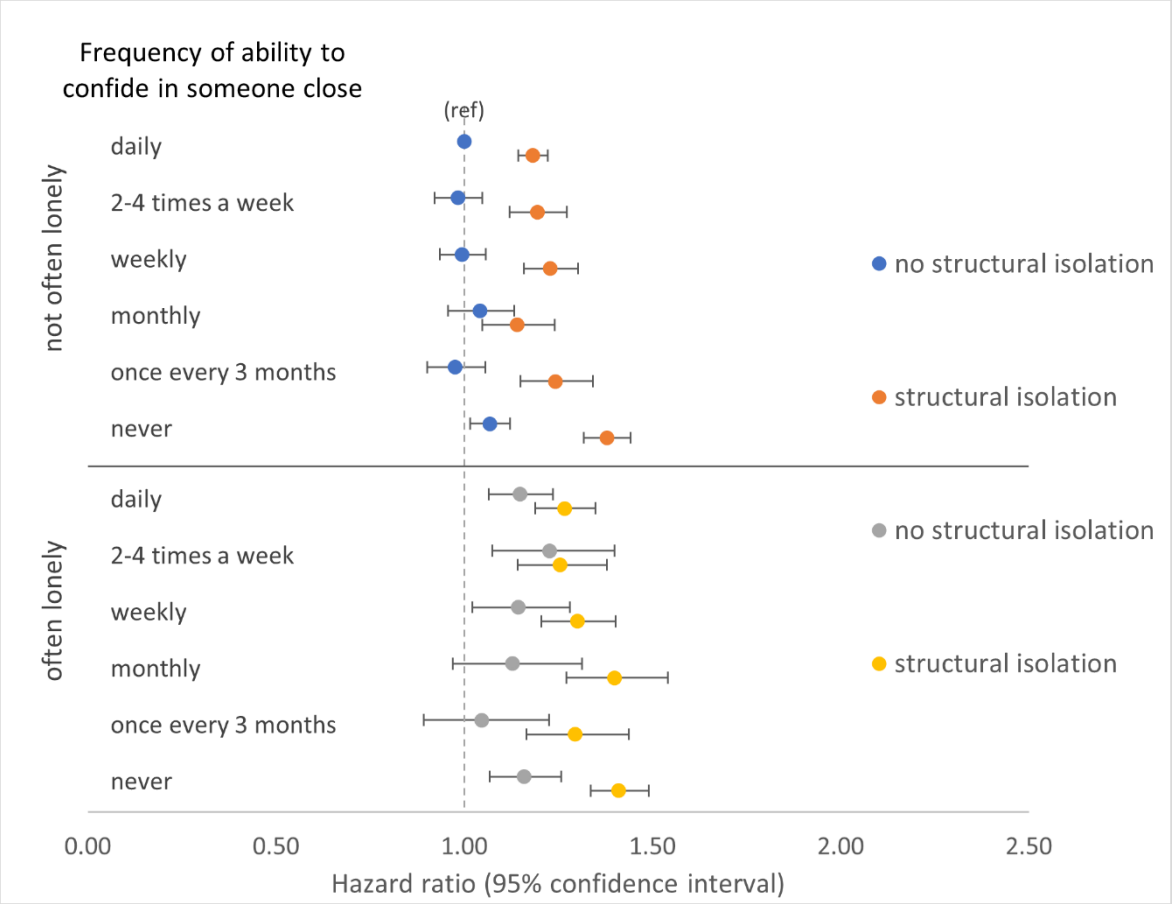


Figure 5-3: Models of combined associations between frequency of ability to confide in someone close, often feeling lonely, structural isolation, and all-cause mortality.

5.5.7.2 Frequency of friends and family visits, weekly group activity, living alone, and functional isolation

Joint associations between all three structural component measures and functional isolation showed that, compared to the reference group of those who reported daily friends and family visits, weekly group activity, not living alone, and without functional isolation, generally, there was a dose-response relationship where the addition of any of the three structural component measures or the addition of functional isolation was associated with higher all-cause mortality (Figure 5-4 & Table S15, Appendix 8). The highest all-cause mortality was observed in those who reported never having friends and family visits, not engaging weekly group activity, and living alone, but without functional isolation (HR 2.34 [1.65-3.30]). However, at this maximal level of structural isolation there were relatively few participants without functional

isolation (n=170) leading to wide confidence intervals in this group and complete overlap with the estimate for otherwise equivalent participants but who did report functional isolation (HR 1.99 [1.71-2.31]). Similarly, there were comparable estimates with wide and almost completely overlapping confidence intervals for those who reported never having friends or family visits and living alone, but who also reported engaging in weekly group activity, either with functional isolation (HR 1.98 [1.72-2.27]) or without functional isolation (HR 2.21 [1.68-2.90]). A similar pattern was present when CVD mortality was modelled as the outcome but with wider confidence intervals making interpretations more challenging (Fig. S3 & Table S15, Appendix 8). Overall, this is consistent with the larger independent effects of never having friends and family visits and living alone compared with weekly group activity or functional isolation (Figure 5-1).

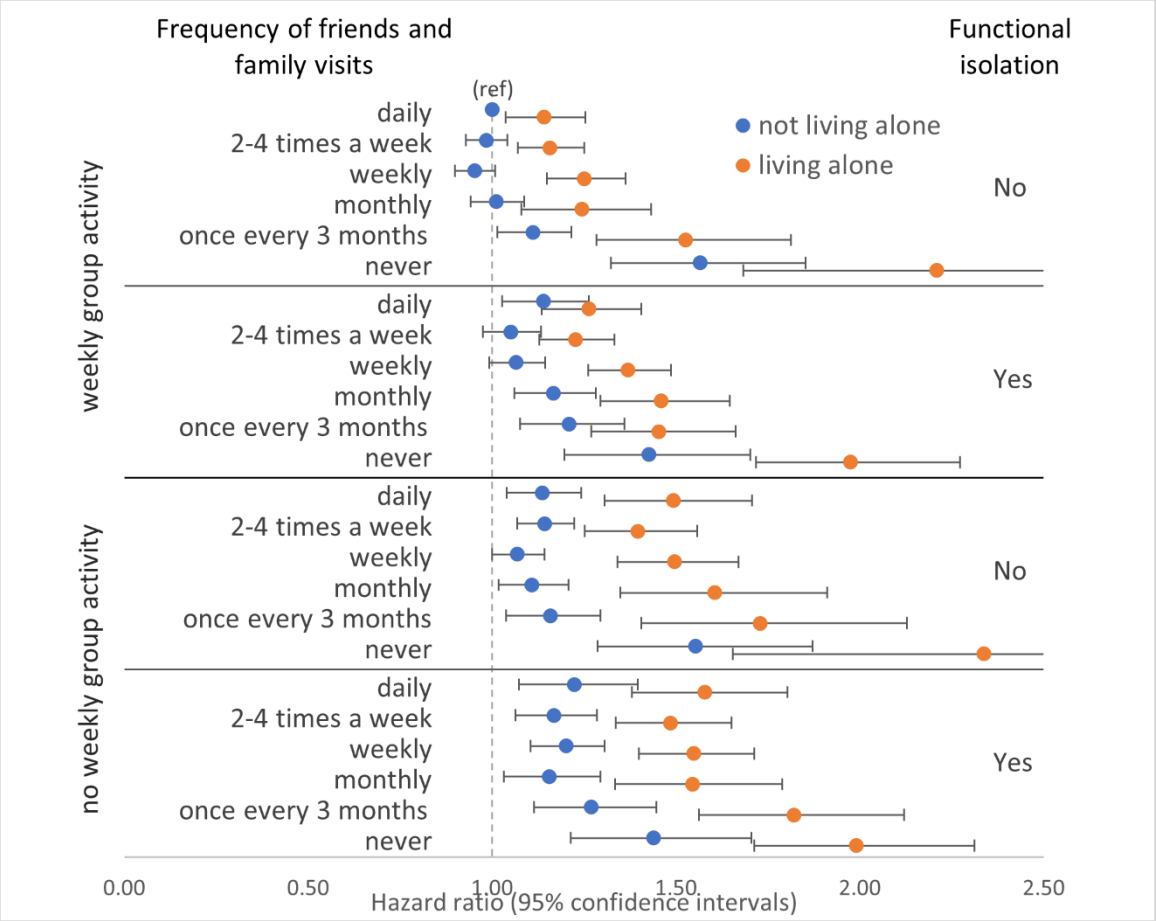


Figure 5-4: Models of combined associations between frequency of friends and family visits, weekly group activity, living alone, functional isolation, and all-cause mortality.

Further, in all categories of each of weekly group activity (yes/no), living alone (yes/no), and functional isolation (yes/no), there was incrementally lower all-cause mortality associated with increasing frequency in friends and family visits up to a level of monthly with further increases in frequency in friends and family visits being associated with similar levels of all-cause and CVD mortality (Figure 5-4 & Table S15, Appendix 8). This is consistent with the independent effect of frequency of friends and family visits (Table 5-5) where visit frequencies less than monthly were associated with adverse health outcomes. This suggests there may be a threshold effect for this type of social contact above or below which the health benefits may be felt or not.

In those not living alone and with no functional isolation, not engaging in weekly group activity was associated with higher all-cause mortality compared to engaging in weekly group activity at each level of friends and family visit frequency apart from those who reported never having friends and family visits where the mortality was similar (Figure 5-4). The same was true in those not living alone but with functional isolation and the pattern was more striking still in those reporting living alone.

5.5.7.3 Functional and structural isolation

Combined associations of functional and structural components overall showed, compared to those with neither functional nor structural isolation, there was higher all-cause mortality associated with structural isolation alone (HR 1.21 [1.17-1.24]) than with functional isolation alone (HR 1.11 [1.06-1.15) (Figure 5-5 & Table S16, Appendix 8). However, participants with both components of isolation were associated with the highest all-cause mortality (HR 1.36 [1.32-.1.40]). Consistent with this were results from tests for interaction which suggested an additive interaction: RERI 0.05 [-0.01, 0.10]; AP 0.03 [-0.01, 0.07]; SI 1.15 [0.97, 1.37] (Table S17, Appendix 8). The pattern was accentuated for CVD mortality and there was evidence of an additive interaction (Figure 5-5 & Tables S16 and S17, Appendix 8).

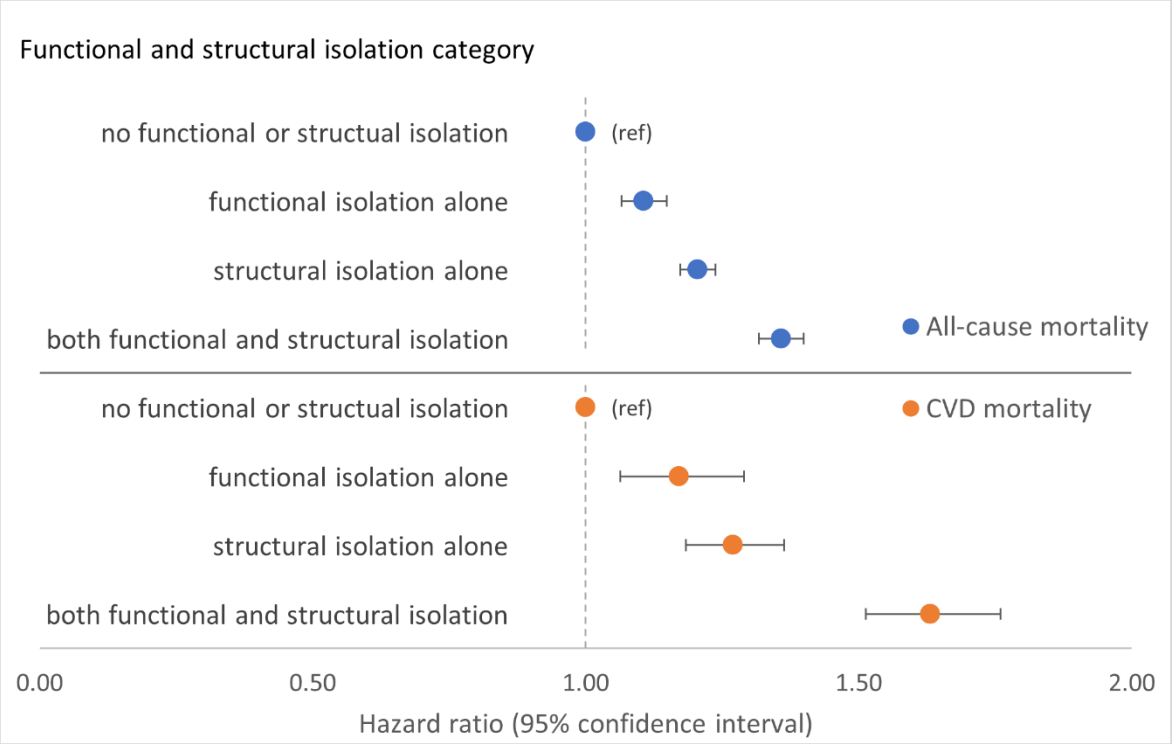


Figure 5-5: Models of combined associations between categories of functional and structural isolation and adverse health outcomes.

5.5.8 Sensitivity analyses

Similar results were seen across all sensitivity analyses where we excluded those with prior CVD or cancer or who died within 2 years of recruitment, albeit often with stronger associations and wider confidence intervals (Tables S18-S35, Appendix 9).

5.6 Discussion

5.6.1 Principal findings

This study shows that two measures of the functional component and three measures of the structural component of social connection were independently associated with all-cause and CVD mortality.

A combination of both measures of the functional component was also associated with adverse health outcomes. Previous studies using these measures to define 'loneliness' in UK Biobank may have underestimated this component of social connection as loneliness was only defined when both criteria were met.[221-223]

The association between never being able to confide in someone close and both adverse health outcomes appeared to be stronger when structural isolation was present irrespective of a subjective feeling of loneliness. When structural isolation was absent, the effect of never being able to confide appeared to be stronger among those who felt lonely. This highlights the complexity present in social connection but also why it may be important to consider different measures when exploring combined effects of functional and structural components of social connection on health outcomes.

Friends and family visit frequency of less than monthly was associated with higher all-cause and CVD mortality suggestive of a potential threshold effect, where visits up to a level of once a month could be required to maximise the benefits associated with these contacts. Further, never having friends and family visits was associated with highest all-cause and CVD mortality of any of the measures examined but associations were markedly lower for those reporting visits once every 3 months. If causal, this could suggest large health benefits may be associated with small changes in certain measures of social connection in those with a complete lack of that type of connection. Replicating this finding in other datasets and or identifying which measures of social connection would be most beneficial to target and the level of change which would maximise benefit could be extremely valuable.

The independent association between living alone and both all-cause and CVD mortality and the interactions with frequency of friends and family visits and with weekly group activity seen here suggests there may be high risks for living alone and even higher risks for living alone with additional levels of structural isolation (e.g., infrequent friends and family contacts or not engaging in regular group activity). Whilst it may be difficult or undesirable to change some individuals' living circumstances, our results suggest further investigation into whether identifying those who live alone (e.g., by front-line clinicians) could be warranted.[345,346]

When three structural component measures were examined in combination with functional isolation, the risks were similarly higher for all those with no friends or family contacts who also lived alone regardless of the presence of functional isolation or whether participants engaged in group activity. This result suggests there may be a hierarchy of components of social connection for those who experience numerous types of social disconnection. For example, our study showed the lower risk of mortality associated with regular group activity appeared to be masked by a lack of friends and family visits and living alone. Exploring this concept in other datasets could highlight targets for intervention for the most isolated in society.

Overall combined associations of functional and structural isolation showed that those defined as isolated by both components had the strongest associations with adverse health outcomes and there was evidence for an additive interaction for CVD mortality. Thus, further highlighting the potential importance of considering both components together.

5.6.2 Strengths and weaknesses

This study shows the added value of examining the adverse health outcomes associated with different individual measures of functional and structural components of social connection and their joint associations and interactions. A major strength of this study includes the large sample size of UK Biobank, which allowed us to examine the combination of different measures and components of social connection whilst adjusting for numerous potential confounders. The large

sample size also allowed us to conduct sensitivity analyses and show that our findings are less likely to be due to reverse causality.

There are some important limitations to our study. UK Biobank has a response rate of 5.5% and is not representative of the UK general population, which means there are risks of collider or selection bias.[247] However, while prevalence estimates may be inaccurate, strengths of association are likely to be more generalisable.[33] There remains a possibility of unmeasured confounding despite adjusting for numerous potential confounders. We have performed numerous analyses in this study which raises the issue of multiple testing. As a result, we only draw general conclusions where patterns of results (or differences in subgroups) are consistent across analyses of both outcomes and sensitivity analyses. The measures of social connection examined are self-reported which means our results could be affected by misclassification bias, leading to under or overestimates depending on the presence of random or systematic misclassification.[216] The measures used here are also relatively crude and binary variables fail to capture severity or dose response relationships. There are numerous alternative measures of functional and structural components of social connection not examined in this study which may be of equal or greater importance.[90] Indeed, relationship quality (e.g., marital strain) is another key component of social connection also associated with mortality.[347-349] UK Biobank, and therefore this analysis, lacks data that assesses relationship quality for the whole cohort. However, our study shows that separate measures of different components of social connection can interact, and further research could examine additional and more complex measures in similar detail.

5.6.3 Comparison with wider literature

Previous UK Biobank analyses examining both functional and structural components of social connection have used the same measures as in our analysis but have coded the item responses into scales of loneliness and social isolation, respectively.[221-223] For example, Elovainio et al. (2017) examined the association between loneliness and social isolation and mortality using the same measures as in our analysis to create social isolation and loneliness scores but they did not examine the association between each measure that comprised the score (frequency of friends and family visits) and for the ordinal variables (frequency

of ability to confide and frequency of friends and family visits) they did not estimate the level at which these measure were associated with outcomes. Further, their study did not examine for interactions between measures or between loneliness and social isolation. Our findings highlight the value of examining separate measures of functional and structural components of social connection.

Previous studies have examined the interaction between functional and structural components of social connection, but their results are mixed and are based on different multi-item scales or indices of each component making comparisons difficult. Some studies found no interaction between the two components,[222,334,350,351] while one found a positive interaction (where higher functional isolation strengthens the association between structural isolation and mortality and vice versa)[352] and another found a negative interaction (where higher functional isolation weakens the association between structural isolation and mortality and vice versa).[353] However, none of these studies examined additive interactions and none examined the associations or interaction between the separate measures that make up the multi-item scales or indices. Our study shows how examining the underlying associations of separate measures that make up each component may be warranted prior to defining isolation for each component. In our study there was evidence of an additive interaction between functional and structural components for CVD mortality and suggestive of the same for all-cause mortality. Overall, our findings highlight why considering both components together may be important, particularly when developing methods for identifying high-risk target populations for intervention.

Our findings differ from those of a meta-analysis of prospective studies examining the association between objective social isolation (e.g., infrequent social contacts), living alone, subjective loneliness and all-cause mortality.[317] In that study, average effects sizes were similar for social isolation, loneliness, and living alone (29%, 26%, and 32% increased likelihood of mortality, respectively). In contrast, we found greater effect sizes for those with the least frequent friends and family visits and for those who live alone compared with the effects sizes for not engaging in weekly group activity or those who felt lonely. The importance of having some friends and family visits highlighted here suggests

that these contacts could represent a more valuable type of social connection than others (e.g., social contact at a weekly group). For example, these contacts could reflect high quality social connections and, therefore, a lack of which would be strongly associated with adverse health outcomes. Additionally, these types of contacts may provide more practical support or be more likely to identify subtle deteriorations in the health and wellbeing of an individual. This is consistent with the smaller effect sizes for weekly group activity in our study, which featured in studies in the meta-analysis but often as part of multi-item measures of structural isolation where its individual impact was not assessed. The relatively lower effect size for functional isolation seen here compared with the equivalent results for loneliness from the meta-analysis, could be explained by a less stringent measure of functional isolation used here albeit with our measure being based on the associations between the individual constituent measures and adverse health outcomes.

Previous work has highlighted a lack of evidence for a threshold effect of measures of social connection, where risk becomes more pronounced at a certain level of isolation.[317] However, our study suggests that a threshold effect may exist as mortality associated with friends and family visits frequency was only higher at 'about once a month' and less often. Although this result may be due to the categories available (the response items in the original questionnaire) and there may indeed be a continuum of risk.

The association with mortality was observed to be stronger for structural than functional isolation. There are various reasons why this may be the case. For example, the measures of the functional component of social connection examined in this thesis were single item measures. Whereas the most common measure in the epidemiological literature for loneliness is the R-UCLA, which has 20 items.[219] If multicomponent measures are required to accurately measure subjective feelings of loneliness, then the strength of association between single item measures of loneliness and mortality outcomes may produce underestimates. This in turn may be due to the 'messier' and less tangible nature of the subjective functional components of social connection compared with more objective structural components. For example, there could be greater variation in the interpretation of 'do you often feel lonely' compared with 'how

often do friends and family visit you', which could bias associations between single item measures of the functional component of social connection and mortality towards the null more than equivalent associations for single item measures of the structural component of social connection.

Another potential reason for stronger mortality associations for structural components is that structural components may be more strongly associated with factors that are 'higher upstream' and are more closely linked to wider socioeconomic factors that also influence the risk of adverse health outcomes. For example, having frequent friends and family visits may be confounded by socioeconomic factors (e.g., having frequent visits from or frequently visiting friends and family may require time and resources) more than frequent feelings of loneliness which may be more influenced by individual-level 'downstream' factors. Therefore, estimates of the observed association between structural components and mortality in this analysis may be overestimates.

5.6.4 Future research

There is no standard measure for social connection. However, the independent risks of living alone, and the interactions with both friends and family visits and weekly group activity, seen here suggests that further work is warranted in ascertaining whether living alone could represent a single and simple measure that could be standardised and included in studies examining social connection.[354] Our findings suggest that the benefits of group activity could be masked by an overriding negative effect of never having friends or family contacts. Further examination into the ways in which components of social connection interact could inform how intervention targets might be prioritised, particularly for those who are most isolated. Finally, more work is required to understand the role of potential mediators (e.g., mental health problems or health behaviours) to further elucidate the mechanistic pathways by which social disconnection might cause adverse health outcomes and inform future interventions.

5.7 Conclusions

This study of UK Biobank is the first to examine two measures of the functional component and three measures of the structural component of social connection both independently and in combination. Our findings suggest that advice, interventions, and policy may need to be tailored to address different aspects of social connection and target highest risk groups. Specifically, we show that separate measures of different components of social connection may contribute different levels of risk of adverse health outcomes. And the combined associations and interactions of the measures examined here suggest that those who live alone with additional concurrent markers of structural isolation may represent a population who could benefit from targeted support.

6 Sociodemographic effects on the association between a weighted lifestyle score and mortality in the UK Biobank cohort.

6.1 Chapter summary

This chapter provides analyses for the development of a weighted lifestyle score using 11 lifestyle factors and to examine the effects of deprivation, sex, ethnicity, and age on the association between the weighted score and all-cause and CVD mortality. This chapter addresses research objectives 3) a-c), which were:

a) to examine the strength of associations between a wide range of individual lifestyle factors (including measures of social connection) and adverse health outcomes

b) use the magnitude of effect sizes for numerous lifestyle factors to create a weighted lifestyle score

c) examine the effect of socioeconomic deprivation on the association between the weighted lifestyle score and adverse health outcomes.

The text and figures are presented as submitted to BMC Public Health on 29th March 2024 (under peer review at the time of writing). Associated Supplementary material 1 and 2 are given here in Appendix 10 and Appendix 11, respectively. Before the manuscript is presented, there follows a brief description explaining how analyses in Chapter 5 informed the work in this chapter.

The social connection analyses showed that each of the available measures of social connection in UK Biobank could be included. However, with the focus on producing a lifestyle score that could be used to support individuals enact healthy lifestyle change it was decided to focus on those measures of social connection that could be more amenable to individual-level action. For example, it is less clear how living alone or having reduced frequency of ability to confide in someone close could be considered lifestyle factors that are

amenable to healthy change via individual-level action. Whereas it is easier to envisage how increasing the frequency of friends and family visits or weekly group activity could be more readily modified by individuals either alone or via community-level interventions. Further, although important for health and inter-related with lifestyle factors, living alone, having reduced frequency of ability to confide in someone close, and feeling lonely do not fit the definition of lifestyle factors used in this thesis. Therefore, frequency of friends and family visits and engaging in weekly group activity, two structural components of social connection, were included in this Chapter and the set of quantitative analyses that address RQ3. Also based on the preceding social connection analysis, friends and family visits less frequent than monthly was classified as unhealthy. This permitted these two measures of social connection to be combined into a dichotomous social participation lifestyle factor variable (as per the other lifestyle factors which were also dichotomised) where a low ('unhealthy') level of social participation was defined as either having friends and family visits less often than monthly or not engaging in weekly group activity.

6.2 Abstract

Background: Unhealthy lifestyles are associated with disproportionate mortality among deprived populations. We aimed to create a weighted lifestyle score and examine the effects of sociodemographic variables on the association between the score and mortality.

Methods: Prospective analysis of 462,235 UK Biobank participants aged 37-73 years. A weighted lifestyle score was developed using 11 lifestyle factors (LFs): smoking/alcohol/physical-activity/TV-time/intake of red-meat/processed-meat/salt/oily-fish/fruit and vegetables/sleep/social participation. Cox models of associations between individual LFs and all-cause mortality determined score weightings. Weightings were combined into a score to explore effects of deprivation/sex/ethnicity/age on the association between weighted score and all-cause/CVD mortality.

Results: Over 12.0 years median follow up, 30,687 (6.6%) participants died including 4,632 (1.0%) CVD deaths. Each LF was independently associated with all-cause/CVD mortality. Weighted score (maximum 30 points indicating unhealthy) comprised 14 points for smoking, one each for unhealthy oily fish/red meat/processed meat/salt intake, and two each for remaining LFs. There was a dose-response association between all-cause/CVD mortality and increasing weighted score, with stronger associations both among more deprived and male participants. Compared with least deprived in the lowest score category, all-cause mortality HRs (95%CIs) for those with highest scores were 2.67 (2.43, 2.92) and 4.71 (4.43, 5.01) among the least and most deprived, respectively. Equivalent HRs but with women in the lowest score category as reference, were 3.07 (2.88, 3.26) for women and 4.66 (4.44, 4.89) for men.

Conclusions: An extended weighted lifestyle score has strong associations with mortality, particularly among deprived and male participants, and could convey personalised risk and inform policy.

Funding: HF and his supervisors (CO'D, FM, JG, DL) were supported by Medical Research Council Clinical Research Training Fellowship (grant number MR/T001585/1), and the funder played no role at any stage in this study

including the decision to publish. Remaining co-authors received no specific funding for this work.

Keywords: Epidemiology, Lifestyle, Health behaviours, Health inequalities, Social medicine

6.3 Introduction

Lifestyle factors (LFs) such as smoking or alcohol intake are among the leading risk factors for death.[1] LFs have greater impact on socioeconomically deprived populations, both via clustering and having disproportionately strong associations with mortality.[15,355] LFs combined into scores can guide and support individuals make healthy behaviour change and delay adverse health.[72,356] However, various barriers prevent widespread use and reduce the effectiveness of lifestyle scores among deprived populations where need for lifestyle support is greatest.[357] Scores often require anthropometric (e.g., weight) or physiological (e.g., blood pressure) measurements, requiring time and resources. Secondly, scores generally assign similar weightings to different LFs despite varying magnitudes of effect on risk. For example, in the American Hearts Association's (AHA) Life's Essential 8, smoking has the same weighting as physical inactivity.[356] Finally, scores fail to account for additional risk associated with deprivation, itself a risk factor inter-related with LFs.[24] Developing a simple lifestyle score that omits physiological measurements, is weighted for the relative effects of different LFs, and accounts for deprivation could provide a new tool for prioritising LFs for intervention.

We previously created a lifestyle score (based on smoking, physical activity, short/long sleep duration, television viewing time, and intake of alcohol, fruit and vegetables, oily fish, red or processed meats) and showed that associations between an unhealthy score and mortality were stronger in more deprived groups.[24] However, this score assumed the same level of risk associated with each LF. Further, there may be demographic variations in risk associated with lifestyle scores as there are sex and ethnic differences in metabolic factor-associated mortality and also higher mortality in older ages with some LFs.[358-360] However, less is known about demographic variation in the mortality associated with combinations of 'behavioural' LFs. Identification of high risk groups could influence policy and interventions designed to address combinations of LFs.[302]

This study aims to: 1) create a simple weighted lifestyle score reflecting the relative associations between a wide range of LFs and mortality, 2) examine whether and how deprivation affects the association between the weighted

score and mortality, and 3) examine whether sex, ethnicity, and age affect the association between the weighted score and mortality.

6.4 Methods

6.4.1 Study design, data, and participants

We prospectively analysed data from UK Biobank, which recruited 502,459 UK adults between 2006–2010.[246] Participants completed questionnaires and nurse-led interviews at recruitment (see https://biobank.ndph.ox.ac.uk/ukb/).[213]

6.4.2 Outcomes

The outcomes were all-cause and cardiovascular (CVD) mortality. Details of deaths were provided by linkage to National Health Service (NHS) Information Centre (England and Wales) and the NHS Central Register (Scotland). Analyses were censored at the latest date available (Feb 28, 2021) or date of death, if earlier. Primary causes of death with International Classification of Diseases (10th Revision) codes I05–I99, Z86.7, G45, and G46 were classed as CVD deaths.[361]

6.4.3 Exposures

Participants self-reported a range of LFs. Based on previous work, 11 LFs were included: smoking, alcohol, physical activity, television (TV) time, sleep duration, social participation level, and intake of red meat, processed meat, salt added to food, oily-fish, and fruit and vegetables (Table S1, Appendix 10).[24] Salt added to food and social participation were included to extend the range of LFs in a future clinical risk score.[362,363] Social participation components considered modifiable on an individual-level (friends/family visits and group activity) were included.[361] Each LF was dichotomised and coded as unhealthy/healthy as determined by available guidelines or latest evidence (Table S1, Appendix 10).

6.4.4 Covariate data

Socioeconomic deprivation was measured by Townsend index, calculated by UK Biobank using 2001 census data and residential postcode (higher scores indicate higher relative deprivation).[341] Townsend index was analysed as a continuous

variable unless otherwise stated. Other potential confounding variables included were sex (female, male), ethnicity (White, mixed, Asian/Asian British, Black/Black British, Chinese, Other), 'cholesterol lowering' or 'blood pressure' medication ('CVD medication'; yes, no), and long-term condition (LTC) count based on a list of 43 conditions.[339] Self-reported LFs may vary by season and so date of assessment was included as winter/summer (Oct–Mar/Apr–Sep) months. Systolic blood pressure (mmHg) and body mass index (BMI, kg/m^2) were analysed as continuous variables.

6.4.5 Statistical analysis

We excluded participants without full data. For main analyses, we used Cox proportional hazard models to examine associations between exposures and outcomes. We used age as the timescale because risk of death increases with age. Follow-up time was calculated as the difference in age (in days) between assessment and the earlier date of censoring or death.

To create a weighted lifestyle score, we examined associations between each LF and all-cause mortality, adjusting for sex, ethnicity, deprivation, month of assessment, blood pressure, BMI, CVD medication, LTC count, and mutually for remaining LFs. The strength of these associations, based on hazard ratios (HRs), determined the weighting for each LF. In view of future clinical utility, we aimed to select a scale of weighting that resulted in a relatively simple score. Therefore, weightings were rounded to the nearest whole integer, so each LF associated with all-cause mortality contributed a whole integer value. We trialled various formulae (e.g., [HR-1]*10, [HR-1]*11…etc.) until the sum of all weightings provided a small total rounded to the nearest 10. The final formula for the weighting scale was (HR-1)*13, which resulted in a score range 0–30. Each participants' weighted score was calculated as the sum of their weighted LFs. We then examined the association between the weighted score and both outcomes, modelling the score as a continuous variable. To examine sociodemographic effects on the association between weighted score and outcomes and maintain power, we categorised the weighted score into very low (<5), low (5-9), moderate (10-14), high (15-19), and very high (>19) score categories. We then modelled the associations between score category and outcomes.

To explore potential deprivation effects, we examined interactions on both multiplicative and additive scales by estimating the relative excess risk for interaction (RERI), attributable portion (AP), and synergy index (SI) with estimates expressed as HRs and 95% confidence intervals (95% CI).[344] Estimating RERI, AP, and SI requires two dichotomous exposures. Therefore, the weighted score was dichotomised using score categories low and very low (≤9)/moderate, high, and very high (≥10). A representative Towsend index median (UK 2001 census) was used to dichotomise deprivation (≤/> -0.83).[364] Models were adjusted as before but with deprivation removed. To explore effect modification, we then examined the combined association between weighted score category, deprivation quartile (as per UK 2001 census), and outcomes.[364] To further explore deprivation effects, we examined the association between weighted score category and outcomes stratified by deprivation quartile.[344]

To explore the effect of sex, models were adjusted for ethnicity, deprivation, month of assessment, blood pressure, BMI, CVD medication, and LTC count and used to examine: 1) interactions between sex and weighted score (≤9/≥10), 2) combined associations between weighted score category, sex, and outcomes, and 3) associations between weighted score category and outcomes stratified by sex.

For ethnicity effects, models were adjusted for sex, deprivation, month of assessment, blood pressure, BMI, CVD medication, and LTC count and used to examine: 1) combined associations between weighted score category, ethnicity, and outcomes and 2) associations between weighted score category and outcomes stratified by ethnicity. Interaction estimates were not calculated because the requisite dichotomous variable would involve combining all non-White groups, which is clearly too heterogenous.

We were unable to examine the effect of age in the same way because including age in models where age is the timescale leads to overadjustment.[365] Therefore, we used a 'time-on-study' timescale and adjusted for sex, ethnicity, deprivation, month of assessment, blood pressure, BMI, CVD medication, and LTC count to examine: 1) interactions between age (≤55/>55 years) and weighted score (≤9/≥10), 2) combined associations between weighted score category, age category (≤45, >45 to ≤55, >55 to ≤65, >65 years), and outcomes,

and 3) associations between weighted score category and outcomes stratified by age.

6.4.6 Sensitivity

To examine the effect of using CVD mortality to create the weighted score, we repeated CVD mortality analyses after creating an alternative weighted score by using associations between individual LFs and CVD-mortality. To reduce the chance of reverse causality being the explanation for findings (i.e., poor health leading to unhealthy LFs and mortality), we repeated main analyses after excluding those with self-reported CVD (diabetes, coronary heart disease, atrial fibrillation, chronic heart failure, chronic kidney disease, hypertension, stroke/transient ischaemic attack, or peripheral vascular disease) or cancer and those who died within two years of recruitment.

6.5 Results

6.5.1 Descriptive statistics

Of 502,459 participants, 40,224 (8.0%) had data missing for at least one covariate and were more likely to be male, of minority ethnic backgrounds, more deprived, assessed during summer months, have higher BMI, take CVD medication, have more LTCs, and report smoking and unhealthy levels of physical activity, TV time, sleep, red meat intake, and added salt (Figure S1 & Table S2, Appendix 10).

Of 462,235 participants with complete data, 179,307 (38.8%), 125,150 (27.1%), 90,135 (19.5%), 67,643 (14.6%) participants were in deprivation quartiles 1 (least deprived) to 4 (most deprived), respectively (Table 6-1). Similar percentages of participants across deprivation quartiles were women and had similar month of assessment and mean blood pressure. However, those from more deprived quartiles were more likely to be younger, from minority ethnic groups, have higher BMI, take CVD medication, have more LTCs, and report unhealthy LFs.

After a median 12.0 years (11.3–12.0 inter-quartile range) follow-up there were 30,687 (6.6%) deaths, of which 4,632 (1.0%) were CVD deaths.

Table 6-1: Characteristics of participants by deprivation quartile

	Deprivation quintile				Total
	1 (least deprived)	2	3	4 (most deprived)	
N (row%)	179,307 (38.8%)	125,150 (27.1%)	90,135 (19.5%)	67,643 (14.6%)	462,235 (100)
Female	97,419 (54.3%)	69,205 (55.3%)	49,974 (55.4%)	36,078 (53.3%)	252,676 (54.7%)
Mean age (SD)	57.1 (7.86)	56.7 (8.04)	55.9 (8.23)	55.2 (8.35)	56.5 (8.08)
Ethnicity					
White	175,769 (98.0%)	121,190 (96.8%)	84,135 (93.3%)	59,509 (88.0%)	440,603 (95.3%)
Mixed	558 (0.3%)	575 (0.5%)	652 (0.7%)	854 (1.3%)	2639 (0.6%)
Asian/Asian British	1497 (0.8%)	1669 (1.3%)	2497 (2.8%)	2051 (3.0%)	7714 (1.7%)
Black/Black British	460 (0.3%)	825 (0.7%)	1584 (1.8%)	3465 (5.1%)	6334 (1.4%)
Chinese	399 (0.2%)	318 (0.3%)	348 (0.4%)	270 (0.4%)	1335 (0.3%)
Other	624 (0.3%)	573 (0.5%)	919 (1.0%)	1494 (2.2%)	3610 (0.8%)
Townsend index					
Mean (SD)	-4.09 (0.793)	-1.99 (0.588)	0.512 (0.851)	4.27 (1.60)	-1.40 (3.03)
Median [Min, Max]	-4.00 [-6.26, -2.90]	-2.05 [-2.90, -0.830]	0.451 [-0.830, 2.13]	3.92 [2.13, 11.0]	-2.22 [-6.26, 11.0]
Winter Assessment	86,331 (48.1%)	58,284 (46.6%)	43,585 (48.4%)	32,399 (47.9%)	220,599 (47.7%)
Systolic BP at assessment, Mean (SD)	135 (18.2)	134 (18.3)	133 (18.2)	133 (18.5)	134 (18.3)
Body mass index (kg/m^2), Mean (SD)	27.0 (4.37)	27.3 (4.62)	27.6 (4.99)	28.2 (5.46)	27.4 (4.75)
Self-reported CVD medication	46,196 (25.8%)	33,471 (26.7%)	24,704 (27.4%)	21,039 (31.1%)	125,410 (27.1%)
Long-term condition count, Mean (SD)	1.11 (1.15)	1.16 (1.20)	1.23 (1.26)	1.40 (1.39)	1.19 (1.23)
Current smoker	11,584 (6.5%)	10,859 (8.7%)	11,218 (12.4%)	13,436 (19.9%)	47,097 (10.2%)
High alcohol intake	14,995 (8.4%)	11,014 (8.8%)	8799 (9.8%)	6796 (10.0%)	41,604 (9.0%)
Low PA level	33,506 (18.7%)	23,643 (18.9%)	17,389 (19.3%)	14,321 (21.2%)	88,859 (19.2%)
TV time ≥4 hours/day	46,178 (25.8%)	35,102 (28.0%)	26,423 (29.3%)	24,428 (36.1%)	132,131 (28.6%)
Sleep <7 or ≥9 hours/day	51,547 (28.7%)	38,497 (30.8%)	30,305 (33.6%)	26,685 (39.4%)	147,034 (31.8%)

Fruit and vegetable intake <400 g/day	31,300 (17.5%)	23,296 (18.6%)	18,670 (20.7%)	16,590 (24.5%)	89,856 (19.4%)
Oily fish intake <1 portion/week	74,811 (41.7%)	54,584 (43.6%)	41,358 (45.9%)	31,928 (47.2%)	202,681 (43.8%)
Red meat intake >3 portions/week	26,060 (14.5%)	17,984 (14.4%)	13,184 (14.6%)	11,512 (17.0%)	68,740 (14.9%)
Processed meat intake >1 portion/week	54,004 (30.1%)	38,242 (30.6%)	27,942 (31.0%)	22,518 (33.3%)	142,706 (30.9%)
Salt added to food usually or always	25,637 (14.3%)	19,391 (15.5%)	15,513 (17.2%)	14,080 (20.8%)	74,621 (16.1%)
Low social participation	57,758 (32.2%)	42,296 (33.8%)	33,259 (36.9%)	27,936 (41.3%)	161,249 (34.9%)
Weighted lifestyle score, Mean (SD)	5.54 (4.48)	6.03 (5.00)	6.82 (5.72)	8.42 (6.79)	6.34 (5.35)
Weighted lifestyle score category					
<5	89,617 (50.0%)	58,298 (46.6%)	38,017 (42.2%)	22,515 (33.3%)	208,447 (45.1%)
5-9	71,980 (40.1%)	51,146 (40.9%)	36,639 (40.6%)	27,569 (40.8%)	187,334 (40.5%)
10-14	6087 (3.4%)	4818 (3.8%)	4238 (4.7%)	4088 (6.0%)	19,231 (4.2%)
15-19	5854 (3.3%)	4988 (4.0%)	4685 (5.2%)	4630 (6.8%)	20,157 (4.4%)
>19	5769 (3.2%)	5900 (4.7%)	6556 (7.3%)	8841 (13.1%)	27,066 (5.9%)
Weighted lifestyle score split at median (4)					
More healthy	113,419 (63.3%)	74,627 (59.6%)	48,962 (54.3%)	30,155 (44.6%)	267,163 (57.8%)
Less healthy	65,888 (36.7%)	50,523 (40.4%)	41,173 (45.7%)	37,488 (55.4%)	195,072 (42.2%)
Townsend index split at UK census median					
Least deprived	179307 (100%)	125,150 (100%)	0 (0%)	0 (0%)	304,457 (65.9%)
Most deprived	0 (0%)	0 (0%)	90,135 (100%)	67,643 (100%)	157,778 (34.1%)

6.5.2 Individual lifestyle factors and creation of weighted lifestyle score

Each individual LF was independently associated with all-cause mortality. Figure 6-1 shows the associations for each LF and their score weighting. Compared with

non-smoking, all-cause mortality HR (95% CIs) for smoking was 2.09 (2.03, 2.15), resulting in the highest score weighting (14 points). Equivalent results for low (versus high) oily-fish intake were 1.02 (1.00, 1.05), which, along with high red and processed meat intake and added salt, had the lowest weighting (1 point each). Associations between high alcohol intake, low physical activity, low social participation, high TV time, short/long sleep, and low fruit and vegetable intake and all-cause mortality were similar (HRs 1.12–1.19), resulting in weightings of 2 points each.

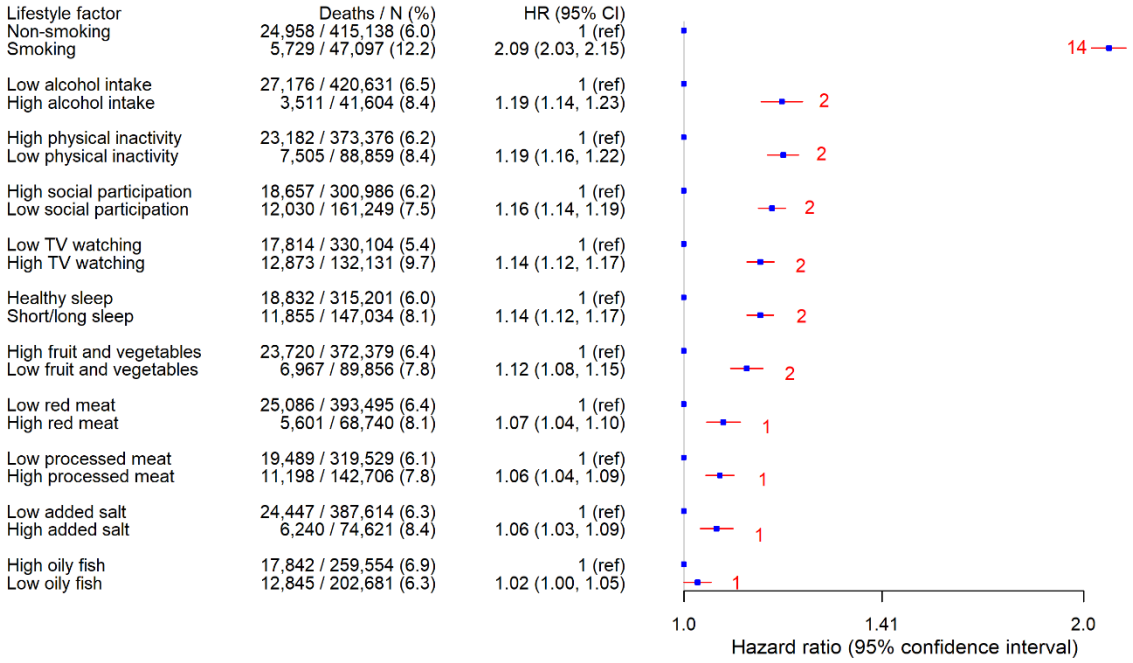


Figure 6-1: Models of independent associations between lifestyle factors and all-cause mortality. Weighting assigned to each lifestyle factor shown in red.

Heavily influenced by smoking, the score had a bimodal distribution (Figure S2, Appendix 10). Most participants (415,138 [89.8%]) were non-smokers and therefore had very low/low scores (0-9). Fewer participants had moderate (10-14) compared with high scores 15-19. The median weighted score was four and no participant had the maximum score 30. The median number of LFs was two for non-smokers and four for smokers (data not shown).

6.5.3 Association between weighted lifestyle score and outcomes

Models of the association between the weighted score and outcomes showed an approximately linear dose-response relationship (Figure 6-2). Compared with those with a score of zero, HRs (95%CI) for those with a score of 29 were 12.38 (7.31, 20.96) and 11.10 (2.75, 44.75) for all-cause and CVD mortality, respectively. Fewer participants and events at scores >25 resulted in wider confidence intervals for these scores. Confidence intervals were also wider between scores 12-15 as there were both few non-smokers that reported high number of other LFs (maximum score for non-smokers was 16) and few smokers without other LFs. The gradient of the association between weighted score and all-cause mortality appeared slightly steeper at scores >16 (i.e., among smokers) implying that additional LFs were associated with higher all-cause mortality among smokers than non-smokers. However, that pattern was less clear for CVD mortality. A clear dose-response relationship was seen modelling the association between weighted score categories and outcomes (Figure 6-3).

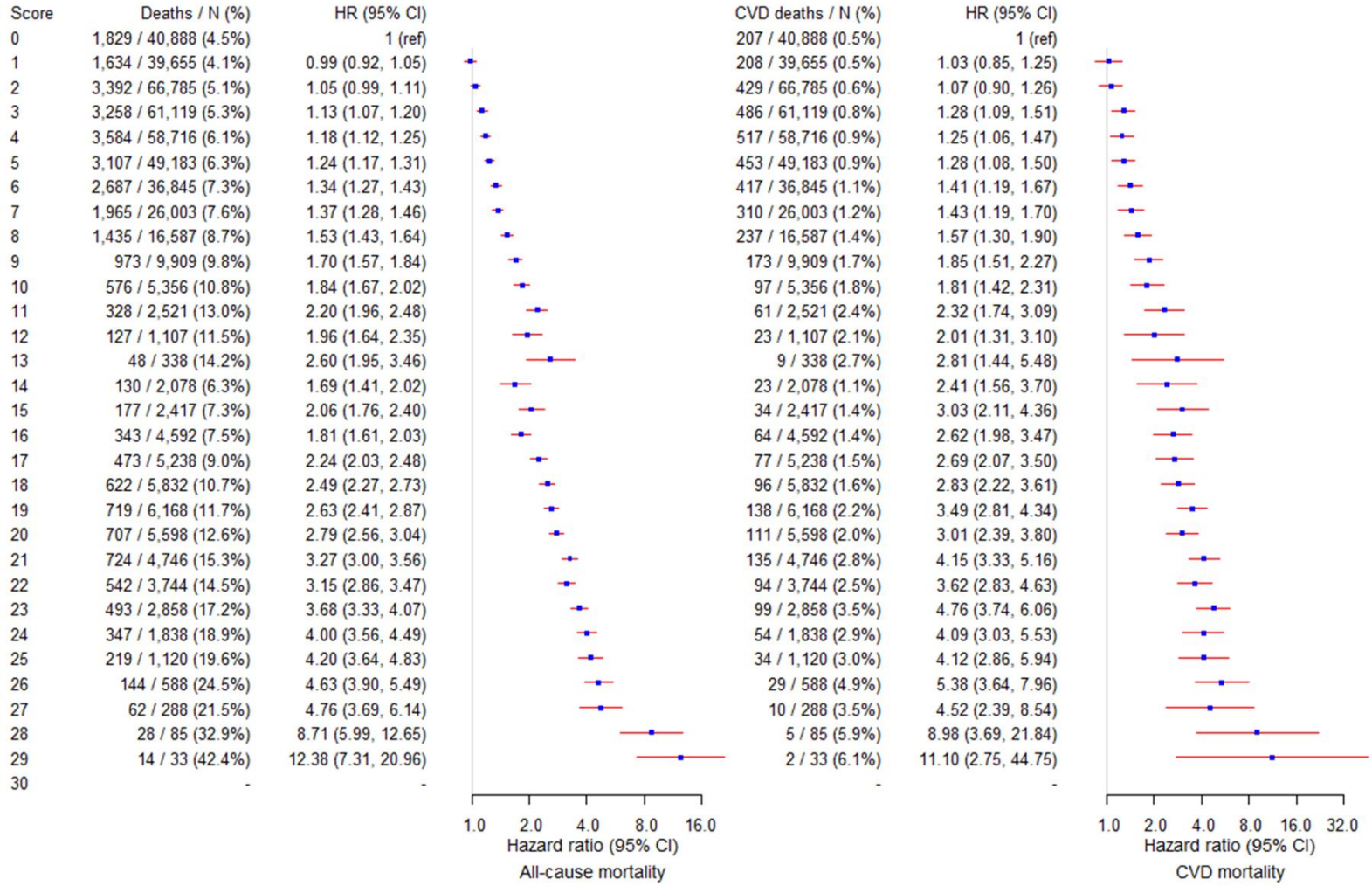


Figure 6-2: Association between weighted lifestyle score and all-cause and CVD mortality.

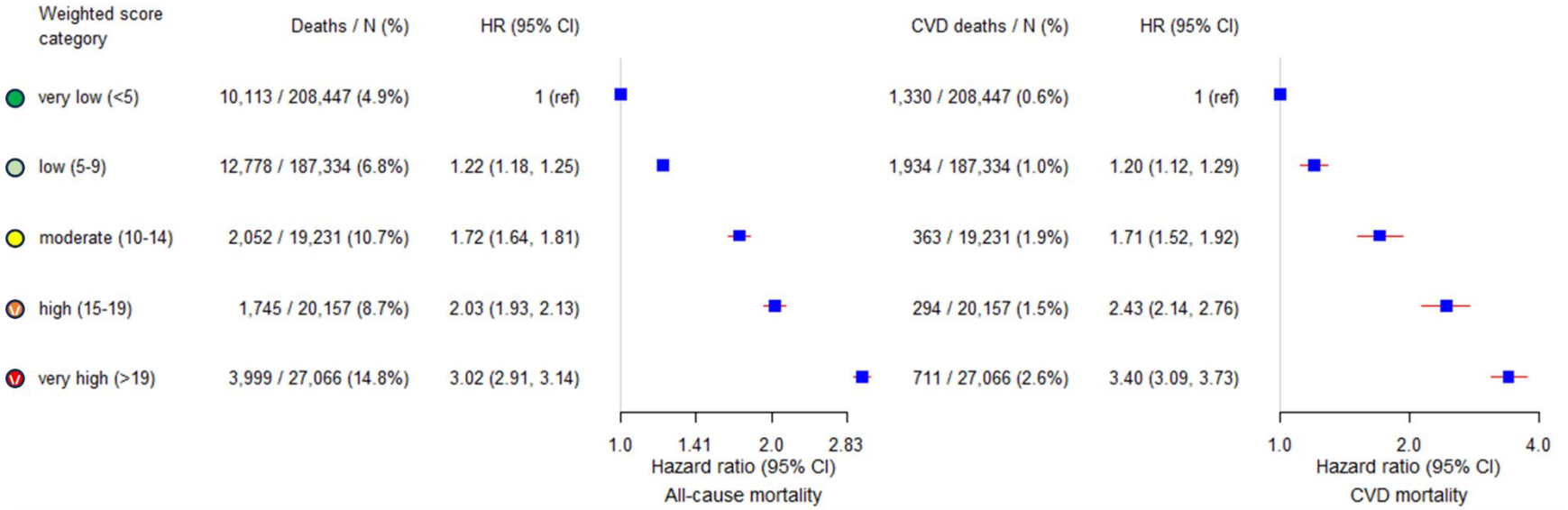


Figure 6-3: Association between weighted lifestyle score categories and all-cause and CVD mortality.

6.5.4 The effect of socioeconomic deprivation

There was evidence of a significant interaction on both additive and multiplicative scales between the weighted score (≤9/≥10) and deprivation (≤/> median) for both all-cause and CVD mortality (Table S3, Appendix 10). Interaction estimates for all-cause mortality were: multiplicative 1.22 (1.16,1.29), RERI 0.73 (0.62, 0.84), AP 0.26 (0.22, 0.29), and SI 1.65 (1.52, 1.78). CVD mortality interaction estimates were similar. These results indicate that a moderate, high, or very high score had greater effect in the more deprived group. Consistent with this, combined associations between weighted score category, deprivation quartile, and outcomes showed a higher HR for the same score category with increasing deprivation (Figure 6-4). For example, compared with those in the least deprived quartile with very low scores, all-cause mortality HRs (95% CI) for those with very high scores were 2.67 (2.43, 2.92) and 4.71 (4.43, 5.01) among the least and most deprived quartiles, respectively. Equivalent results for CVD mortality revealed a more accentuated pattern, albeit with wider and overlapping confidence intervals. With participants stratified by deprivation, compared with those with very low scores, stronger mortality associations were seen among those with very high scores with increasing level of deprivation (Figure S3, Appendix 10). Within each quartile of deprivation, compared with those with very low scores, all-cause mortality HRs (95% CI) for their counterparts with very high scores were 2.55 (2.35, 2.77), 2.92 (2.70, 3.16), 3.27 (3.02, 3.54), and 3.54 (3.27, 3.82) in the least to most deprived quartiles, respectively.

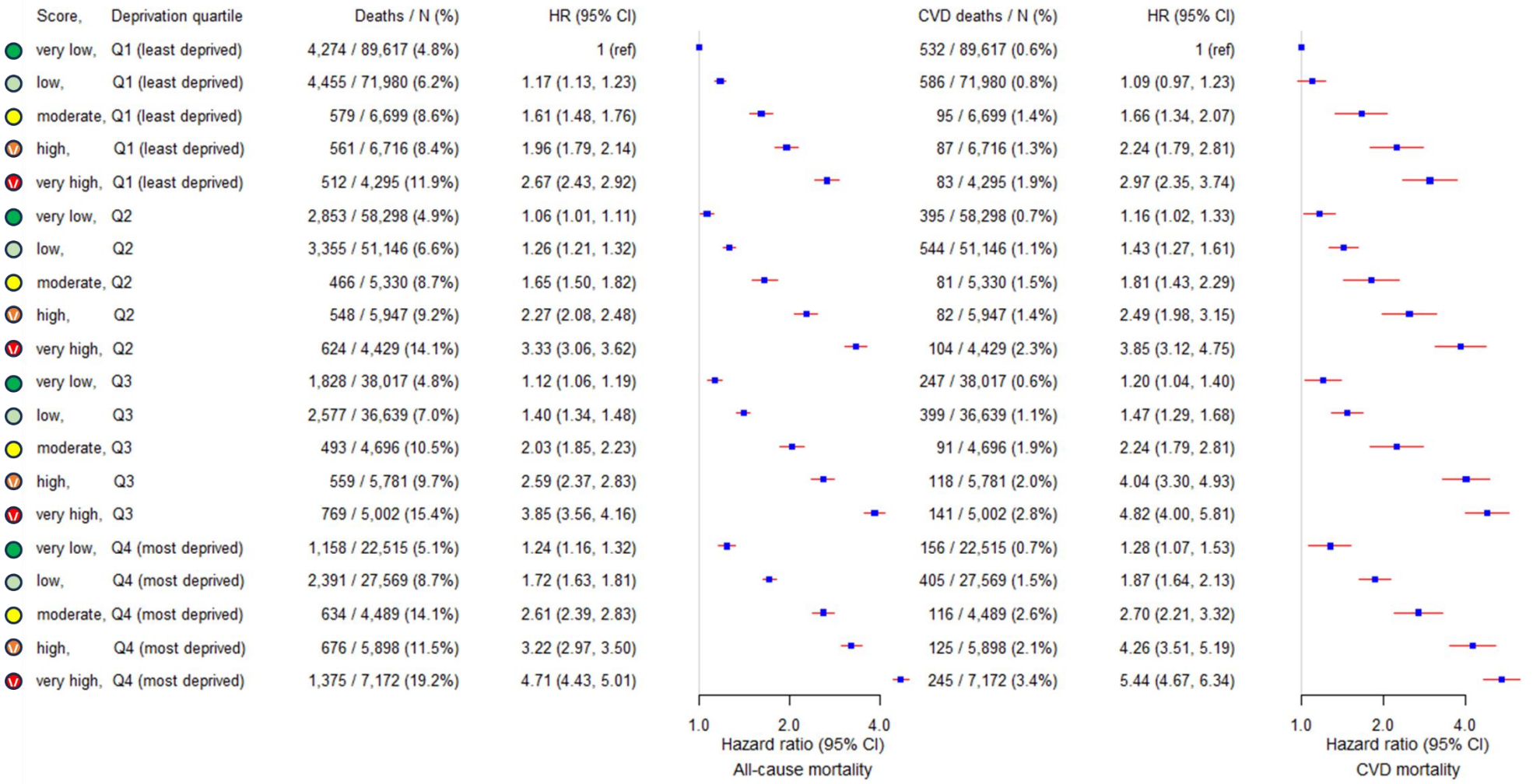


Figure 6-4: Combined association between weighted lifestyle score, deprivation, and all-cause and CVD mortality.

6.5.5 The effect of sex

Evidence for interaction between sex and the weighted score for all-cause and CVD mortality was observed on the additive scale only and indicated the relative effect of a weighted score ≥10 was similar in men and women (Table S4, Appendix 10). However, models of combined associations showed men had markedly higher hazards than women with the same score category (Figure S4, Appendix 10). The pattern was most pronounced for CVD mortality where results highlighted the effect of male sex alone. For example, compared with women with very low scores, men with very low scores had similar CVD mortality HRs (95%CI) to women with very high scores: 3.63 (3.22, 4.10) and 3.77 (3.09, 4.60), respectively. While the highest absolute rate of deaths was seen in men, when associations were stratified by sex, the effect of increasing weighted score category appeared slightly greater in women (Figure S5, Appendix 10). However, the pattern was similar for women and men and there were fewer deaths and wider confidence intervals of estimates among women.

6.5.6 The effect of ethnicity

Models of combined associations showed that participants with similar scores had broadly similar hazards across ethnicities with a dose-response relationship with increasing score category (Figure S6, Appendix 10). However, there were relatively few participants from minority ethnic groups compared with White participants resulting in wide confidence intervals, which made results difficult to interpret in some groups. Stratified by ethnicity, the weighted score showed similar effects across ethnic groups (Figure S7, Appendix 10).

6.5.7 The effect of age

There was evidence of interaction on both multiplicative and additive scales between age (≤55/>55) and weighted score (≤9/≥10) for both all-cause and CVD mortality (Table S5, Appendix 10). Here a negative multiplicative interaction (all-cause mortality HR 95% CI 0.88 [0.83, 0.94]) indicated that the effect of scores ≥10 was greater in those aged ≤55 than those >55. However, there remained evidence of additive interaction for both all-cause and CVD mortality with a much larger ‘effect’ for age >55 than scores ≥10 (Table S5, Appendix 10). The absolute rate of deaths increased markedly with age and models of

combined associations showed, compared with the youngest with very low scores, HRs were highest in the oldest group with very high scores (Figure S8, Appendix 10). Stratified by age, the pattern of associations was broadly similar across age categories (Figure S9, Appendix 10). However, consistent with the negative multiplicative interaction above, compared with those in the very low score category, HRs for those with very high scores were generally lower for those aged >65 and highest in those aged >45 to ≤55 for both all-cause and CVD mortality.

6.5.8 Sensitivity analyses

Supplementary material 2 (Appendix 11) contains a more detailed description of sensitivity analysis results. Using CVD mortality associations to create the score (Table S7 and Figure S10, Appendix 11), again resulted a bimodal distribution of the weighted score comprising non-smokers and smokers (Figure S11, Appendix 11). There was a dose-response association between this weighted score and CVD mortality and with larger increments among smokers (Table S8, Appendix 11).

Similar results were observed after excluding those with CVD, cancer, or an early death (Tables S9-S11, Figures S12-S22, Appendix 11).

6.6 Discussion

6.6.1 Summary of findings

This study shows how a new weighted lifestyle score comprising 11 LFs is associated with all-cause and CVD mortality. The strength of association between the weighted score and mortality outcomes was greater among more deprived participants. While the strength of association between weighted score and mortality outcomes was similar among women and men, markedly higher risks overall were seen in men compared with women. Higher mortality risks associated with the weighted score were seen in the oldest (>65) compared with youngest (≤45) group overall. However, within age categories, there were

relatively stronger associations with mortality in a middle-aged group (>45 to ≤55). Results were similar across sensitivity analyses.

6.6.2 Comparison with previous literature

Consistent with previous studies, the 11 LFs examined had varying strengths of independent associations with all-cause and CVD mortality.[366] Thus, a weighted lifestyle score could guide prioritisation of LFs for support. Smoking had markedly stronger associations compared with other LFs, consistent with smoking as a leading risk factor for adverse health outcomes.[1] However, the distribution of the weighted score highlighted the clustering of LFs, underscoring potential benefits of supporting behaviour change across a range of LFs.[82] It also suggests how the lifestyle score may be especially helpful for smokers, among whom there was greater co-occurrence of LFs, as these could be prioritised for smokers not yet ready to quit. Further, the stronger dose-response relationship between the score and mortality among smokers than non-smokers, indicated a potential interaction between smoking and the other LFs. Healthy change for other LFs could see a disproportionate benefit for smokers, which could be motivating for smokers not yet prepared to quit.

The disproportionately strong associations between weighted lifestyle score and mortality outcomes in deprived participants in this and other studies further strengthens arguments to target deprived populations for lifestyle support.[24,293,357] A measure of deprivation could be included in lifestyle scores to convey lifestyle risk more accurately. Adding deprivation to risks scores is not new (e.g., ASSIGN and QRISK) but has not been applied to lifestyle scores of behavioural factors.[367,368] While the powerful role of social determinants of health was cited as influencing the recent update of AHA's lifestyle score, a socioeconomic measure was not included.[356]

Lifestyle score risks here were greater overall in men, but in stratified analyses we found the relative effect of the weighted lifestyle score appeared slightly greater in women, particularly for CVD mortality. This differs from results of a meta-analysis which found similar effect sizes for LF combinations in both women and men for all-cause and CVD mortality.[72] Our result is consistent with known sex differences in the magnitude of excess risk conferred by LFs, such as

smoking having a relatively greater impact on CVD risk in women than men.[358] The implication is that there may be disproportionate benefits for women who make healthy behavioural change across these 11 LFs but greatest population health benefit would come from addressing multiple LFs in men. This chimes with the recognised need for interventions targeting physical activity and diet in men, but our study suggests there would be benefits in broadening this to include additional LFs.[369]

Our analyses of the effect of age showed a strong association between weighted score and mortality across all age groups but also suggested that middle-aged groups (e.g., <65 years) could be targeted. Consistent with this, sub-group analysis from a meta-analysis showed stronger associations between combined LFs and CVD mortality in those <60 compared with those aged >60.[72]

6.6.3 Strengths and limitations

Rich and linked prospective data provided by UK Biobank allowed assessment of all-cause and CVD mortality associated with a wide range of LFs, including less considered factors (e.g., social participation). We were able to use individual LF effect sizes to create a weighted lifestyle score while adjusting for numerous potential confounders. Despite potential for interactions between LFs and, therefore, the importance of mutual adjustment, few studies control for as many LFs as this study. The large cohort size permitted sensitivity analyses where those with poor baseline health were excluded whilst retaining statistical power.

As with all observational studies, there is a chance of residual confounding. There is also a risk of collider bias as UK Biobank participants are more affluent, less ethnically diverse, and have fewer LFs than the UK general population.[247] The lack of ethnic diversity curtailed our assessment of the effects of ethnicity. Multiple testing is an issue as we performed numerous tests and, therefore, we only draw on findings that were consistent across analyses. Misclassification of LFs through misreporting could have influenced associations. LFs can change over time and the assessment of LFs was on average 12 years prior to the study end.[370] However, lifestyle trajectories generally remain stable as those with unhealthy LFs tend to retain them over time. Differences in misreporting by

characteristics (e.g., sex) could have affected our results but whether these differences exist is unclear. Dichotomisation of LFs oversimplifies relationships between LFs and health and interactions between LFs. Future studies could assess LFs in greater detail, both in terms of trajectories with multiple assessments over follow up and in terms of exploring LFs as continuous variables. A direct comparison with other lifestyle scores would be important as scores vary in their performance of predicting adverse outcomes. Further, many other outcomes (e.g., cancer/dementia) are associated with LFs and could be included when examining these relationships and designing lifestyle scores. Lifestyle scores could be part of precision medicine that includes precision over LFs and socioeconomic circumstances.[6] With widening health inequalities and parallel widening LF inequalities there is an urgent need to support improving multiple LFs in deprived populations.

6.7 Conclusion

This study has created an extended weighted lifestyle score that could convey personalised lifestyle mortality risk and inform policy and practice for harm reduction. It also shows the higher mortality risk associated with a weighted lifestyle score both among more deprived populations and men. Replication of these findings in different datasets would strengthen the case for use of this lifestyle score in clinical practice.

7 Exploring public, practitioner, and policy maker perspectives of unhealthy lifestyle factors in the context of socioeconomic deprivation: A qualitative study.

7.1 Chapter summary

This chapter addresses research objective 4) to explore how key stakeholders view combinations of lifestyle factors in the context of deprivation. The text is presented as the manuscript submitted to BMC Public Health in May 2024.

7.2 Abstract

7.2.1 Introduction

Understanding public and professional views on unhealthy lifestyle factors in the context of socioeconomic deprivation is critical to intervention development and reducing harm associated with such lifestyle factors in the most disadvantaged sections of society. The aim of this study was to explore the views of the public, healthcare professionals, and policymakers in relation to wider combinations of unhealthy lifestyle factors in the context of socioeconomic deprivation.

7.2.2 Methods

Adults in Scotland were recruited between August 2022 and June 2023 via social media, third sector organisations, and professional networks. Twenty-five members of the public took part in four focus groups. Eighteen semi-structured interviews were conducted with professionals: 12 primary care practitioners and six public health practitioners and policy makers. Reflexive thematic analysis was undertaken.

7.2.3 Results

Four main themes were developed from the data: 1) Evolving number and complexity of lifestyle factors – the number of lifestyle factors which adversely impact health has grown, with increasingly complex interactions and impact on people's lives, 2) Social determinants of lifestyle – numerous and directly influencing links were made between socioeconomic conditions and unhealthy lifestyle factors by all participant types, 3) Poverty as a fundamental social determinant – poverty was identified as a core and exacerbating factor in influencing unhealthy lifestyle factors, 4) Agency versus structure in influencing lifestyle – individual agency to address lifestyle factors was seen as limited by the structural constraints of people's lives. Among professionals, understanding the challenging social determinants of unhealthy lifestyle factors was countered by a desire to support healthy change in individuals affected by socioeconomic deprivation.

7.2.4 Conclusion

Public and professional views around lifestyle factors highlight an evolving complexity of new and more lifestyle factors and their interplay. Their views around the social determinants of lifestyle and structural limits to agency around lifestyle strengthen arguments for a reduced emphasis on individual-level responsibility for unhealthy lifestyle factors as well as deeper integration of social determinants into lifestyle interventions. In addition to addressing poverty and socioeconomic inequalities directly, innovative policy, planning, and legislation that incorporates wider approaches could tackle upstream determinants of numerous unhealthy lifestyle factors simultaneously.

7.2.5 Patient or Public Contribution

Members of the public who participated in this study have made contributions by sharing their views and perspectives. The National Health Service Research Scotland Primary Care Patient and Public Involvement (NRS PPI) Group contributed to the development of this work. The NRS PPI Group was consulted as part of preparatory work for HMEF's doctoral thesis funding application The findings of the qualitative work were presented to them and they informed the interpretation of those findings and related work presented at conferences and public engagement events.

7.3 Introduction

Unhealthy lifestyle factors such as smoking, alcohol consumption, poor diet, and physical inactivity are leading risk factors for noncommunicable diseases and premature death.[371] A wider range of lifestyle risk factors are emerging,[7] including poor sleep,[57] sedentary behaviours,[372] updated dietary measures,[373] and markers of social connection (e.g., social participation levels).[17]

Unhealthy lifestyle factors tend to cluster in individuals and combinations of unhealthy lifestyle factors have stronger associations with adverse health outcomes than single ones.[15,374] Further, there is a well-recognised and steep socioeconomic gradient in both mortality and lifestyle factors: more deprived populations have higher mortality and higher prevalence and clustering of unhealthy lifestyle factors.[15,23] Unhealthy lifestyle factors are also associated with disproportionate harm in socioeconomically deprived populations.[375]

Therefore, the evidence strongly suggests that even greater lifestyle support for individuals and communities affected by deprivation would tackle a wider range of lifestyle factors and their clusters and would maximise health benefits. While population-level interventions to improve lifestyle factors can have disproportionate benefit among more deprived populations,[376,377] successful interventions that are proportionate to need ('proportionate universalism') at individual or community-levels are rare.[142,378,379] One reason for this lack of success is likely to be the tension inherent in how interventions and policies are operationalised, namely at the level of individuals, communities, or wider society.[138] Barriers to progress may also be due to the lack of integration or acknowledgement of 'upstream' drivers in interventions targeting individuals or communities.[380]

One way of considering the drivers of unhealthy lifestyle factors so that they may be better incorporated into individual and community-levels lifestyle interventions is through the lens of structural or social determinants of health.[142,143] The ability to change and sustain healthier ways of living requires not only individual agency, the capacity of individuals to make lifestyle choices, but also wider social and political environments that enable healthy living.[124] These wider 'upstream' structural conditions determine the extent to which

individuals can modify behaviours and sustain healthier ways of living over a lifetime.[381] For example, socioeconomic status and financial, employment, and environmental factors all intersect in ways which influence choices and opportunities.[382] However, lifestyle factors are often viewed as matters of individual agency or responsibility; thus lifestyle interventions have often drawn on psychology or economics with an aim to inform individuals why and how to make healthier choices.[30,383] Indeed, even when interventions start with a focus on upstream structural determinants, they often 'drift' downstream to rely on individual-level motivation and shift responsibility to individuals.[384,385]

At least three issues may contribute to the focus on and drift towards individual responsibility for lifestyle. Firstly, the commercial determinants of health mean that corporations are incentivised to resist control and perpetuate the paradigm that targeting individual responsibility is the optimal mechanism by which to reduce harm from their products (e.g., tobacco or energy-dense-nutrient-poor food).[386] Secondly, individual responsibility for lifestyle is often the predominant narrative in print media, a potential source of influence on public perspectives and, therefore, on support for policies and interventions.[387,388] Thirdly, there remains a lack of appreciation among decision makers that lifestyle factors are social practices carried out within peoples' lives and characterised by social networks embedded in local contexts that facilitate and limit individual and group choices.

Similarly, shared characteristics of individuals and communities, which affects their 'risk of risks', are often viewed as only distantly or not directly or causally related to lifestyle factors.[389] For example, poverty and racism are shared experiences for populations which act as 'fundamental causes' of disease but are rarely addressed directly in lifestyle interventions.[390,391] Therefore, lifestyle interventions with a focus on individuals and communities affected by deprivation could be improved by accounting for how the intervention will work when fundamental causes or upstream drivers remain operational.

New understanding of how to optimise support for healthy living for individuals and communities experiencing deprivation could be found: 1) in the growing range of lifestyle factors identified as important risk factors, 2) by considering

several lifestyle factors simultaneously to help individuals prioritise a suite of risk factors and increase awareness of the interactions between lifestyle factors,[392] 3) via the clustering of unhealthy lifestyle factors among individuals and communities affected by deprivation, which presents opportunities to address these collectively, and 4) via deeper integration and consideration of the social and commercial determinants of health into individual and community-level interventions. Exploring key stakeholders' views of these concepts could inform development of future lifestyle interventions and contribute to overcoming existing barriers to successful interventions.[197] However, few studies have examined stakeholders' perceptions of the growing range of lifestyle factors and their interaction with social contexts. This qualitative study aimed to explore the views of the public, healthcare professionals, and policymakers in relation to wider combinations of unhealthy lifestyle factors in the context of socioeconomic deprivation. The aim was not to ask for views on how to improve lifestyle factors, although those views were expected and welcomed, but to explore views around lifestyle factors more generally and understand how they are conceptualised and considered from key stakeholder perspectives.

7.4 Methods

Work reported here formed the final part of a mixed methods study, which had the overall aim of understanding the risks of mortality and morbidity associated with combinations of unhealthy lifestyle factors across the socioeconomic spectrum.

7.4.1 Study design

This study used focus groups with members of the public and one-to-one semi-structured interviews with professionals conducted between August 2022 and June 2023.

7.4.2 Study recruitment and participants

We used social media to recruit adult members of the public who were resident in Scotland. The study was advertised via the authors' personal and wider

professional social media channels including University affiliated channels and those of third sector partners (The ALLIANCE and The Poverty Alliance). We recruited healthcare professionals ('practitioners') via professional networks including The Scottish Deep End Project using a combination of social media and group email adverts. Policymakers and public health professionals (PPHPs) were recruited via email using a combination of snowball sampling, existing professional networks, and contact with relevant organisations (Convention of Scottish Local Authorities). Professional participants had to be working in Scotland to be eligible. We focussed on Scotland to provide potentially more contextually relevant findings; health is a devolved responsibility with often clear policy and organisational differences between the UK's devolved nations.

We collected data from 25 members of the public in four focus groups. Public participants were aged between 24 and 78 years; 18 were women, 7 were men. Occupations included managers/senior officials, professional occupations, skilled trades, sales occupations, plant and machine operatives, and elementary occupations (major groups of the Standard Occupational Classification 2000),[393] while some worked part time or were unemployed, students, or retired. For interviews with professionals, we aimed to recruit those who either provide face-to-face lifestyle or health behaviour advice or who have day-to-day public health or policy making experience directly applicable to unhealthy lifestyle factors or social determinants of health. Semi-structured interview data were collected from 18 professionals (six community links workers [CLWs: non-clinical practitioners based in primary care in areas of socioeconomic disadvantage, who work with patients to navigate health and community services],[394] one community nurse, one community pharmacist, four general practitioners [GPs], three public health professionals, and three policymakers).

7.4.3 Data collection

Three focus groups and all interviews were conducted online via video conferencing software (Zoom or MS Teams). One focus group took place face-to-face on university premises. Focus groups lasted approximately 60-80 minutes; interviews lasted up to 60 minutes. Focus group guides and interview schedules were developed iteratively, drawing on findings from the project's related studies,[24,375,395,396] from review of pertinent literature, and from the knowledge

and interests of the study team, which included two academic GPs and a non-clinical scientist with an interest in health prevention and health inequalities.

7.4.4 Data analysis

Audio data were transcribed, anonymised, and analysed using reflexive thematic analysis.[261,397] A reflexive approach to the data was chosen for several reasons. Firstly, it provided flexibility when analysing the pre-defined wide range of topics discussed in the focus groups and interviews. Secondly, because this qualitative work was informed by the wider project's related quantitative research (a systematic review and prospective cohort analyses),[24,375,395,396] a reflexive and interpretative approach allowed the researchers to acknowledge and utilise their prior understanding, varying fields of expertise, and literature bases to interpret the data.[397]

Transcripts were imported into NVIVO 14 to aid analysis. HF read all the transcripts in detail, with COD and FM reading a selection of early interviews. All three noted codes in the data, in part resulting from questions asked during data collection but also information that came from participants without prior elicitation. Regular analysis sessions were held to discuss the codes. HF continued to code the remaining transcripts, meeting regularly with COD to discuss and refine these. An initial 'long-list' of codes was refined, for example merging codes that described similar data. Codes were then grouped into common areas or themes. Later stages of coding and refinement included discussing the coding and interpretation with FM and presenting the analysis to an advisory board and to a PPI group.

7.5 Results

Data are presented in relation to four generated themes – evolving complexity of unhealthy lifestyle factors, the social determinants of lifestyle, poverty as a fundamental determinant of lifestyle, and agency versus structure in relation to lifestyle factors. Example quotes are provided verbatim (in italics) for each theme. The corresponding participant type, gender (F/M), and, for public participants, age in years is given in parentheses following each quote.

7.5.1 Evolving complexity of unhealthy lifestyle factors

All participants described a wide array of unhealthy lifestyle factors and how this has evolved over time to become more complex. While public participants often referred to 'conventional' unhealthy lifestyle factors such as smoking and diet, they also described a broader range of unhealthy lifestyle factors including social media use, use of digital technology/smartphones, and inadequate sleep. Public participants frequently suggested direct links between unhealthy lifestyle factors and individual-level social circumstances (e.g., shift work affecting sleep), but also characterised wider societal changes such as new technology or increasing levels of loneliness as unhealthy lifestyle factors themselves.

> *I think loneliness has a big impact, and I think it's getting worse, because there's a lot more single people. And aye that came to roost, sort of during COVID, because people suddenly realised that they were alone, and you know, it's very hard. I live alone, and loneliness can be, how can I put it, it can swamp you, you know, it can overwhelm you.* (Public participant, 58F)

Practitioners also recognised a wider range of unhealthy lifestyle factors and described a shift from a relatively narrow focus on the more 'traditional' factors of smoking, alcohol consumption, diet, and physical activity to a much wider concept that encompassed anything that people do regularly which could influence their health. This included sleep, medication concordance, health seeking behaviours (e.g., attending medical appointments), breastfeeding, gambling, sexual risk behaviour, caffeine or energy drink consumption, social media and mobile phone use, social isolation, exposure to sunshine, and time spent in green or blue space. However, smoking and alcohol consumption were still cited as the most important lifestyle factors affecting health.

> *'Yeah. I mean I think the very traditional…maybe the bit that I would have said if I was ten years ago is, yeah, things about diet, activity, which I would have called exercise, smoking, alcohol, drugs, I may or may not have mentioned, but I think those are the core bits.'* (GP, F)

PPHPs had the broadest view of unhealthy lifestyle factors, discussing unhealthy lifestyle factors in terms of both individual lifestyle factors and upstream social drivers that influenced opportunities and available choices.

> *So when I think about health behaviours, I think actually about what's driving those behaviours and the environment that are leading them to do that.* (PPHP, M)

The growing complexity of unhealthy lifestyle factors was characterised as an expansion of the number of unhealthy lifestyle factors to include more facets of daily living as opposed to a shift of focus away from more traditional unhealthy lifestyle factors. The language used to describe unhealthy lifestyle factors had also evolved. Practitioners described how 'exercise' was now referred to as 'physical activity', to acknowledge an awareness that physical movement and activity in a broader sense was associated with health benefits. CLWs used language such as 'social connection', highlighting awareness of broader and newer concepts that encompass social isolation and loneliness. Public participants also described awareness of more 'up to date' terminology for conventional factors (e.g., 'excessive alcohol consumption' as opposed to 'drinking').

> *Yeah, you know, I would say, negatively, you know, I wouldn't just say, drinking of alcohol, yeah, I would use the word, excessive consumption of alcohol. You know, which some people, it has become a habit for them, and yeah, so this excessive consumption of alcohol actually brings harm to the body's system, yeah.* (Public participant, 26F)

For practitioners, the new language appeared to help them consider what patients might perceive as feasible in terms of healthy lifestyle change. For CLWs, this paralleled a more critical appreciation of social circumstances as they would weigh up what level or type of lifestyle change was feasible for individuals given their wider social context. This involved exploring and understanding relevant financial and practical considerations as well as the individuals' own perceptions of what was achievable while using a language of broader and more nuanced descriptions of unhealthy lifestyle factors.

> *…we really want to encourage people to become more active. And […] something is better than nothing, you know, like just start? Okay two minutes, it's fine, it's two minutes more than you were doing before…people are coming from a baseline of nothing'* (CLW, F)

> *...you know, that would be a great thing for [them] to go and do, it'd get [them] out, socialising, would help [their] mental health, but at the end of the day [they] can't afford to do it; and there's nothing on [their] doorstep.'* (CLW, F)

There was also an awareness of the complex interplay between unhealthy lifestyle factors. Public participants recognised that unhealthy lifestyle factors impact one another (e.g., screen time affecting sleep) but also that multiple unhealthy lifestyle factors increased the risk to health. Although single unhealthy lifestyle factors were often a focus for practitioners, often when specific factors were felt to be most relevant to a health problem, unhealthy lifestyle factors were rarely considered in isolation and practitioners seemed to perceive unhealthy lifestyle factors as social practices. For example, smoking, alcohol consumption, and socialising were described as a linked triad of behaviours whereby reducing one unhealthy lifestyle factor could impact on others. Likewise, discussions around sleep invariably led to consideration of other unhealthy lifestyle factors that impact sleep, such as screen time, and caffeine or alcohol consumption.

> *You will get people who will say, I can't do this because of that, and I can't do that because of this. Or if I cut down on this, I can't see my… Like, if I don't drink, I can't see my friends, because you only socialise through the pub. Or if I stop smoking, I can't see my friends because then I can't go for a drink, because if I do, I'll want to have a cigarette.* (Nurse, F)

PHPs and policymakers, while acknowledging clustering of unhealthy lifestyle factors within individuals, also described lifestyle factor interplay at the population-level. By considering the shared drivers for numerous unhealthy lifestyle factors, PPHPs suggested that several unhealthy lifestyle factors could

be addressed simultaneously via consideration of health and wellbeing more broadly and at the level of communities rather than via specific unhealthy lifestyle factors enacted by individuals.

> *What we haven't been able to do, has been to join those up and think, alright, well actually what does a healthy community look like, what would you have? What do you need in order to be healthy, and how do you mitigate some of these unhealthy influences in order that they don't cluster because there's nothing else in communities?* (PPHP, F)

7.5.2 The social determinants of lifestyle

All participants saw difficult social and economic circumstances as a direct challenge to healthy ways of living. From the financial cost of a healthy diet or a gym pass, through to transport availability to access health and wellbeing services, and ease of access to green space, participants drew numerous and detailed links between wider social circumstances and unhealthy lifestyle factors. Their perceptions of inextricable pathways from socioeconomic conditions to unhealthy lifestyle factors were voiced from the outset and some participants' definitions of lifestyle factors were indistinguishable from socioeconomic factors.

> *So, I think it's* [unhealthy lifestyle factors] *everything, exercise, what you eat, drinking, smoking, drugs, just your family history. And how you've been taught about cooking, and what education you've had. And just what kind of stress you're under, and the environment that you live in, I think'* (Public Participant, 58F)

Members of the public gave occupational status and education as examples of unhealthy lifestyle factors, which are not generally considered unhealthy lifestyle factors by healthcare and research professionals. There was also a distinction between perceptions of the public and practitioners with those of PPHPs. While public participants and practitioners often implied causal links between wider social circumstances and unhealthy lifestyle factors, they more readily discussed unhealthy lifestyle factors, and behaviour change, at the level of individual action. In contrast, PPHPs consistently took a broader view and

were reluctant to use the terms ‘lifestyle factor’ or ‘health behaviour’. Instead, PPHPs responded to questions around unhealthy lifestyle factors with reference to what they considered to be upstream drivers. ‘Lifestyle’ or ‘health behaviour’ were seen as unhelpful as they referred to downstream proxies that drew attention away from the upstream influences and drivers.

> *I suppose there’s also a general sense that the behaviour sits within a wider context of other things, that affect your health, and so I suppose, one of the things that I’m keen to do, is not overtly focus on the behaviour, at the expense of all the other things, that wider context that that behaviour is occurring in, that is also a factor in the outcome of poor health, whatever.’* (PPHP, F)

A focus on downstream unhealthy lifestyle factors was perceived as shifting blame to individuals and risked both vilifying people in difficult circumstances and increasing health inequalities. PPHPs saw unhealthy lifestyle factors, and the choices that individuals make, as inseparable from their social circumstances. The value of focussing on upstream drivers of unhealthy lifestyle factors was further apparent in PPHPs’ feelings of frustration by the siloing of funding and legal frameworks around unhealthy lifestyle factors (e.g., tobacco legislation, alcohol or fast-food licencing, local council planning regulations) which typically follow ‘traditional’ lifestyle factor boundaries. This was viewed as reducing opportunities to address a wider range of unhealthy lifestyle factors and healthy living more generally.

> *So alcohol licensing sits apart, tobacco licensing sits apart and the regulation of fast food premises and so on. What we haven’t been able to do, has been to join those up and think, alright, well actually what does a healthy community look like, what would you have?* (PPHP, F)

Related to blame and stigma, all participants linked the experience of discrimination in forms of racism, bullying, and perceived stigma to unhealthy lifestyle factors. Unhealthy lifestyle factors seen as providing ‘comfort’ or ‘a safe space’ from discrimination.

…sometimes people pick up these lifestyles of smoking and drinking as a result of being depressed from things like this [racism]. They see unhealthy lifestyles as a jacket of comfort to wear to just get on. I think inclusiveness and elimination of racism will go a long way. (Public Participant, 26M)

7.5.3 Poverty as a fundamental determinant of lifestyle

Poverty emerged as an all-encompassing factor that presented challenges to healthy living. Connections were made between poverty and more conventional lifestyle factors like smoking, alcohol, and drug use and also between poverty and emerging factors such as social participation levels, sedentary behaviours, and screen time.

I guess, constant stress and lack of resource. So, I mean, access to transport, things like that, people know they can't get around. So people tell me, well I can't do that [physical activity class], because I can't afford the bus, and the bus doesn't run, and therefore that's not an option for me...Or, you know, it's, I do want to eat more healthily, but I really can't afford it. (CLW, F)

Poverty meant not only a lack of financial and material resources which curtailed healthy lifestyle opportunities but also a lack of time and 'headspace' to consider healthier options which could even make an unhealthy choice a rational choice given the circumstances.

If you're a single parent, you know, managing the household and children, dealing with the precariousness of poverty and insecure work. So the time to then think about eating healthily and looking after yourself and going and doing some exercise is really limited… (PPHP, F)

…what struck me was that there's a logic here to making choices, and this is how the conditions you find yourself in financially, actually influence your choices about [lifestyle], it's not that you're necessarily choosing to have an unhealthy diet or whatever. (PPHP, F)

Poverty was also deeply linked to unhealthy lifestyle factors via its impact on mental health. From a lack of money leading to anxiety, to benefits sanctions causing distress, or a lack of hope due to poor housing, all made unhealthy lifestyle factors more likely. Links between mental health and unhealthy lifestyle factors ranged from the severe, such as drug or alcohol dependence following adverse childhood experiences, to less severe such as the 'highs' felt from sugary or processed food to *'give you a kick in an otherwise quite bleak day'* (GP, F). The links made between adverse childhood experiences or psychological trauma, deprivation, and unhealthy lifestyle factors were widespread and perceived as a challenge to lifestyle interventions.

> *There's reasons why people can't, you know, aren't being physically active and everything, and it's not normally just a simple fix. Most of the people we see are extremely complex and quite often have that past trauma and it's, it is then, it takes it away from, you know, just the simple…yeah, let's have a wee plan and this is what we do and we work to that. Because of so much that they're dealing with and all the things for us.'* (CLW, F)

Unhealthy lifestyle factors were seen as 'coping mechanisms' and as understandable responses to adverse living conditions and past experiences. Therefore, to make healthy lifestyle change, new coping mechanisms were required before people could address unhealthy lifestyle factors they may be relying on.

> *…So like alcohol or smoking, can also offer respite from things that are happening in someone's environment too. […] Like, say, for example, if somebody has like high anxiety and smoking helps them with that, then it's helping their health in another way, but it may be kind of […] Like if they've got kids or if they drink a lot of alcohol and they have a stressful upbringing or something like that, a trauma response, then that can subdue that.* (Public participant, 26F)

Poverty was thus a key driver that not only directly led to unhealthy lifestyle factors but also exacerbated or primed a host of other factors that negatively

influenced unhealthy lifestyle factors. For example, poverty was seen as impacting people's self-esteem and lowering their expectations for health, thereby increasing fatalism and acceptance of poorer health at a younger age. Poorer housing associated with poverty was felt to affect peoples' sense of safety and hope and linked to insufficient facilities to prepare healthy food.

> *So it's no coincidence that people in their sixties are walking round with sticks in certain areas, whereas in other areas they're not. We can all see that, but I guess the expectation about health and what people can accept, to a certain degree, is maybe influenced by what we can see around us. And I guess that's from little kids, right up. If you're used to seeing people in their sixties, with high levels of disability, it's not so shocking if that happens to you.* (CLW, F)

7.5.4 Agency versus structure in relation to lifestyle factors

Despite detailed descriptions of how challenging social circumstances made healthy ways of living difficult, most participants felt that individuals maintain some level of agency with which to make healthy choices. However, the opportunities for healthy choices or lifestyle change and the level of agency to make those choices were seen on a spectrum contingent on wider social circumstances. Circumstances felt to facilitate agency were not just socioeconomic factors but also psychological and related to timing, with practitioners often citing components of the Transtheoretical (Stages of Change) Model.[398]

> E*ven people who have been in the depths of substance use, can go on and make an amazing recovery, and live some of the healthiest lifestyles. But that's a journey they have to go on themselves. And it relies on people being with them at the right stage, at the right time, to open that understanding.* (CLW, F)

The awareness of structural challenges to agency around unhealthy lifestyle factors generated a tension or dissonance, particularly among practitioners and PPHPs, as they tried to estimate the level of agency that should be accorded to individuals considering the structural challenges that they knew the individuals

faced in daily life. Among practitioners, this tension manifested as a balancing act between their understanding of the social determinants of unhealthy lifestyle factors and trying to support and improve the health and wellbeing of individual patients. On one side there was an appreciation of the lack of opportunities for healthy ways of living, while on the other there was a strong sense of duty to instil hope that healthy change was possible and to cultivate any residual agency despite difficult circumstances. For PPHPs, although the role of the social determinants of unhealthy lifestyle factors dominated their views, an apparent tension between agency and structure was still manifest in their views. This was expressed in terms of PPHPs' concerns of a disproportionate focus on helping people already living with unhealthy lifestyle factors at the expense of wider prevention efforts and a desire to redress that balance without neglecting those with unhealthy lifestyle factors.

> *You need both prevention and the cessation, so we can't just prevent our way out of smoking and we can't just write off these tens of thousands of people who are smoking and basically leave them to die, so we have to help them. I think it's about having both approaches, so I think we need to think about what the proportions are that are appropriate to the issue. So it's not a one prevention is good, upstream is good, downstream is bad, that actually you do need to have both of them. But I think that a lot of the focus has been on the downstream stuff, and so a lot of what you see in public health is to make the case for upstream and prevention, and I think that's right. But we need to not forget about the downstream still needs to actually happen as well.* (PPHP, M)

7.6 Discussion

We have presented key stakeholders' perspectives on unhealthy lifestyle factors in the context of socioeconomic deprivation. This includes perceptions of an evolving complexity of unhealthy lifestyle factors, with a greater and wider number of health and social factors considered 'unhealthy lifestyle factors', a greater appreciation of interactions between lifestyle factors, and their understanding of the influence of socioeconomic conditions. Poverty emerged as a fundamental social determinant alongside the influence of social context on

people's lives in shaping both the prevalence of unhealthy lifestyle factors and the ability to enact and sustain healthy lifestyle change. The role of individual agency was viewed within these structural constraints, leading to tensions between a desire to support individuals lead healthier lives while recognising the limits posed by wider structural forces.

7.6.1 Comparison with literature

The expansion of unhealthy lifestyle factors over time from a few specific 'traditional' unhealthy lifestyle factors to a wider range of factors mirrors the changes observed in epidemiological literature.[7,10] Here, among practitioners, the expansion paralleled an increasingly nuanced and detailed exploration of people's lives and wider social circumstances in order to support healthy living and change, which highlights a growing understanding of the inextricable role of social determinants on lifestyle. This reflects the academic literature which has long considered the impact of poverty and other social determinants of health on lifestyle and on the differential impact on the health of poorer communities.[124,390,391]

The various perceptions of the role of social determinants and the extent of individual responsibility or agency for unhealthy lifestyle factors are important to discuss in view of the reliance on individual responsibility in previous interventions.[30] All participants recognised that limits to healthy living could be outside the control of individuals, while also accepting a role for individual responsibility. Views captured here suggest high levels of understanding of the influence of social contexts on lifestyle, including among members of the public, which is in contrast to dominant media portrayals.[388] A qualitative study exploring public views around food choice and related public health policies found there were numerous examples of an 'illusion of food choice' especially in difficult social and economic circumstances.[399] However, the same study also highlighted the range of views in terms of responsibility for healthy eating versus government regulation or influence on food choice. Practitioners' views here straddled both individual responsibility and wider social influences as their sense of duty to support individuals was balanced by their understanding of the individuals' wider social context. PPHP views aligned with a greater role for social determinants as they expressed a need to redress the individual-level

unhealthy lifestyle factors albeit without wanting to neglect those living with unhealthy lifestyle factors. The discrepancy of views between professional types is likely to be linked to their professional perspectives: individual versus population. Therefore, a distinction between 'sick individuals and sick populations' is seen here in the perceptions of professionals dealing with unhealthy lifestyle factors of individuals versus those addressing unhealthy lifestyle factors of populations.[400] A recent qualitative study of GPs discussing obesity found, perhaps unexpectedly, that those working in more deprived areas were more likely to ascribe individual responsibility for obesity than GPs working in more affluent areas.[401] The authors suggest that more challenging working conditions, greater 'empathy fatigue', and more frequent experiences of treatment failure (ineffective weight management options) among those working in more deprived areas could result in feelings of frustration and blaming patients for obesity. This was not apparent in this study, despite all the GPs and most of the other practitioners working in areas of high deprivation. However, practitioners in this study acknowledged both individual-level responsibility as well as social circumstances as influences over unhealthy lifestyle factors.

Participants' views of the greater role of social determinants for lifestyle and of agency limited by structure fit with the health lifestyle theory conceptualised by Cockerham.[120] Cockerham develops previous sociological theories to describe a paradigm where 'life choices (agency)' and 'life chances (structure)' interact to influence individuals' 'dispositions to act (habitus)' and explain unhealthy lifestyle factors as wider social practices. Cockerham also describes sociological legacies that have resulted in an over-emphasis on individual-choice over opportunities and develops health lifestyle theory with the aim to 'bring structure back'.[120] The desire to support individual-level lifestyle change seen among practitioners and PPHPs here may also contribute to 'lifestyle drift', which is:

> 'the tendency for policy to start off recognizing the need for action on upstream social determinants of health inequalities only to drift downstream to focus largely on individual lifestyle factors'.[30,384]

Further, the desire to support agency around lifestyle and instil hope at the same time as appreciating the structural limits to life chances resonates with a

similar finding of ‘contradictory’ beliefs in an ethnographic study of health-promoting community workers.[402] Moreover, even terms concepts such as ‘upstream driver’ or ‘social determinants’ could also contribute to lifestyle drift if they are perceived by decision makers as concerned with issues more distantly related to unhealthy lifestyle factors when in fact, and in line with views seen here, they are much closer and causally linked.

7.6.2 Implications

The increasing complexity in unhealthy lifestyle factors – the increasing number, type, definition/concept, and deeper appreciation of the interactions between lifestyle factors - seen here suggests that practitioners and those who support individuals with unhealthy lifestyle factors would benefit from increased resources and training on supporting healthy lifestyle change. The perceived causal influence of social determinants for unhealthy lifestyle factors means greater integration and direct acknowledgement of these determinants in individual and community-level interventions would be acceptable but rarely exists in extant interventions.[380] For example, perceived interpersonal and structural racism is associated with unhealthy lifestyle factors (e.g., poorer diet) but rarely acknowledged or directly addressed in health behaviour interventions.[403-405]

One way of greater integration of social determinants into lifestyle interventions might be in considering the conscious workload required by an intervention. Marteau et al. propose that interventions can target either conscious (e.g., providing risk information or signposting to resources) or non-conscious (e.g., taxation incentivising reformulation of harmful products) processes. Interventions targeting conscious processes requires conscious effort from individuals to enact change compared with those that target non-conscious processes.[406] Highlighting the conscious processes involved in interventions could help policymakers and researchers consider the resources required of individuals to make and sustain healthy lifestyle change. This could help mitigate against the risk of exacerbating health inequalities via interventions that target conscious processes as these ‘make higher demands on people’s cognitive, social, and material resources’.[406] For example, the lack of time and head space seen in this study’s participants’ characterisation of poverty is consistent with

lower conscious processing resources in communities experiencing socioeconomic deprivation and explain why an intervention might exacerbate health inequalities. Similarly, highlighting the burden of work required to lead a healthy life or a 'prevention burden' offers a way of viewing and supporting unhealthy lifestyle factors that could account, at least in part, for the social determinants of unhealthy lifestyle factors.[407,408] Alternatively, and consistent with how participants viewed the interrelationships between social circumstances and lifestyle and between lifestyle factors themselves to form social practices, Frohlich et al. suggest studying the relationships between agency, lifestyle (social practices), and structure to provide deeper understanding of what health inequalities are and propose an idea of 'collective lifestyles' that can bridge structure and agency and explain unhealthy lifestyles.[409]

The wide and deep links between wider socioeconomic circumstances and unhealthy lifestyle factors identified in all key stakeholders' views supports a reduced emphasis on individual-level responsibility for unhealthy lifestyle factors. This strengthens an already strong evidence base that shows while individual or community-level lifestyle interventions may have an important role to play, policy and decision makers should strongly prioritise population-level interventions to improve lifestyle factors. Extensive evidence for an 'effectiveness hierarchy' exists for numerous unhealthy lifestyle factors (smoking, alcohol consumption, diet, and physical activity) whereby population-level interventions are more effective, equitable, quicker, and cost efficient than individual or community-level interventions.[410] The views captured here suggest that these key stakeholders would support a focus on population-level lifestyle interventions.

The perspectives of agency versus structure in unhealthy lifestyle factors identified here suggests future research examining whether and how practitioners balance expectations for healthy living and their desire to support healthy change with their deep appreciation of the structural determinants of unhealthy lifestyle factors could inform practice and interventions. For population-level interventions, perceptions here support focussing on broader concepts of healthy living and wellbeing and rethinking the siloing of legislation,

policies, and funding streams by different unhealthy lifestyle factors. This qualitative study also has direct implications for further planned work, including the development of a weighted lifestyle score for clinical use through co-design with communities affected by deprivation.

7.7 Conclusion

Public and professional views of unhealthy lifestyle factors in the context of socioeconomic deprivation delineate nuanced views of evolving complexity and interplay of unhealthy lifestyle factors and their underlying socioeconomic drivers. Among practitioners, deep appreciation of the social determinants of unhealthy lifestyle factors coexisted with a sense of duty to instil hope for healthy lifestyle change in individuals affected by socioeconomic deprivation. However, perceptions across participants captured here diminishes the role of individual-level responsibility and agency for healthy choices and supports closer integration of social and behavioural determinants of health. Innovative policy, planning, and legislation is required to incorporate wider approaches that can tackle upstream determinants of numerous unhealthy lifestyle factors simultaneously.

8 Discussion

8.1 Chapter summary

This chapter provides a discussion of the findings presented in Chapters 4-7 and is structured as follows: 1) summary of the findings for each of the four research questions (RQ1-4), 2) synthesis of all the findings contextualised in relation to existing literature to inform current theory and practice, 3) strengths and limitations of this thesis, 4) implications for practice and policy, 5) suggestions for future research, and 6) final conclusion.

8.2 Summary of findings

8.2.1 RQ1: What is the existing evidence for socioeconomic influences on the association between combinations of lifestyle factors and adverse health outcomes?

A systematic review without meta-analysis was used to collate and synthesise the available evidence to answer this question. Six prospective studies of five cohorts from Japan, UK, and USA were found that fit the inclusion criteria and that examined the potential influence of SES on the association between combinations of unhealthy lifestyle factors and adverse health outcomes. In those studies, there was wide heterogeneity in:

1. the lifestyle factors and SES indicators examined,

2. the thresholds or definitions used to classify unhealthy levels for each lifestyle factor,

3. the methods used to examine for interactions between SES and combinations of lifestyle factors, and

4. study quality.

However, all studies examined a measure of alcohol intake and physical activity, and smoking was missing from relevant analysis in only one study. Further, all studies examined all-cause mortality as at least one of the outcomes and all

used Cox-proportional hazard ratios to model the associations between exposures and outcomes.

The results of included studies were mixed, with higher mortality and morbidity risks associated with combinations of unhealthy lifestyle factors seen in the most advantaged SES group in some studies and in the least advantaged group in others. Therefore, with few and heterogenous studies and results to draw on, the literature available to answer this RQ is limited. Nevertheless, in four studies examining five cohorts, and via their results modelling the combined associations between combinations of lifestyle factors, SES, and adverse health outcomes, the evidence indicated that SES had an additive effect. This would suggest that the detrimental health effects from combinations of unhealthy lifestyle factors experienced by populations can be felt in addition to (and are not masked by) further detrimental health effects due to less advantaged SES.

8.2.2 RQ2: What are the associations and interactions between components of social connection and adverse health outcomes?

Analysis of the prospective cohort UK Biobank was used to address RQ2. This involved examining the independent and combined associations as well as interactions between two functional and three structural components of social connection and all-cause and CVD mortality. This study found:

1. that each measure of social connection was, after mutual adjustment, independently associated with all-cause and CVD mortality.

2. a potential threshold effect for one of the structural components, whereby a lower risk of mortality associated with increasing friends and family visits was maximal at a level of monthly visits i.e., visits more frequent than monthly were not associated with any lower mortality risk.

3. an interaction between living alone and other structural components. While living alone itself was associated with higher risks of all-cause and CVD mortality, the combinations of living alone and never having friends and family visits or living alone and not engaging in weekly group activity were associated with even higher risks.

4. that strong associations between living alone or never having friends and family visits and adverse health outcomes appeared to mask beneficial lower risk of mortality associated with regular group activity.

5. that while structural components of living alone and never having friends and family visits appeared to have stronger associations with all-cause and CVD mortality compared with functional components, a lack of both functional and structural components of social connection had even stronger associations with mortality.

Together these findings highlight that even with somewhat crude measures of social connection it is possible to explore some of the complexity of social connection in its association with adverse health outcomes. Further, the results showing evidence of an additive interaction highlight the importance of considering the combined influence of functional and structural components together.

In terms of the wider thesis objective of exploring components of social connection so that a measure of social connection could be included in a weighted lifestyle score, these analyses showed that each of the available measures of social connection in UK Biobank could be included. However, with the focus on producing a lifestyle score that could be used to support individuals to enact healthy lifestyle change it was decided to focus on those measures of social connection that could be more amenable to individual or community-level action and which could be supported by health and allied professionals. Therefore, these two structural components of social connection, namely, frequency of friends and family visits and engaging in weekly group activity, were included in the next set of quantitative analyses that address RQ3. Further, based on the results of the social connection analysis, friends and family visits less frequent than monthly was classified as unhealthy.

This informed the creation of the dichotomous social participation lifestyle factor variable with low levels of social participation defined as either having friends and family visits less often than monthly or not engaging in weekly group activity.

8.2.3 RQ3: How does socioeconomic status influence the association between a weighted lifestyle score and adverse health outcomes?

Analyses of the prospective cohort UK Biobank was again used to address this research question. There were three main aspects to these analyses, the findings of which do not all directly address RQ3 but are nonetheless included here as they relate to the wider thesis aim of understanding the associations between lifestyle factors and adverse health outcomes and understanding how those associations might vary in different socioeconomic contexts.

The first main aspect of the analysis involved creating the weighted lifestyle score. The weightings to be applied to each lifestyle factor in the score were generated by estimating the independent associations of 11 individual lifestyle factors and all-cause mortality. Each of the lifestyle factors, after mutual adjustment, were associated with all-cause mortality. Smoking had the strongest association, resulting in a score weighting of 14 points. High alcohol intake, low physical activity, low social participation, high TV time, short/long sleep, and low fruit and vegetable intake all had similar strengths of association with all-cause mortality resulting in weightings of 2 points each. Low oily-fish, high red meat, and high processed meat intakes, and high added salt frequency had weaker associations and a weighting of 1 point each. Weightings of each unhealthy lifestyle factor were summed so that each participant was assigned a weighted lifestyle score (maximum unhealthy score was 30). Both the continuous score and the score as a 5-catgeory variable (<5, 5-9, 10-14, 15-19, >19) had near linear dose-response relationships with all-cause mortality.

The second main aspect of the analysis involved examining how socioeconomic deprivation affected the association between the weighted lifestyle score categories and all-cause and CVD mortality. The results showed deprivation heightened the risk associated with an unhealthy weighted lifestyle score. Specifically, there was a multiplicative interaction between weighted score and deprivation for both outcomes and stratified analyses showed associations between weighted score and outcomes were stronger with increasing deprivation. Further, combined associations between weighted score, deprivation, and outcomes showed participants with the unhealthiest scores

from the most deprived backgrounds had the strongest associations with the mortality outcomes.

The third main aspect of the analysis involved examining the effect of sex, age, and ethnicity on the association between weighted lifestyle score and outcomes. These mortality analyses showed that for an unhealthy compared with a healthy weighted score there were 1) markedly higher risks overall for men compared with women, 2) higher risks overall for older rather than younger participants but relatively higher risks in the middle age groups (>45 to ≤55 and >55 to ≤65), and 3) no discernible differences by ethnic group.

8.2.4 RQ4: What are the perceptions of key stakeholders around combinations of lifestyle factors in the context of socioeconomic deprivation?

To answer RQ4, perspectives of key stakeholders were collected via four focus group discussions with members of the public and 18 semi-structured interviews with relevant professionals. Transcripts were analysed qualitatively using reflexive thematic analysis. Four major themes were identified and summarised below:

1. **Evolving complexity of unhealthy lifestyle factors**

   Stakeholders described an increasing number of lifestyle factors or health behaviours known to be associated with adverse health outcomes and which were considered for healthy change in both practice and policy. Evolving complexity included changing terminology and concepts around lifestyle factors (e.g., exercise now known as a broader concept of physical activity) as well as the appearance of new technology or other behaviours only more recently considered lifestyle factors (e.g. social media use or time spent immersed in natural environments). Complexity was also perceived in an increasingly complex interplay between lifestyle factors. This interconnectedness of lifestyle factors was often expressed

as a problem or challenge related to healthy change as addressing one lifestyle factor could be contingent on or inadvertently affect another.

2. **The social determinants of lifestyle**

The social determinants of unhealthy lifestyle factors were perceived to be widespread with difficult social and economic circumstances seen as a direct challenge to healthy ways of living. Among all participants, perceptions of lifestyle factors were closely intertwined with socioeconomic factors so that some definitions given for lifestyle factors were indistinguishable from what are more usually considered to be socioeconomic factors. Among public health practitioners, lifestyle factors were considered proxies for or extensions of socioeconomic factors. These practitioners also perceived an historic focus on the individual and behavioural aspects of lifestyle factors that distracted attention away from addressing the fundamental socioeconomic or structural drivers for unhealthy lifestyle factors.

3. **Poverty as a fundamental determinant of lifestyle**

Poverty was frequently used to mean a single and all-encompassing factor that explained unhealthy lifestyle factors. Poverty was seen to bring about unhealthy ways of living via numerous dynamic and interacting mechanisms including reduced access to financial and material resources. Other examples of mechanisms in which poverty either directly or indirectly led to an unhealthy lifestyle were loss of hope, previous psychological trauma, poorer mental health, lower self-esteem, lower health expectations, stigma, and the lack of ‘head space’ in which to consider healthy choices due to competing priorities. Through these mechanisms unhealthy lifestyle factors were seen as a logical or reasonable response to, or as offering relief from, or as functioning as a coping mechanism for challenging socioeconomic circumstances.

4. **Agency versus structure in relation to lifestyle factors**

   A theme of agency versus structure in relation to lifestyle factors was characterised by perceptions of a spectrum of what healthy lifestyle change was possible for individuals in view of their deep and wide appreciation of socioeconomic drivers for unhealthy lifestyle factors. For some professionals this created a tension that manifested as seeing a need to and wanting to offer hope for positive and healthy change on one side versus a detailed understanding of the major challenges faced by individuals due to structural inadequacies that were far beyond individual control. This also manifested as a balancing act for the professionals that support individuals with lifestyle changes as they weighed up what level of agency over lifestyle could be expected of individuals given the practitioner's understanding of their socioeconomic circumstances.

8.3 Synthesis to inform theory and practice

As stated in Chapter 3 (Methodology) this synthesis was informed by methods that attempt to combine findings from mixed methods research (quantitative and qualitative findings).

8.3.1 Differential vulnerability to multiple unhealthy lifestyle factors

The systematic review identified mixed evidence for differential vulnerability to a combination of unhealthy lifestyle factors in lower SES populations. Of the six included studies, four reported multiplicative interactions between a combination of unhealthy lifestyle factors and adverse health outcomes. However, two of those studies identified a negative multiplicative interaction where the relative risk of a combination of unhealthy lifestyle factors was higher in the higher SES group (and only in men in one of the studies).[289,292] Whereas the other two studies that observed a significant multiplicative interaction found the relative risk was greater in the lower SES group (and only in one cohort in one of the studies that examined two cohorts).[24,293]

The weighted lifestyle score analysis that forms Chapter 6 now adds to this evidence base, showing how deprivation can heighten the risk associated with

unhealthy lifestyle factors in a weighted lifestyle score comprised of 11 lifestyle factors.

Evidence from the weighted lifestyle score analysis in this thesis as well as the two studies in the systematic review that report multiplicative interactions in support of differential vulnerability all come from analyses conducted in UK Biobank.[24,293] Therefore, evidence from other datasets is needed to increase confidence in the findings around the potential synergy or multiplicativity of interaction between multiple unhealthy lifestyle factors and lower SES.

Further evidence for differential vulnerability came from the two studies that performed formal statistical tests for additive interactions.[24,293] Significant additive interactions were observed in both studies where the highest overall risks were seen in the unhealthiest lifestyle-lowest SES group. However, Zhang et al. performed analyses of two cohorts in their study and only observed the significant additive interaction in analysis of UK Biobank and not in analysis of NHANES.[293]

These additive interactions indicate an interaction on the additive scale or 'super-additivity' whereby the combined effects of lower SES and a combination of unhealthy lifestyle factors is greater than the sum of their independent effects.[178,411,412]

Comprehensive assessment of additive interactions are rarely reported by studies that examine interactions despite there being a public health argument that positive interactions on an additive scale highlight how the public health relevance of one exposure is dependent on the other interacting exposure.[344] For example, in this case, an additive interaction indicates that the number of adverse health outcomes due to multiple unhealthy lifestyle factors depends on the number of people with unhealthy lifestyle factors who are also from lower SES backgrounds (or that the number of adverse health outcomes due to lower SES depends on the number of individuals from lower SES backgrounds who also have multiple unhealthy lifestyle factors).[413] The implication is that where intervention resources are limited, these additive interactions highlight potential at-risk groups in whom most benefit will be felt from reduction in one

of the interacting exposures (multiple unhealthy lifestyle factors and lower SES).[178,411]

More consistent with additivity, rather than departure from additivity or super-additive effects, there was more evidence of higher overall risks associated with the combined effect of both lower SES and unhealthy lifestyle factors (see Tables S9 and S10, Appendix 7 for results of combined analyses for studies included in the systematic review). Compared with those with no unhealthy lifestyle factors who were also in the highest SES group, the greatest mortality risks were more consistently seen in those with the unhealthiest lifestyle who were also in the lowest SES group.

The systematic review suggests, therefore, that there is currently a small amount of evidence for an SES-based differential vulnerability to a combination of unhealthy lifestyle factors. Together with evidence for an independent effect of deprivation on top of an effect of multiple unhealthy lifestyle factors this strengthens arguments to target both lifestyle and socioeconomic factors to improve population health. Moreover, the results from the systematic review and weighted lifestyle score analysis indicate that addressing lifestyle factors (behavioural determinants of health) alone will not fully diminish socioeconomic gradients in adverse health outcomes.

8.3.2 Differential vulnerability – integrating qualitative findings and implications for interventions and policy

When the epidemiological risks from the combined and potentially interacting exposures of multiple unhealthy lifestyle factors and lower SES are considered in conjunction with the views of key stakeholders seen in the qualitative work package in this thesis, the need for interventions to consider and or address both exposures simultaneously is further highlighted. Participants' views of the manifold paths of influence from difficult social and economic circumstances to all types of unhealthy lifestyle factors, implies that interventions or policies that target unhealthy lifestyle factors with little or no consideration for the socioeconomic factors, which limit or facilitate people's choices, will be less successful. For example, perceptions of greater empathy (i.e., acknowledging the context or underlying reasons for unhealthy lifestyle factors) have been

associated with greater patient 'enablement' and improved health outcomes and may be particularly important for health promotion interactions with those living with socioeconomic deprivation.[414-416]

However, there is also a risk that explicitly targeting resources and interventions towards those experiencing deprivation could increase perceptions of stigma. For example, there may be risks inherent in giving 'lifestyle advice', particularly when the social determinants of lifestyle are acknowledged directly as it may deter those who fear blame or feel shame for their illness from seeking help.[44]

A meta-ethnographic review of 17 qualitative studies that examined UK residents' (mainly from lower SES backgrounds) views around links between socioeconomic circumstances and health or health inequalities identified perceptions that were similar to those captured in the study in Chapter 7.[417] Reduced access to facilities (e.g., play parks, and poor transport links) were linked to poor diets and low levels of physical activity. However, the authors also identified how difficult socioeconomic circumstances (e.g., poor housing) were a source of shame and perceived stigma that impacted behaviour (e.g., unwilling to host friends and family). The authors suggest that the sense of shame due to being from a deprived area was a reason why participants rarely acknowledged the existence of socioeconomic health inequalities directly even though they readily described socioeconomic influences on health and lifestyle. The authors conclude that highlighting socioeconomic health inequalities could itself increase peoples' feeling of injustice and stigmatisation and thereby worsen a psychosocial factor thought to exacerbate health inequalities.

8.3.3 Social connection as an unhealthy lifestyle factor in the context of deprivation

The social connection analyses shows that different measures of social connection are each important for health and that their combinations could be even more impactful. With the reasons given in Chapter 2 for why social connection should be considered as a lifestyle factor or alongside other lifestyle factors, the analysis informed which components of social connection could be included in a lifestyle score. There are many reasons for including a measure of social connection in a lifestyle score that aims to support more deprived

communities. For example, the inter-relatedness of social connection and unhealthy lifestyle factors are likely to be even stronger in areas of deprivation, where unhealthy lifestyle factors cluster, which means that improving social connectedness could support healthy lifestyle factors in areas of deprivation.[17,323,418-420]

Further, a review of literature examining relationships between a range of sociodemographic characteristics and functional components of social connection provides evidence for the association between lower SES (unemployment status and worse financial situations) and loneliness.[91] While research examining SES inequalities in social isolation is lacking, there is a general trend that those of lower SES are more likely to be isolated.[421]

There is direct evidence too of the inter-relationship between social connection, unhealthy lifestyle factors, and deprivation. For example, in a cross sectional survey, 5113 residents from more deprived areas of Denmark had higher odds of reporting loneliness compared with individuals from a nationally representative sample dataset.[422] And in the same study, both social isolation and loneliness were independently associated with higher odds of reporting unhealthy lifestyle factors (smoking, low fruit and vegetable intake, physical inactivity) and even more strongly associated with having multiple unhealthy lifestyle factors. However, compared to those from the highest SES group who were not socially isolated/lonely, those both from the lowest SES group and who were also socially isolated/lonely had the highest odds of having single and multiple unhealthy lifestyle factors. A cross-sectional study with data from 4,814 middle to older aged adults in Germany also found a heighted chance of having unhealthy behaviours with the combination of both lower levels of social connection and higher levels of deprivation than either one alone.[423]

It stands to reason therefore, that addressing social connection alongside unhealthy lifestyle factors with targeted support to more deprived areas could have additional benefits and therefore represent opportunities for intervention and policy development. For example, supporting physical activity groups in more deprived communities may have a dual benefit of increasing social connection and simultaneously increasing physical activity. Indeed, a recent systematic review of studies that evaluated interventions for either loneliness or

social isolation found that social aspects of interventions that target physical activity (e.g., group exercise) could alleviate loneliness.[424]

The qualitative findings in this thesis also support this argument. For example, the analysis of stakeholders' views highlights how and why unhealthy lifestyle factors are part of complex and interacting social practices implicating the critical nature of social connections and human interactions.[17,120,122]

However, problems that have higher prevalence in more deprived areas, such as childhood maltreatment or multimorbidity (i.e. having more than one long-term health condition), are also associated with both social disconnection and deprivation, which highlight the complexity and difficulty of addressing social connection in more deprived areas.[425-427] Moreover, by definition, resources that could support healthy levels of both social connection and other lifestyle factors are scarcer in more deprived areas (e.g., access to green spaces), which heightens the challenge.[428]

Nevertheless, the potential benefits of integrating attempts to improve social connection with those that address other unhealthy lifestyle factors are clear. The inclusion of measures of social connection measures in a lifestyle score are part of wider attempts to achieve that and to raise awareness of the health risks associated with a lack of social connection.[19,422]

8.3.4 Challenges and opportunities in addressing multiple unhealthy lifestyle factors

The qualitative work also supports a wider recognition of the dynamic landscape of lifestyle factor epidemiology.[7,10] The increasing numbers of lifestyle factors considered important to health and their interplay potentially represents a challenge that could add to perceptions of confusion and contradictory health messages around lifestyle factors.[429-431] However, the increasing number of lifestyle factors under consideration also represent new or additional factors for individuals, communities, and policy makers to consider and therefore address. Acknowledging and integrating understanding of the complex interplay of lifestyle factors into interventions and policy could facilitate improvements in

lifestyle factors where lifestyle factor interplay could lead to dual or synergistic benefits (e.g., physical activity and social connection).[424]

Further, as stated in Chapter 6, a weighted lifestyle score that considers multiple lifestyle factors provides a tool which can assist individuals and support health professionals to prioritise which factor(s) to address. Thus, a weighted lifestyle score could help resolve some of the challenges that could be associated with the evolving complexity of lifestyle factors.

8.3.5 Multiple unhealthy lifestyle factors in the context of socioeconomic deprivation

Attempting to address multiple unhealthy lifestyle factors in the context of socioeconomic deprivation is itself both a challenge and opportunity to support healthy living. Firstly, the socioeconomic gradient of unhealthy lifestyle factors means that focussing on the context of deprivation is a way of targeting lifestyle resources and tailoring lifestyle interventions to where they are most needed.[15,161] Secondly, considering multiple lifestyle factors together acknowledges the idea of the complex interplay between lifestyle factors developed in this thesis and begins to integrate the understanding of lifestyle factors as interacting social practices. Thirdly, the social determinants of unhealthy lifestyle factors and viewing poverty as a fundamental determinant manifest in stakeholder views in this thesis, suggest that in order to be effective, policies and interventions for lifestyle factors as social practices must fully integrate the understanding of the influence of poverty and social context.[2]

As stated previously, much prior effort to address unhealthy lifestyle factors have overlooked wider social contexts. For example, systematic reviews of lifestyle factor interventions designed to reduce smoking, increase physical activity, or increase healthy eating in low-income groups highlight the individual-level nature of many existing interventions and the focus on behavioural change techniques.[26,28] A related systematic review and meta-analysis of interventions targeting the same lifestyle factors in low-income groups focussed on identifying which components of the intervention (behavioural change techniques) and ‘context’ or delivery of intervention were associated with effectiveness.[27] The authors identified that the specific

techniques (e.g., encouraging self-monitoring, face-to-face contacts with facilitators) associated with effectiveness were different for different lifestyle factors. However, here 'context' relates to details of the intervention and nature of intervention delivery (i.e., who provided the intervention and how, where, and when was it provided). Whereas the wider social context of the participants or intervention settings is largely missing from evaluations of these interventions. Interestingly, this meta-analysis, while not focussing on multiple lifestyle factor interventions, found there was a suggestion that interventions targeting more than one lifestyle factor were associated with greater effectiveness when the overall aim was promoting weight loss but not when the aim was increasing physical activity.[27] This again highlights the interplay between lifestyle factors and suggests that improvements in certain unhealthy lifestyle factors may be contingent on the aims of the intervention (e.g., an overall aim of making lifestyle change versus improving general wellbeing).

8.3.6 Agency, structure, and the social determinants of lifestyle

The qualitative WP in this thesis raised important issues around agency versus structure in relation to lifestyle factors. The discussion in Chapter 7 describes a tension or dissonance among professionals who want to support individuals and populations make healthy lifestyle change and lead healthier lives whilst also appreciating the limits to people's agency imposed by the social determinants of lifestyle. This reemphasises the need to address social and contextual factors that limit healthy choices. As stated by Michael Marmot, 'if the major determinants of health [lifestyle] are social, so must be the remedies.'[432]

Views of participants in WP3, where unhealthy lifestyle factors were seen as natural responses to or coping mechanisms for difficult circumstances, also fit with an alternative explanation for the social patterning of unhealthy lifestyle factors where opportunities to make healthy lifestyle choices are not so much considered limited but viewed as rational or optimal from a given perspective and context.[278,417,433-435]

However, whether an individuals' choices are limited by social contexts or are rational given their context, both indicate that there should be greater emphasis on altering the wider context to facilitate healthy living.[194] Indeed, structural

interventions for structural problems are increasingly being implemented (e.g., governmental and institutional mechanisms to address commercial determinants of health) and their success would facilitate individual and community-level interventions that support healthy living.[436] Additionally, framing structural interventions as ways of increasing agency for healthy living could capitalise on the perceptions of agency for lifestyle choices seen in this thesis and increase support for wider structural change.[437]

8.3.7 Integrating social determinants of lifestyle into clinical practice

This thesis reinforces arguments to increase the focus on the links between context and lifestyle and to deepen the integration of our understanding of social determinants of lifestyle when trying to support healthy living. This is especially the case for populations experiencing socioeconomic deprivation.

From a clinician's perspective, there is growing appreciation of the need to integrate SDH into clinical tools and guidelines. For example, the additional risks associated with deprivation have been incorporated into CVD risk scores (ASSIGN, QRISK) that lower treatment thresholds for CVD medication.[367,368] However, there is little consideration on how these principles could be applied to lifestyle interventions or guidance.

The 2019 American College of Cardiology/American Heart Association guideline on the primary prevention of CVD recommends that clinicians evaluate the SDH to inform management.[438] Here, the incorporation of SDH in recommendations for clinicians focusses on considering how wider social contexts could influence individuals' ability to adhere to healthy lifestyle factors (e.g., reduced potential to adhere to a 'heart-healthy diet' by those from more socioeconomically disadvantaged backgrounds). Related US-based recommendations on how clinicians and health care systems should address SDH again mainly focus on screening for and awareness of SDH and subsequent tailoring of management advice (e.g., simplifying medications for those with lower health literacy).[439] While tailoring expectations for adherence to lifestyle or medication advice could be important for effective clinician-patient relationships, it is, by itself, unlikely to lead to either healthier living or a reduction in inequalities.

Existing single lifestyle factor guidelines, such as the UK Chief Medical Officers' Physical Activity Guidelines, only briefly acknowledge awareness of 'health inequalities in relation to physical inactivity'.[338] Equivalent US guidelines explicitly focus on 'selected aspects of health-related quality of life' and do not include 'other aspects of quality of life, such as those related to finances, relationships, or occupations.'[303] However, the risks of differential vulnerability to lifestyle and limits posed by social determinants of lifestyle identified in this thesis suggest that lifestyle factor guidelines could do more to directly acknowledge and address lifestyle inequalities. Also, failing to address the interrelationship of single lifestyle factors with others, including social connection, and interplay with the wider social context more explicitly, existing lifestyle guidelines could contribute to 'lifestyle drift' by focussing explicitly on behaviours rather than social context.

For example, there are currently no lifestyle recommendations or interventions that provide increased and proportionate lifestyle support for more socioeconomically disadvantaged groups, nor do recommendations alter thresholds for lifestyle interventions based on the higher risks experienced by those in lower SES populations.[378] However, similar to CVD risk scores that guide CVD medication management, findings from the weighted lifestyle score analysis in this thesis indicate a way of integrating the additional risk associated with deprivation with the risks of unhealthy lifestyle factors to inform lifestyle support.[367,368]

Further, existing lifestyle recommendations and guidance contain no mentions of support or advocacy for population-level interventions that can influence the ability of individuals and communities to adhere to healthy living guidance (e.g., affordability, accessibility, and availability of health harming products like tobacco or energy dense and nutrient poor foods). Lifestyle factor guidance for individuals and supporting clinicians could better incorporate our understanding around the social determinants of lifestyle by promoting examples of policies and interventions that facilitate healthy lifestyle change in more deprived communities (e.g., minimum unit pricing of alcohol or levies on sugar sweetened beverages).[376,377,440,441]

8.3.8 Integrating lifestyle into measures of SES

Views of participants in WP3 of the inextricable links between social circumstances and unhealthy lifestyle factors resonate with some of the original and underlying theories and development of SES measures where lifestyle factors were integral to the social structuring and ranking of societies.[36,116,120,144] And, as alluded to in the introduction, lifestyle factors have long been considered part of a definition of deprivation and as a key social determinant of health.[119,146,442] Therefore, conceptualising multiple unhealthy lifestyle factors as a part of deprivation rather than as explanatory mediators of health inequalities, would be another way of 'bringing structure back' into the forefront of lifestyle related policymaking and intervention design.[120] This could be part of an effort to reappropriate the term 'lifestyle' to include fuller consideration of opportunities as well as choices as it was perhaps originally intended by Weber.[36,37,124]

Considering multiple unhealthy lifestyle factors as part of socioeconomic deprivation could help shift the focus from individual responsibility for lifestyle towards the social and contextual drivers of unhealthy lifestyle factors. This shift in perspective could help stem the 'lifestyle drift' of policy and interventions because addressing unhealthy lifestyle factors would then mean addressing socioeconomic deprivation.[384]

Others have recognised that differences in adverse health outcomes observed in those that share similar levels of SES could indicate there may be important aspects missing from existing measures of SES.[443] This implies there are other experiential aspects of poverty and deprivation that are not currently captured in existing measures of SES (e.g., stigma). Findings from this thesis around reduced agency due to structural limits for healthy living suggest that multiple unhealthy lifestyle factors could be considered part of the experience of poverty and therefore could be included in new measures of deprivation. For while SIMD includes a measure of alcohol or drug-related hospital admissions, the common indices of deprivation described in Chapter 2 (Townsend, Jarman, Carstairs, EIMD, and SIMD) do not include direct measures of lifestyle factors in their components or domains.[151-154,341]

8.4 Strengths and limitations of this thesis

The strengths and limitations pertaining to each work package of this thesis are discussed in the relevant results chapters (4 to 7). This section presents a reflection on overall strengths and limitations of the wider thesis.

Searches for the systematic review (Chapter 4) were originally conducted from database inception to March 2020 but were updated after submission for publication to December 2021. Articles published since then would have been missed. Therefore, an updated search of PubMed was conducted in September 2023 and a forward citation search of the most highly cited included study (Zhang et al. 2021)[293] was conducted in October 2024 to scope the extent of newer literature. These searches identified more recent eligible articles that examined additional outcomes (e.g., site specific cancers)[311] and were based in different settings (e.g., Korea,[311] and China [312,444]). While some heterogeneity exists in the detail of these newer papers, their results still broadly suggest that the detrimental effect of socioeconomic deprivation on health is in addition to that of a combination of unhealthy lifestyle factors. Updating the systematic searches in full would likely reveal further studies which would influence the synthesis of evidence for RQ1.

After identifying the importance of the independent effects of both functional and structural components of social connection in Chapter 5, only two structural components were included in the weighted lifestyle score analyses. This does overlook the importance of subjective elements of social connection. However, the definition of lifestyle factors used in this thesis, while necessary to delineate the object and target of interest in the research, made it difficult to justify including subjective measures of social connection as lifestyle factors. Similarly, measures of the quality of social connections are lacking in UK Biobank and therefore from this thesis. However, future work, both in terms of development of a lifestyle score to be used clinically and in terms of new quantitative analyses, incorporating both functional and quality components of social connection into measures and tools for healthy living could be fruitful.

8.4.1 UK Biobank analyses

The use of UK Biobank as a resource for the quantitative analyses represents both a strength and limitation. As discussed in Chapter 3, there are advantages and disadvantages to UK Biobank. However, conducting similar quantitative analyses in different datasets is required to increase confidence in the generalisability of these results. As well as additional datasets, examination of different demographic groups would be important to extend the type of research and analyses in this thesis. For example, it would be important to examine the potential combined impact of lifestyle and SES in datasets with improved representation of minority ethnic groups and datasets with participants from younger age groups.

Another example where UK Biobank data resource appears to be both an advantage and disadvantage is the rich variety of baseline exposure and health related outcome variables available. There are numerous additional variables available within UK Biobank that could have been examined as exposure variables, such as objectively measured physical activity (accelerometer data), different alcohol measures, measures of relationship quality, more detailed dietary assessments, and repeated baseline assessments. Each variable has advantages and disadvantages and therefore, bar examining all of them, which would be unwieldy, they represent both opportunities for additional examination at the same time as somewhat arbitrary choices. The main advantage of the baseline variables used in this thesis is that they are available for nearly all UK Biobank participants. Whereas subsamples with more detailed data are only available for fewer participants. For example, repeat baseline assessment is only available for around 20,000 participants. And, as with the other subsamples that had additional and more detailed assessments (e.g., online follow-up dietary questionnaire or accelerometer data) participants with additional or repeat assessments are even more likely to be of White ethnicity, older, of higher SES, and therefore represent an even more highly selected population than UK Biobank as a whole, which already has a significant healthy volunteer bias.[247,445]

Therefore, while there may have been enough participants within subsamples (e.g., 20,000 for repeat assessment or 231,287 for happiness and satisfaction variables) to sustain statistical power across the large number of interaction

analyses in this thesis, there is also a greater risk of selection and collider bias. I chose to have higher numbers of participants and lower risk of selection bias.

With additional time and resources, additional analyses could have been performed that further utilise the rich UK Biobank resource. For example, examining available hospital admission data could provide results on the strength of association between social connection variables or combinations of lifestyle factors and additional outcomes, such as CVD and cancer incidence. Therefore, future work could test whether similar results are observed for numerous different exposure and outcome variables that are available in UK Biobank.

The lack of representativeness of the UK Biobank cohort has been highlighted in both Chapters 5 and 6 but is worth discussing again here. As with many research cohorts, it is well recognised that UK Biobank participants are, on average, more affluent, more likely to be from a White ethnic background, and have fewer recognised disease risk factors than the general population.[247] However, the UK Biobank has an atypically low response rate (5%) compared with other research cohorts (response rates ~60%) whilst also being uncommonly large and rich in data.[446] This raises questions about the validity or generalisability of UK Biobank findings.

To discuss this issue further here, I use an example of CVD deaths in both quantitative analyses. However, this discussion could apply to many other outcome or baseline variables available in UK Biobank or its linked registry data.

In the social connection analysis (Chapter 5) there were 33,135 (7.2%) deaths, of which 5,112 (1.1%) were CVD deaths. Similarly, in the weighted lifestyle score analysis (Chapter 6) there were 30,687 (6.6%) deaths, of which 4,632 (1.0%) were CVD deaths. These results are similar to those reported in the UK Biobank Death Summary Report (March 2024) because, in each case, the percentage of all deaths due to CVD is approximately 15%.[447]

Office for National Statsitics (ONS) data for 2021 (the final year of follow up in the UK Biobank analyses in this thesis) shows that, in England and Wales, 15% of all deaths in all age groups were due to ischaemic heart disease (ICD-10 codes

I20 to I25) and cerebrovascular diseases (ICD-10 codes I60 to I69).[448] These causes likely represent most but not all deaths defined as CVD deaths in analyses in Chapters 5 and 6, which included more diseases of the circulatory system (i.e., ICD-10 codes I05 to I99, Z86.7, G45, and G46). However, these ONS data relate to only one year rather than 12 or so years of follow up in the UK Biobank analyses.

Making a fairer comparison using the ONS downloadable dataset for 2021 by limiting the data to 45–74-year-olds (an age group more similar to most UK Biobank participants) and expanding the definition of CVD deaths to all ICD-10 codes I00 to I99 (all diseases of the circulatory system), ONS data show that 38,456 (22%) of all 173,092 deaths in this age group were CVD deaths.[449] Therefore, compared with all 45–74 year olds in England and Wales in 2021, a population more representative of the UK general population but of similar age to those in UK Biobank, UK biobank participants appear to have a lower percentage of deaths due to CVD (15%).

Comparing the 2021 ONS report to the 2024 UK Biobank death summary report highlights other differences in the commonest causes of death.[447,448] For example, the ONS report shows that 10% of all deaths (total population) are due to dementia while only 5% of UK Biobank deaths are dementia deaths. However, there are some figures that are more similar to UK Biobank in the 2021 ONS downloadable data: 13-16% of deaths among 50-64 year-olds were due to CVD (ICD-10 I00 to I99); and equivalent percentages for 65-79 year-olds and those aged 80 years or more were 11-16% and 15%, respectively.[449]

The selective nature of UK Biobank (e.g., participants who are younger, more educated, and more predominantly of a White ethnic background than the UK general population) very likely explains the discrepancies observed in the commonest causes of death and the percentage of all deaths that are CVD. This has implications for the observed associations within UK Biobank and remains a topic of epidemiological debate. Some authors argue that the selective nature of UK Biobank and healthy volunteer bias pose a risk of collider bias, which can lead to spurious associations being observed that do not exist among the general population.[248,450] However, others assert that both a sufficient range within variables (e.g., for Townend index, enough participants from affluent, deprived,

and mid-range backgrounds) and a sufficiently large number of participants mean that the strength of associations observed between exposures and outcomes are likely to be generalisable to the wider population even if the prevalence of exposures and outcomes are not.[33,245,247,451]

Batty et al. have tested this assertion by comparing UK Biobank with pooled data from 18 more representative cohorts (health surveys in England and Scotland) and examined the associations between common risk factors and cause-specific deaths.[446] They found, compared with the pooled survey data, CVD mortality rates (as well as those from all cancers, and tobacco and obesity linked cancers) were markedly lower in UK Biobank while the rate of suicide was higher. However, they also found that the direction and strength of association between well recognised risk factors (e.g., age, sex, smoking, physical inactivity, and alcohol consumption) and CVD mortality were similar in both UK Biobank and the pooled survey data.

While it is important to avoid reliance on evidence from a single dataset with which to change practice or policy, the UK Biobank remains a rich epidemiological resource with which to test hypotheses and contribute to evidence that can influence practice and policy. As others have stated, there can be utility in non-representative cohorts as exemplified by the Framingham study and the British Doctors' study which have had major influence on practice and policy.[9,452] However, as per any other single dataset or study, findings from UK Biobank must be interpreted with keen consideration of the limitations and wider evidence.

One limitation of the UK Biobank analyses presented in this thesis is a lack of testing of the proportional hazards assumption. A fundamental assumption of the Cox proportional hazard model is that the hazard for a study participant is a fixed proportion of the hazard for any other participant and that ratio remains constant over time.[453] Therefore, the hazard ratios presented in this thesis represent a weighted average of the true hazards over the study periods.[454] Testing for violations of the proportional hazards assumption could provide additional valuable information for time-varying adverse health effects of the key exposure variables in this thesis (i.e., a lack of social connection, combinations of unhealthy lifestyle factors, and deprivation). For example, the

impact of the feeling of loneliness on mortality may not be linear and may be greater the longer that feeling is present. And the effect of a combination of unhealthy lifestyle factors on CVD mortality could increase over time if protective compensatory mechanisms degrade over time. However, exploration of the time-varying effects of these key exposure variables would also best be done with data that includes many repeated assessments for exposures throughout follow-up, which was not available for UK Biobank participants.

The main results from the quantitative analyses of UK Biobank in this thesis rely on the outputs of Cox Proportional hazard models. The models provide estimates of the strength of association between exposures and outcomes (hazard ratios) as well as information on the uncertainty of those estimates (confidence intervals). However, p-values, which are also an output from Cox models, were not provided in this thesis. There is a long history of misinterpretation and application of p-values in epidemiology, with the particularly common problem of assigning arbitrary significance thresholds to p-values.[240,241,455] Nevertheless, while p-values and confidence intervals are closely related, reporting p-values in this thesis could have provided further information regarding hypothesis testing.

Whilst the covariates included in the Cox models in the quantitative analyses are important potential confounders of the examined associations, they are not the only potential confounders and their measurement (e.g., single baseline measure or self-report error) and nature (e.g., researcher derived categories) could have influenced results. For example, whilst the list of 43 long-term conditions is comprehensive and based on previous evidence, it takes no account of the severity of conditions or treatment received.[339,456] Further information that would permit model adjustment for disease severity and treatment would add rigour to analyses by taking account not only of the number of long-term conditions but also the severity or level of control of each condition. And, although BMI, which is strongly correlated with many dietary measures, was controlled for in the models, there may be specific dietary factors (e.g., ultra processed food) that could confound the association between exposures and outcomes. Therefore, controlling for specific dietary factors in future social connection analyses and examining alternative dietary measures as part of a lifestyle score would be important. Additionally, a list of long-term conditions is one of many ways to measure multimorbidity (and, therefore, baseline health in

observational studies) and future analyses could explore the effect of adjusting for other common measures of multimorbidity such as weighted indices e.g., the Charlson Comorbidity index.[457]

The sensitivity analyses (excluding those with baseline self-reported CVD or cancer or who died soon after recruitment) performed in the quantitative analyses were done to assess for reverse causality. The fact that results were similar in both the main and sensitivity analyses suggests that it was not poor baseline health that explained the observed associations (i.e., poor baseline health did not appear to lead to both exposures and outcomes). The list of conditions excluded in the sensitivity analyses, while by no means exhaustive, included CVD and cancer diagnoses which are among the most common causes of death in men and women in the UK.[448] Excluding those who died within 2 years of recruitment would, by association, also exclude some participants with poor baseline health due to mental health problems or musculoskeletal diseases (or any other diseases) that are likely to act as mediators between exposures and outcomes. However, it would be important for future studies to examine the impact of excluding participants with other baseline self-reported conditions that are likely to lie on the complex and multidirectional causal pathways (e.g., depression or arthritis) that link these exposures to adverse health outcomes. Indeed, future analyses that incrementally exclude different sets of diseases could explore the impact of excluding various baseline health conditions on the strength of association between social connection and mortality or between combinations of unhealthy lifestyle factors and mortality. Such future studies could shed light on the potential causality implicated in the exposure-outcome associations observed in this thesis.

The use of a single area-based measure of SES in the weighted lifestyle score analysis is a limitation. For example, components of Townsend index (e.g., overcrowding or car ownership) have been criticised as now being less relevant for many sections of society.[458] Therefore, examining the effect of different SES measures, including individual-level measures (e.g., income, education level), on the weighted lifestyle score-health outcome associations would be important as different effects by SES measure could have implications for causal inferences.

This thesis did not look at the link between lifestyle (including markers of social connection), deprivation, and other adverse health outcomes that are of growing importance in view of shifting demographic patterns and ageing populations. For example, both social disconnection and other unhealthy lifestyle risk factors are associated with dementia, diabetes, and frailty.[58,93,459,460]

8.4.2 Qualitative findings

The qualitative work did not examine perceptions of deprivation or examine the ways in which stakeholders conceptualise poverty. Rather, the topic guide steered discussion towards how participants felt financial circumstances or deprivation were linked to unhealthy lifestyle factors, how much control individuals have over lifestyle given their social circumstances, and how people from more disadvantaged backgrounds could be supported to lead healthy lives. Therefore, this means additional insights into why participants perceive the links they do between deprivation and lifestyle may have been missed. Finally, public participants were not purposively recruited from lower SES groups and therefore the views analysed here may not reflect those with experience of socioeconomic deprivation. However, participants were recruited with assistance from third sector organisations (The ALLIANCE and The poverty Alliance) with close links to members of the public who may have experience of challenging social and economic circumstances and the range of occupational types of recruited participants was broad.

There was a discrepancy in views by professional types: public and practitioner participants more readily discussed lifestyle factors at the level of individual action whereas PPHPs were keener to focus on wider societal influences on lifestyle. Nevertheless, all participants described the strong influence of socioeconomic circumstances on lifestyle. Indeed, this is what lead to a tension in practitioners' views; a desire to support individuals and instil hope for healthy change on one side and their understanding of the structural limits to lifestyle on the other. This tension highlighted a balancing act that patient-facing practitioners attempted to perform by accounting for socioeconomic conditions when supporting and advising individuals. There is a risk that this finding of a desire to instil hope for healthy lifestyle change could be interpreted as best practice without due consideration of the structural limits on lifestyle. As a

result, practitioners could be encouraged to promote healthy lifestyle change excessively, which could perpetuate feelings of inadequacy in individuals or victim blaming. However, it remains an empirical question whether, for example, practitioners should encourage hope for healthy lifestyle change in the face of challenging conditions of poverty. As there is also a risk of not intervening on a lack of hope, which, as highlighted in Chapter 7, was linked with unhealthy lifestyle factors. All of this motivates further research into if and how primary care and community workers should support individuals affected by socioeconomic deprivation and unhealthy lifestyle factors.

8.4.3 Integration of qualitative and quantitative findings

A major strength of this work is the integration of results from both qualitative and quantitative studies that examine complex aspects of the associations and interactions between behavioural and other social determinants of health. It highlights, in new ways, the importance of considering socioeconomic context with lifestyle factors and this will be of critical relevance to policymakers and health professionals.

This mixed methods thesis synthesises broad types of new evidence. Firstly, via a systematic review of the wider literature on differential vulnerability, which has not been done before, this thesis highlights that the wider evidence suggests an additive effect of deprivation on the association between a wide combinations of unhealthy lifestyle factors and adverse health. Secondly, it highlights the complexity and health importance of social connection by examining different components of social connection in a widely used research dataset (UK Biobank) in a new way. Thirdly, it uses the new quantitative findings on social connection to include measures of social connection in the development of a novel weighted lifestyle score. This, in turn, raises the profile and awareness of social connection by considering it a ‘lifestyle factor’ which is rarely done. Fourthly, it examines the mortality associated with the new lifestyle score and thereby adds to the evidence around differential vulnerability for an unhealthy lifestyle in more derived groups. Finally, views of key stakeholders on these issues highlight the inextricable links between poverty and wider socioeconomic circumstances in explaining unhealthy lifestyle factors and, therefore, the need for policy and interventions to focus on the social determinants of lifestyle.

Considering the qualitive findings around professionals' perceptions (evolving complexity of lifestyle factors, social determinants of lifestyle, and a desire to support individuals) with results from the weighted lifestyle score analysis suggests a tool such as a weighted lifestyle score that also accounts for deprivation could support professionals as they guide individuals living in more deprived contexts through a complex array of lifestyle factors. This has positive implications for future intervention development, implementation, and acceptability. Further, stakeholder perspectives on lifestyle factors in deprived contexts has informed the next steps in how to develop the lifestyle score as a clinical tool, namely that contextual factors (e.g., the individual's local area and resources, their financial wellbeing, their mental health, and their wider social support), need to be a core consideration of any related advice and support. Therefore, any response or action to using the score needs to embed the understanding of social determinants and structural drivers for each lifestyle factor.

While this thesis discusses the social determinants of lifestyle and cites political contexts as an example of a structural determinant of health, a discussion around political ideology as a structural determinant of lifestyle could have added another important area for contextualising the findings from this thesis. Political ideology (e.g., a right-left spectrum with conservatism at one end and liberalism or progressivism at the other) could influence policy responses to lifestyle factors. For example, conservative governments may favour policies that impose fewer infringements on individual liberties versus more left-wing governments. This has direct implications for the findings of this thesis which support a reduced emphasis of individual-level approaches for lifestyle policies whilst highlighting the social determinants of lifestyle. This would imply a greater likelihood of victim blaming with conservative-based lifestyle policies which rely on individual-level responsibility. However, a unidimensional left-right political spectrum is potentially oversimplistic with multidimensionality or spectra of political ideologies more likely representative of the political environment and policy responses.[461,462] Therefore, how political values, ideologies, and parties influence lifestyle policy and the wider social determinants of lifestyle warrants detailed discussion and investigation that is largely beyond the scope of this thesis.

8.5 Implications for practice

The implications for practice from the findings of each the results chapters are summarised in turn here.

Firstly, an SES differential vulnerability to multiple unhealthy lifestyle factors could encourage health workers and allied professionals that health benefits for addressing unhealthy lifestyle factors may be higher in more deprived groups. However, tailoring lifestyle advice and interventions to integrate knowledge of the social determinants of lifestyle are crucial.

Secondly, results from the social connection analysis suggest clinicians and community workers should consider the combined risks to individuals who are classified as isolated by both subjective and objective measures of social connection. And identifying those who live alone could be a helpful standardised marker of isolation and signifier of risk.[345,346]

Thirdly, a weighted lifestyle score could assist individuals and health professionals in addressing a wide range of lifestyle factors. Additionally, analysis here shows that the additional risk associated with deprivation could be integrated into health workers assessments of risk from multiple unhealthy lifestyle factors.

Fourthly, increasing lifestyle support and training for front line workers could assist them in dealing with the evolving complexity of lifestyle factors and therefore in supporting health living in individuals and communities.

8.6 Implications for policy

The main implication for policy from this thesis is the need to further deepen the integration of our understanding of the social determinants for unhealthy lifestyle factors into relevant policy and guidelines. Existing policies have focussed on single lifestyle factors, on the individual, and not on social context, and have overseen a widening of health and lifestyle inequalities in the UK.[29,170] The combined and potentially synergistic risks associated with multiple unhealthy lifestyle factors and socioeconomic deprivation suggest that policies

concerned with lifestyle factors should focus on supporting healthy living in more socioeconomically deprived populations.

Further, the identification of poverty as a fundamental determinant of unhealthy lifestyle factors along with its well-recognised direct impact on health means more is required at policy-level to address poverty. The entanglement of life circumstances and lifestyle factors highlighted in this thesis provides further evidence for the need to alleviate poverty directly in order to improve health outcomes, especially those health outcomes that are strongly associated with lifestyle factors.

Policy concerned with social connection or social isolation and loneliness should address numerous facets or types of these problems to identify those most at risk and reduce the associated adverse health. Further, this thesis highlights how and why social connection could be considered alongside other lifestyle factors.

Finally, Chapter 7 highlights poverty and the social determinants of lifestyle and supports the argument for reduced individual-responsibility for unhealthy lifestyle factors. This has potential implications for the word 'lifestyle' itself and perceptions of what 'lifestyle' means in policy and decision making. Using the word 'lifestyle' throughout this thesis while discussing the additional risks and influences of wider socioeconomic and structural forces beyond individual-level control promotes an understanding of the word that acknowledges limits to agency and the role of structure. As stated before, this more nuanced conceptualisation of 'lifestyle' is potentially as originally intended by Weber to whom the word and its meaning is often attributed in sociology.[37] However, with its current understanding as being a suite of 'modifiable' risk factors it could be that the word lifestyle will continue to be understood simply in terms of individual-level choice and therefore perpetuate 'lifestyle-drift' and risk victim blaming. Other terms such as 'risk factors' or 'health behaviours' appear to have the advantage of being less clearly associated with individual-level choice. And other terms offer other advantages of promoting positive benefits such as 'healthy living' and 'wellbeing'. Nevertheless, the word 'lifestyle' remains widely used in medicine and epidemiology and promoting a more nuanced and accurate understanding of the word and its influences could reduce the risks of victim blaming and 'lifestyle drift'. Interestingly, 'lifestyle' was used throughout

discussions with the qualitative research participants who clearly felt there were socioeconomic limits to individual choice for healthy ways of living. This suggests that based on the understanding of the word ‘lifestyle’ among these key stakeholders, which included members of the public, the risks associated with using it could be overstated.

8.7 Suggestions for future research

Suggestions for future research is noted in each of the results chapters and are summarised again here.

The evidence base for differential vulnerability to multiple unhealthy lifestyle factors across SES gradients could be strengthened not only with examination for differential vulnerability in a variety of datasets and population types but also by examining a wider range of adverse health outcomes.[463] Further, measures of absolute difference in risk or additive hazard models would provide evidence for the scale of adverse health outcomes due to differential vulnerability.[464] More evidence for differential vulnerability would motivate further research into potential explanatory mechanisms.

Definitions of unhealthy lifestyles could be refined by exploration of the non-linear associations and interactions for a wide range of lifestyle factors to define unhealthy lifestyle across the SES spectrum. This could provide new targets for interventions and policies that aim to address unhealthy lifestyle factors in more deprived populations.

From the social connection analyses, the interactions between living alone and both friends and family visits and weekly group activity suggest that future research should explore whether living alone could be developed as a standardised measure for examining social connection. Similarly, future work examining interactions between components of social connection could inform interventions for those who have multiple markers of social disconnection.

The weighted lifestyle score created in this thesis could be further refined and validated through similar analyses in additional datasets that could provide evidence for generalisability and further motivate its use as a clinical tool.

Examining the impact of various measures of SES on the association between weighted lifestyle score and adverse health outcomes could be informative. More research examining modes of delivery of lifestyle scores that address multiple lifestyle factors could inform how to improve effectiveness e.g., addressing lifestyle factors in sequence or simultaneously.

Finally, future research could examine how practitioners balance a desire to support healthy lifestyle change in socioeconomically deprived contexts given their appreciation of the social determinants of lifestyle. Research like this would inform practice, for example, on how to adjust expectations and set goals around healthy living.

8.8 Conclusion

Adverse health outcomes associated with multiple unhealthy lifestyle factors may vary across the socioeconomic spectrum due to the differential effects of lifestyle factors. This differential vulnerability, together with evidence for the steep socioeconomic gradient in lifestyle factors (differential exposure), could help to explain some of the gradients in SES-based health inequalities. Therefore, maximal population health benefits could arise from targeting resources and support to improve multiple unhealthy lifestyle factors in more deprived areas.

A weighted lifestyle score comprised of multiple traditional and emerging lifestyle factors and that accounts for the additional risks associated with socioeconomic deprivation could be developed as a clinical tool to support individuals and professionals around healthy lifestyle change.

However, lifestyle factors and socioeconomic circumstances are deeply intertwined and therefore healthy living interventions and policies cannot afford to overlook wider social contexts. Combining lifestyle factor epidemiology with a sociological understanding of lifestyle factors could lead to more effective interventions. Fundamental integration of the understanding for contextual drivers for healthy living in all levels of policy and interventions may still be lacking. Redefining deprivation to incorporate unhealthy lifestyle factors more explicitly in current SES measures could be one way to redress that.

Appendix 1: Manuscript for systematic review protocol

Open access Protocol

BMJ Open

Understanding the influence of socioeconomic status on the association between combinations of lifestyle factors and adverse health outcomes: a systematic review protocol

Hamish Foster ,[1] Peter Polz ,[1] Frances Mair,[1] Jason Gill,[2] Catherine A O'Donnell [1]

To cite: Foster H, Polz P, Mair F, *et al*. Understanding the influence of socioeconomic status on the association between combinations of lifestyle factors and adverse health outcomes: a systematic review protocol. *BMJ Open* 2021;**11**:e042212. doi:10.1136/bmjopen-2020-042212

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-042212).

Received 28 June 2020
Accepted 04 April 2021

[1]General Practice and Primary Care, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK
[2]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK

Correspondence to
Professor Catherine A O'Donnell; Kate.O'Donnell@glasgow.ac.uk

ABSTRACT

Introduction Combinations of unhealthy lifestyle factors are strongly associated with mortality, cardiovascular disease (CVD) and cancer. It is unclear how socioeconomic status (SES) affects those associations. Lower SES groups may be disproportionately vulnerable to the effects of unhealthy lifestyle factors compared with higher SES groups via interactions with other factors associated with low SES (eg, stress) or via accelerated biological ageing. This systematic review aims to synthesise studies that examine how SES moderates the association between lifestyle factor combinations and adverse health outcomes. Greater understanding of how lifestyle risk varies across socioeconomic spectra could reduce adverse health by (1) identifying novel high-risk groups or targets for future interventions and (2) informing research, policy and interventions that aim to support healthy lifestyles in socioeconomically deprived communities.

Methods and analysis Three databases will be searched (PubMed, EMBASE, CINAHL) from inception to March 2020. Reference lists, citations and grey literature will also be searched. Inclusion criteria are: (1) prospective cohort studies; (2) investigations of two key exposures: (a) lifestyle factor combinations of at least three lifestyle factors (eg, smoking, physical activity and diet) and (b) SES (eg, income, education or poverty index); (3) an assessment of the impact of SES on the association between combinations of unhealthy lifestyle factors and health outcomes; (4) at least one outcome from—mortality (all cause, CVD and cancer), CVD or cancer incidence. Two independent reviewers will screen titles, abstracts and full texts of included studies. Data extraction will focus on cohort characteristics, exposures, direction and magnitude of SES effects, methods and quality (via Newcastle-Ottawa Scale). If appropriate, a meta-analysis, pooling the effects of SES, will be performed. Alternatively, a synthesis without meta-analysis will be conducted.

Ethics and dissemination Ethical approval is not required. Results will be disseminated via peer-reviewed publication, professional networks, social media and conference presentations.

PROSPERO registration number CRD42020172588.

Strengths and limitations of this study

- ► This review protocol lays out a comprehensive search strategy and a rigorous data extraction and synthesis plan to collate the evidence for the effect of socioeconomic factors on the association between combinations of unhealthy lifestyle factors and adverse health outcomes.
- ► The inclusive nature of the eligibility criteria, which is necessary as there are likely to be few studies in this area, means included studies may be heterogeneous in design and methodology and this may preclude meta-analysis.
- ► The wide range of possible socioeconomic indicators and combinations of lifestyle factors likely to be included due to the inclusion criteria may make firm conclusions difficult.
- ► However, the wide range of possible socioeconomic indicators and combinations of lifestyle factors likely to be included due to the inclusion criteria will permit a comprehensive overview of both sets of exposures and, therefore, highlight evidence gaps.
- ► Synthesising a broad evidence base to provide an overview of the potential influence of socioeconomic status (SES) on associations between combinations of unhealthy lifestyle factors and adverse health outcomes could indicate which combinations of unhealthy lifestyle factors are associated with the greatest risks for lower SES groups.

INTRODUCTION

Background

Globally, unhealthy lifestyle factors (eg, smoking, physical inactivity) are among the main risk factors for mortality and non-communicable diseases (NCDs).[1] Socioeconomically deprived populations have the highest mortality and morbidity rates from NCDs but this is only partially explained by higher prevalence of unhealthy lifestyle factors.[2–4] Deeper explanations for

lifestyle-related health inequalities include both the synergistic interactions between individual lifestyle factors themselves and interactions between lifestyle factors and socioeconomic status (SES).[5 6] However, to date, there has been limited examination of how the association between combinations of unhealthy lifestyle factors and adverse health outcomes is moderated by SES.

It is critical to note that the word 'lifestyle' implies choice and control over health behaviours. However, lower SES (more deprived, lower income or less educated) groups are less likely to have as much control over 'lifestyle' or health behaviours than higher SES groups. Further, 'choosing' unhealthy health behaviours may be entirely rational given specific socioeconomic contexts.[7] However, the word lifestyle is employed here as this is widely understood in the context of potentially modifiable health behaviours.

Combinations of lifestyle factors

While single lifestyle factors are clearly associated with adverse health outcomes, meta-analyses provide evidence for how combinations of lifestyle factors have stronger associations with mortality and NCDs.[5 8 9] The evidence for the health impact of single lifestyle factors now also includes 'new' or emerging lifestyle factors, such as sleep duration,[10] television viewing time[11] and social participation levels.[12] When 'new' lifestyle factors are combined with 'conventional' factors (eg, smoking, physical inactivity, high alcohol intake or poor diet) associations with mortality are strengthened further.[12] Moreover, combinations of lifestyle factors can have a multiplicative or synergistic effect on adverse health outcomes. For example, the mortality associated with smoking and high alcohol intake together is more than the sum of the risks for each lifestyle factor alone.[13] Therefore, investigating the impact of broad combinations of lifestyle factors is necessary for comprehensive understanding of lifestyle-associated harm. Particularly so when the prevalence of three or more lifestyle risk factors is high.[14] For example, 55% of the Scottish population has three or more risk factors.[15] Furthermore, the additional risks associated with combinations of unhealthy lifestyle factors would motivate work to determine which combinations have the highest risk. For example, if a combination of high sedentary time together with short sleep duration and poor diet is highlighted as particularly high risk then interventions could be targeted at this specific behavioural combination.

SES and lifestyle

There is a significant body of research that focusses on investigating the extent to which the greater prevalence and nature of unhealthy lifestyle factors in lower SES populations can explain the well-known socioeconomic gradient in adverse health—so-called 'differential exposure'.[3 4 16–21] These studies estimate that 30%–50% of socioeconomic inequalities in all-cause and cause-specific mortality are attributable to the differential exposure to unhealthy lifestyle factors. Typically, these studies examine conventional lifestyle factors only, although often alongside metabolic factors such as blood pressure or body mass index (BMI).

However, despite being independently associated with mortality and NCD at levels commensurate with those of unhealthy lifestyle factors, socioeconomic factors are often omitted from lifestyle policy.[22] Furthermore, many studies appear to lack an assessment of the interaction between unhealthy lifestyle factors and SES. There is some evidence for interactions between single lifestyle factors and SES, whereby, for the same level of exposure, lifestyle factors have different effects across socioeconomic spectra—that is, 'differential vulnerability'.[21] For example, in a Scottish cohort, lower (as opposed to higher) SES (measured by education level, social class, household income and area-based deprivation) had stronger associations with alcohol-related hospital admissions and alcohol-related deaths at the same level of alcohol intake even after controlling for drinking patterns, smoking and BMI.[23] Similarly, excess harm in lower SES groups has been associated with the single lifestyle factors of smoking and physical inactivity.[24] The underlying mechanisms that explain differential vulnerability remain unclear but could include interactions between lower SES and other harmful factors associated with low SES (eg, other unhealthy lifestyle factors, stress, reduced access to healthcare) or be due to accelerated biological ageing in lower SES groups due to greater cumulative life course risks (eg, increased frequency of adverse childhood experiences, poorer childhood health).[25 26] However, differential vulnerability shown in these observational studies may also represent an artefact of residual confounding or could be due to lack of detail in survey or interview measurements of lifestyle factors which fails to fully capture greater intensiveness (differential exposure) of unhealthy lifestyle factors in lower SES groups (eg, lower SES groups who drink heavily may drink more than heavy drinkers in higher SES groups).[27]

More recently, there has been investigation of the interaction between combinations of lifestyle factors and SES.[3 6 21 28 29] Some studies show lower SES being associated with disproportionately higher cardiovascular disease (CVD) and all-cause mortality with combinations of unhealthy lifestyle factors.[6 30] Examining the evidence for SES influence on adverse health associated with combinations of unhealthy lifestyle factors would help unpack the evidence for and against differential vulnerability and improve our understanding of wider lifestyle associated risks across SES spectra. However, to our knowledge, there has been no systematic review of the evidence for interactions between SES and combinations of unhealthy lifestyle factors in terms of adverse health outcomes. This paper describes the protocol for a systematic review of the effects of SES on the association between combinations of unhealthy lifestyle factors and adverse health outcomes. This review will highlight evidence gaps and deepen our understanding of the complex interplay between lifestyle, SES and adverse health outcomes. Findings from this review will inform the development

Table 1 PICOS inclusion and exclusion criteria

PICOS element	Description
Population	Studies of any general population type will be included. Eligibility will not be restricted by age, sex, or other sociodemographic characteristics. Cohort studies focusing on participants with an index condition/disease will be excluded.
Exposure	Studies that examine two main exposures of interest will be included: 1. Lifestyle factor combinations Combinations must include at least three lifestyle factors and may include any combination of either conventional or emerging lifestyle factors. Combinations may include metabolic or intermediate risk factors (eg, blood pressure, cholesterol, or body mass index) but at least three factors included in the combination must be behavioural lifestyle factors (eg, smoking, physical activity and diet) as opposed to intermediate or metabolic factors. 2. Socioeconomic status All SES measures will be permitted. Anticipated variables include but are not limited to individual or area-based measures of education, employment, occupation, income and deprivation or poverty indices.
Comparator	Studies will be included where reported findings allow an assessment of the impact of SES on the association between combinations of lifestyle factors and adverse health outcomes. Comparisons of effects for available outcomes will be made, for example, HRs of participants with the 'unhealthiest' lifestyle factor combination in the most affluent SES group will be compared with the HRs of participants with the unhealthiest lifestyle but in the least affluent SES group (ie, unhealthy + high SES vs unhealthy + low SES). We will compare results for tests of interaction between lifestyle factor combinations and SES measures.
Outcomes	Primary outcome: ► All-cause mortality Secondary outcomes: ► CVD and cancer mortality ► CVD and cancer incidence
Study	Prospective observational cohort studies. Studies published in English language.
Exclusions	Ineligible publication/study design (eg, reviews, conference abstracts, case–control and cross-sectional studies, intervention studies, qualitative studies). Studies lacking exposures or outcomes of interest (eg, combinations of fewer than three lifestyle factor or SES not examined). Studies that do not provide an assessment of the impact of SES on the association between combinations of lifestyle factors and adverse health.

CVD, cardiovascular disease; PICOS, population, intervention, comparator, outcome, study design; SES, socioeconomic status.

of policy and research that aims to better support and understand healthy lifestyles and contribute to reducing the excess lifestyle-related mortality and morbidity in lower SES populations.

Aims

This review aims to identify, appraise and synthesise the findings from studies that examine the effects of SES on the association between combinations of unhealthy lifestyle factors and adverse health outcomes. This review has two key questions:

1. What are the characteristics of studies that examine the effect of SES on the association between combinations of unhealthy lifestyle factors and adverse health outcomes?
2. What is the evidence for whether and how the association between combinations of unhealthy lifestyle factors and adverse health outcomes is moderated by SES?

METHODS AND ANALYSIS

This systematic review is registered with the international database of prospectively registered systematic reviews.[31]

Eligibility criteria

Inclusion criteria are presented in table 1 according to an adapted population intervention, comparator, outcome, study design framework from the Cochrane Handbook, where 'I' (intervention) is replaced with 'E' (exposure).[32]

Population

This review will focus on the impact of SES on lifestyle associated adverse health outcomes in the general population. Of the studies included in previous systematic reviews investigating the adverse health outcomes associated with combinations of lifestyle factors, very few included an evaluation of the impact of SES.[5 8 9] Therefore, because it was anticipated that few studies would fit the inclusion criteria, the population type was not restricted in order to identify as many studies as possible.

Exposure

Only studies that examine both combinations of lifestyle factors and SES as exposure variables will be included. Studies that examine the combined influence of at least three lifestyle factors will be included. It was decided that three lifestyle factors represented a balance between identifying the evidence for combinations of lifestyle factors as opposed to single lifestyle factors (two lifestyle factors

Table 2 PubMed search strategy

Search	MeSH terms and keywords	Theme
#1	combination*[tiab] OR combined[tiab] OR composite[tiab] OR integrated[tiab] OR interaction*[tiab] OR joint effect* OR merged effect*[tiab] OR score*[tiab] OR adhere* to[tiab] OR collective[tiab] OR cumulative[tiab] OR multiple[tiab]	combined
#2	life style[MeSH] OR life style*[tiab] OR lifestyle*[tiab] OR risk reduction behavior[MeSH] OR risk reduction behaviour*[tiab] OR health behavior[MeSH] OR health behaviour[tiab] OR health factor*[tiab] OR low risk*[tiab] OR prevention guideline*[tiab] OR protective factor*[tiab] OR risk reduction behaviour*[tiab] OR health* behaviour*[tiab] OR risk behaviour*[tiab] OR modifiable factors[tiab]	lifestyle factors
#3	healthcare disparities[MeSH] OR healthcare disparities[tiab] OR Health Status Disparities[MeSH] OR disparate[tiab] OR disparit*[tiab] OR inequal*[tiab] OR health inequalities[tiab] OR unequal[tiab] OR health inequities[tiab] OR inequit*[tiab] OR socioeconomic factors[MeSH] OR socioeconomic factors[tiab] OR socio-economic*[tiab] OR socioeconomic*[tiab] OR social-economic[tiab] OR Social Determinants of Health[MeSH] OR social determinant*[tiab] OR poverty[MeSH] OR poverty[tiab] OR depriv*[tiab] OR sociological factors[MeSH] OR sociological factors[tiab] OR social medicine[MeSH] OR social medicine[tiab]	SES
#4	cohort studies[MeSH] OR cohort[tiab] OR incidence[MeSH] OR incidence[tiab] OR survival analysis[MeSH] OR survival[tiab] OR early diagnosis[MeSH] OR early diagnosis[tiab] OR prospective*[tiab] OR follow* up[tiab] OR longitudinal[tiab] OR nested case-control[tiab] OR nested case control[tiab] OR predict*[tiab]	study design
#5	#1 AND #2 AND #3 AND #4	-
#6	Death[MeSH] OR death*[tiab] OR mortality[MeSH] OR mortalit*[tiab] OR fatal*[tiab] OR life expectanc*[tiab] OR surviv*[tiab]	mortality outcome
#7	cardiovascular diseases[MeSH] OR cardiovascular[tiab] OR CVD[tiab] OR heart disease*[tiab] OR myocardial ischaemia[tiab] OR AMI[tiab] OR IHD[tiab] OR CHD[tiab] OR coronary artery disease*[tiab] OR CAD[tiab] OR myocardial infarction[tiab] OR heart infarction[tiab] OR acute coronary syndrome[tiab] OR ACS[tiab] OR heart failure[tiab] OR sudden cardiac death[tiab] OR cerebrovascular disorder*[tiab] OR cerebrovascular accident*[tiab] OR cerebrovascular attack*[tiab] OR CVA[tiab] OR cerebrovascular disease*[tiab] OR CBVD[tiab] OR cerebral arterial disease*[tiab] OR stroke*[tiab] OR apoplex*[tiab]	CVD outcomes
#8	neoplasms[MeSH] OR neoplas*[tiab] OR cancer*[tiab] OR carcinoma*[tiab] OR tumour*[tiab] OR malignanc*[tiab]	cancer outcomes
#9	#6 OR #7 OR #8	-
#10	#9 AND #5 (final search)	-

(MeSH)=Medical Subject Heading; (tiab)=contained in either title or abstract; underlined=both UK and American spellings will be searched; *=any group of letters/characters, including no character.
CVD, cardiovascular disease; SES, socioeconomic status.

was felt to be too narrow) while ensuring that a sufficient number of studies are included (there are fewer studies examining the risks of >3 lifestyle factors). In order to help identify as much literature as possible we decided that all definitions of SES variables will be accepted.

Comparator

Studies will be included if they examine the effect of SES on the associations between combinations of lifestyle factors and adverse health outcomes. Results for effects may be reported in different ways: HRs, ORs, incidence rates. Where possible, comparisons of effects for similar outcomes will be made across studies. For example, the HRs of participants with the least healthy lifestyle factor combination in the most affluent SES group will be compared with the HRs of participants with the least healthy lifestyle combination but in the least affluent SES group. Where reported, we will compare results for tests of interaction between lifestyle combinations and SES measures.

Outcomes

Our primary outcome of interest is all-cause mortality. However, lifestyle-associated adverse health is well recognised to be strongly linked to CVD and cancer outcomes. Therefore, we will include studies that examine the following outcomes: CVD and cancer mortality; CVD and cancer incidence. Studies examining specific CVD or cancer outcomes such as stroke, angina or site-specific cancer will also be included. The International Classification of Diseases (10th revision) codes I05–I89.9 and C00-C97 will be used to define CVD and cancer outcomes, respectively.

Study design

We aim to identify prospective observational cohort studies. Case–control and cross-sectional studies, intervention studies, qualitative work and review articles will be excluded. Only full-text published articles will be included and conference abstracts, dissertations,

editorials or papers without data will be excluded. Studies not published in the English language will be excluded.

Study identification

Electronic searches

A systematic search of PubMed, EMBASE and CINAHL databases will be performed. The search strategy will incorporate a combination of Medical Subject Heading (MeSH) terms and keywords. The search strategy of a recent systematic review examining combined lifestyle factors and the risk of incident type 2 diabetes was used as a template and adapted to incorporate SES related MeSH terms and keywords.[8] The search strategy has been developed with assistance from a specialist university librarian. Table 2 shows the search strategy that will be used for PubMed. The search strategy will be adapted and applied to other databases and will be available from PROSPERO once the review is complete.

Searches will be from database inception (PubMed 1966; EMBASE 1947; CINAHL 1984) to 3 March 2020. Searches will be supplemented by handsearching of reference lists of included papers, forward citation searching and a search for grey literature using the following sources:

- Charities/health organisations: for example, The King's fund, The Health Foundation, Cancer Research UK, WHO, American Heart Association, American Cancer Society.
- Databases such as OpenGrey, the Healthcare Management Information Consortium, the National Technical Information Service.
- Google and Google Scholar.
- Literature compiled by governmental organisations for example, Department of Health in England, Office for National Statistics, Centers for Disease Control and Prevention.

Data collection and analysis

Study selection

Studies identified by the search strategy will be uploaded to 'DistillerSR' software and duplicates will be removed. Two reviewers will independently screen titles and abstracts using the inclusion criteria above. Any conflicts shall be resolved by discussion and if an agreement cannot be made the study shall be included for full-text screening.

Full texts will be reviewed using a piloted checklist based on the inclusion and exclusion criteria. Conflicts will be resolved by discussion and will include a third reviewer if no consensus is reached. All studies excluded at the full-text stage will be listed with reasons for exclusion given.

Data extraction

Each study that meets inclusion criteria after full-text screening will go through the data extraction phase. Data extraction will be carried out by two reviewers working independently using a piloted data extraction form (box 1). Data will be extracted for the following study characteristics: author, publication year, title, study cohort, number of participants, proportion female, mean age, ethnicity, setting, country, date of recruitment, duration/follow-up. Details of the lifestyle or metabolic factors and SES measures used as exposure variables will be extracted. Where possible, this detail will include how and when exposure variables were measured or assessed. Data will be extracted for any included metabolic factors such as BMI, blood pressure or cholesterol levels. The number of study participants with unhealthy lifestyle factors will be recorded and reported. Details of confounder variables and the level of missing data will be extracted. Health outcome definitions and ascertainment will be recorded. The type of analysis, statistical methodology and confounder adjustment will be extracted. Study results, the nature of adverse health outcome associations identified and the effect sizes that measure the impact of SES on lifestyle associated outcomes will be recorded.

Box 1 Data extraction

Article identifiers
- Author.
- Publication year.
- Title.
- Journal, Vol, Issue, Page numbers.
- Source (eg, Database, Grey Literature source, handsearching of references, etc.).
- Study funding.

Study characteristics
- Study aims and objectives.
- Cohort name.
- Number of participants.
- Proportion female.
- Mean age (SD).
- Ethnicity.
- Country.
- Setting (eg, general population, occupational cohort, etc.)
- Study duration/follow-up.
- Study start and end dates.

Exposures, confounders and outcomes
- Lifestyle and/or metabolic factors (definition; when and how measured/assessed).
- SES measures (definition; when and how measured/assessed).
- Confounder variables or covariates included.
- Outcome definition.
- Outcome ascertainment.
- Number of participants with missing data.

Analysis characteristics
- Type of analysis (statistical methods).
- Sensitivity analysis conducted.
- Methods to deal with missing data.

Results, conclusions, and quality
- Effect of SES on lifestyle-associated adverse health: yes or no.
- Size of effect.
- Study conclusions.
- Newcastle-Ottawa Scale and justification.

Box 2 Adapted Newcastle-Ottawa scale†

Selection

1. Representativeness of the cohort
   - Truly representative of those about whom the study makes conclusions (adult/general/low income population).*
   - Somewhat representative.*
   - Selected group of users, for example, nurses, volunteers.
   - No description of the derivation of the cohort.
2. Ascertainment of exposures (lifestyle factors and SES measures)
   - Secure record (eg, healthcare records).*
   - Structured interview.*
   - Written self-report.
   - No description.
3. Analysis rigour
   - Impact of >1 socioeconomic measure assessed.*
   - Impact of one socioeconomic measure assessed.
4. Demonstration that participants were disease free at start of study
   - Yes/sensitivity analysis excluding those with prior disease.*
   - No.

Comparability (two stars available)

1. Comparability of exposed versus unexposed on the basis of the design or analysis
   - Analysis includes adjustment for key variables and maximises the use of available data.*
   - Study offers reasonable justification for adjustment variables.*

Outcome

1. Assessment of outcome
   - Independent blind assessment.*
   - Record linkage.*
   - Self-report.
   - No description.
2. Was follow-up long enough for outcomes to occur
   - Yes (≥6 months).*
   - No.
3. Adequacy of follow-up of cohorts
   - All participants accounted for and missingness <50%, or assessment of impact of missingness given, or description of those with missing data allows assessment of impact of missingness.*
   - Some participants unaccounted for or missingness >50%, or no assessment of impact of missingness, or insufficient description of those with missing data prevents assessment of impact of missingness.

* = star awarded.

†A study can be awarded a maximum of one star for each numbered item within the selection and outcome categories. A maximum of two stars can be given for comparability. A total of nine stars are available: selection (four stars), comparability (two stars) and outcome (three stars). If a single study combines/harmonises multiple cohorts for analysis, then the item that best fits with the majority of the cohorts included should be given for that study.

Whether and how sensitivity analyses were conducted will be noted. Techniques for dealing with missing data will be recorded. Studies' overall conclusions will be extracted.

Quality

Included studies will be assessed for quality using an adapted version of the Newcastle-Ottawa Scale for cohort studies,[33] a tool that has been used extensively for the appraisal of observational studies as described here. This scale, after piloting, has been adapted to include an assessment of confounder adjustment, sensitivity analyses, and dealing with missing data (box 2).

Data synthesis

The process and results of study identification and selection based on inclusion and exclusion criteria will be displayed as a Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram.[34]

To aid comparisons across studies, we will present at least the following summary data for each included study in tabular format:

- Cohort characteristics (eg, number of participants, proportion female, mean age, setting, date of recruitment, length of follow-up).
- Lifestyle or metabolic factors included in combination.
- SES measures used in assessment of SES effects
- Health outcomes and outcome ascertainment.
- Risks for health outcomes and their statistical significance.
- Inconsistent findings within each study.
- Study quality.

Studies will be grouped together according to our outcomes of interest (all-cause mortality; CVD and cancer mortality; CVD and cancer incidence). For all outcomes, we will describe details of how the outcome was assessed (eg, administrative data or questionnaires) and approach used for analysis (eg, time-to-event). We will provide details of the association between SES, lifestyle and the outcome (eg, OR, HR, etc) including the length of follow-up. We will report studies attempts to deal with confounding and discuss whether resulting associations are likely confounded. Where possible, comparisons of effects for similar outcomes will be made across studies. For example, we will compare HRs of participants in the least healthy category and most affluent SES category (reference group) with HRs of participants in the least healthy category but in the most deprived SES category.

If studies are sufficiently homogeneous in terms of participant characteristics, exposures and outcomes, we will standardise study findings for similar outcomes, provide justification for our transformation methods and combine results by a random-effects meta-analysis.[35] We will then calculate I^2 to describe the proportion of effect estimate variance due to study heterogeneity rather than chance.

Initial scoping of the literature has identified significant exposure, outcome and methodological heterogeneity across studies. Therefore, a synthesis without meta-analysis (SWiM) will likely be the most appropriate method to synthesise study findings.[36] As per SWiM, we will provide justifications for the method and presentation of study findings.

Irrespective of whether a meta-analysis is conducted, we will provide a transparent and full account of any limitations of our synthesis. Further, in conducting and reporting this systematic review, we will endeavour to fulfil,

where possible, all items proposed by the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) Group.[37] Any amendments to the review protocol will be identified and justified on completion.

Patient and public involvement

This Systematic Review constitutes a primary aspect of HF's doctoral thesis. National Health Service Research Scotland Primary Care Patient and Public Involvement (NRS PPI) Group was consulted twice as part of preparatory work for the doctoral thesis funding application.[38] The NRS PPI Group were not involved in the design of the study but have influenced how results of this review, as well as other aspects of the doctoral thesis, will be presented at two planned public engagement events over the course of the thesis.

ETHICS AND DISSEMINATION

This review will not require ethical approval as it will not involve individual-level patient data. Results will be disseminated via peer-reviewed publication, professional networks, social media, public events and conference presentations.

DISCUSSION

Both combinations of multiple unhealthy lifestyle factors and SES play major roles in mortality, CVD and cancer.[1 5 8 9 22] Numerous studies have investigated the mediating influence of lifestyle factors in attempts to explain the socioeconomic gradient in adverse health outcomes.[3 4] However, fewer studies appear to examine the relationships between lifestyle, SES and adverse health outcomes from the perspective of interactions between combinations of lifestyle factors and SES.[28] Understanding the evidence for whether and how SES influences the association between lifestyle and adverse health outcomes could inform policies and interventions that aim to support healthy lifestyles.

Scoping the literature suggests that evidence for a moderating influence of SES is mixed. Eguchi *et al*[28], with data from 42 647 Japanese adults aged 40–79 years and approximately 20 years follow-up, examined the risks associated with a lifestyle score (comprised of eight lifestyle factors: smoking, alcohol, physical activity, sleep, dietary intake of fruit, fish and milk, and BMI) stratified by age (≥16 or <16 years) at last formal education. The authors reported a 44% higher all-cause mortality risk for those with a higher level of education but with the least healthy lifestyle, compared with those with a higher level of education but with the healthiest lifestyle. When the same comparison was made in those with the lower level of education, participants with the least healthy lifestyle had a 40% higher all-cause mortality risk compared with those with the healthiest lifestyle. Namely, the level of elevated risk associated with the least healthy lifestyle was similar in both higher and lower education groups.

Foster *et al*[6], performed similar analyses with data from 328 594 UK adults 40–69 years and approximately 5 years follow-up to examine the risks associated with a lifestyle score (comprised of nine lifestyle factors: smoking, alcohol, physical activity, television viewing time, sleep duration, and dietary intake of fruit/vegetables, oily fish, and red and processed meat) stratified by quintiles of socioeconomic deprivation (Townsend index). The authors observed a 65% higher all-cause mortality risk for the least deprived with the least healthy lifestyle, compared with the least deprived but with the healthiest lifestyle. However, when the same comparison was made in the most deprived participants, those with the least healthy lifestyle had a 145% higher risk than those with the healthiest lifestyles.

These highly comparable studies report opposing results. Eguchi *et al*[28] found no interaction between lifestyle and SES with similar lifestyle risks in the least and most educated groups. Whereas Foster *et al*[6] found an interaction between lifestyle and SES with disproportionately raised lifestyle risk in the most deprived group. Several methodological differences (lifestyle and SES measures; follow-up time; population characteristics) could explain the two studies' conflicting results but additional evidence from similar studies would help to clarify whether there is any moderating influence of SES on lifestyle-associated risks. This systematic review will help unpack such associations in more detail. In addition, included studies may identify specific combinations of unhealthy lifestyle factors that pose the highest risks for lower SES groups. However, we suspect there is likely to be a lack of studies which identify the combinations that pose the greatest risk for lower SES groups and this may be one of the evidence gaps that this review identifies.

To our knowledge, this systematic review will be the first to synthesise the evidence for whether and how SES influences the association between combinations of lifestyle factors and adverse health. We will describe the lifestyle factors, SES measures and adverse health outcomes that have been examined thus far. We will collate and interpret the findings considering both the type of analyses and the quality of studies to provide a comprehensive synthesis of available evidence and highlight gaps in current evidence.

We have developed a comprehensive search strategy with broad inclusion criteria in order to identify all available evidence and reduce the chance of omitting relevant studies. However, our scoping of the literature suggests that not only will there be few studies that attempt to examine this problem but also that previous studies will be widely heterogeneous both in terms of the lifestyle factor and SES variables examined and in terms of the statistical methods employed. This will likely preclude a meta-analytical synthesis of the evidence, which may be a limitation of our review. However, SWiM will likely highlight important gaps in available evidence and direct future research in this sphere. This review will adhere to SWiM reporting guidelines and will be guided by MOOSE recommendations to improve transparency and clarity.[36 37]

All screening, data extraction and quality assessment will be performed independently by two reviewers to improve study rigour.

This comprehensive and rigorous systematic review will improve our understanding of the complex interaction between SES and lifestyle and has the potential to inform research, interventions and policy.

Twitter Hamish Foster @hamishfoster, Frances Mair @FrancesMair and Catherine A O'Donnell @odo_kate

Contributors HF, CAO, FM and JG were involved in study concept. HF, CAO, FM, PP and JG developed the study design. HF, CAO, FM, JG and PP were involved in acquisition, analysis or interpretation of data. Drafting of the manuscript was led by HF and PP with supervision and support from CAO, FM and JG. All authors were involved in critical revision of the manuscript for important intellectual content. CAO is the guarantor.

Funding HF is funded by a Medical Research Council Clinical Research Training Fellowship (MRC CRTF; MR/T001585/1) and this body of work constitutes a primary aspect of HF's doctoral thesis.

Disclaimer The funder had no role in developing this protocol.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Hamish Foster http://orcid.org/0000-0002-0224-7125
Peter Polz http://orcid.org/0000-0003-1524-8482
Catherine A O'Donnell http://orcid.org/0000-0002-5368-3779

REFERENCES

1 GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the global burden of disease study 2017. *Lancet* 2018;392:1923–94.
2 Meader N, King K, Moe-Byrne T, *et al*. A systematic review on the clustering and co-occurrence of multiple risk behaviours. *BMC Public Health* 2016;16:657.
3 Laine JE, Baltar VT, Stringhini S, *et al*. Reducing socio-economic inequalities in all-cause mortality: a counterfactual mediation approach. *Int J Epidemiol* 2020;49:497–510.
4 Marmot MG, Shipley MJ, Hemingway H, *et al*. Biological and behavioural explanations of social inequalities in coronary heart disease: the Whitehall II study. *Diabetologia* 2008;51:1980–8.
5 Loef M, Walach H. The combined effects of healthy lifestyle behaviors on all cause mortality: a systematic review and meta-analysis. *Prev Med* 2012;55:163–70.
6 Foster HME, Celis-Morales CA, Nicholl BI, *et al*. The effect of socioeconomic deprivation on the association between an extended measurement of unhealthy lifestyle factors and health outcomes: a prospective analysis of the UK Biobank cohort. *Lancet Public Health* 2018;3:e576–85.
7 Pepper GV, Nettle D. The behavioural constellation of deprivation: causes and consequences. *Behav Brain Sci* 2017;40:e314.
8 Zhang Y, Pan X-F, Chen J, *et al*. Combined lifestyle factors and risk of incident type 2 diabetes and prognosis among individuals with type 2 diabetes: a systematic review and meta-analysis of prospective cohort studies. *Diabetologia* 2020;63:21–33.
9 Zhang Y-B, Pan X-F, Chen J, *et al*. Combined lifestyle factors, incident cancer, and cancer mortality: a systematic review and meta-analysis of prospective cohort studies. *Br J Cancer* 2020;122:1085–93.
10 Jike M, Itani O, Watanabe N, *et al*. Long sleep duration and health outcomes: a systematic review, meta-analysis and meta-regression. *Sleep Med Rev* 2018;39:25–36.
11 Sun J-W, Zhao L-G, Yang Y, *et al*. Association between television viewing time and all-cause mortality: a meta-analysis of cohort studies. *Am J Epidemiol* 2015;182:908–16.
12 Krokstad S, Ding D, Grunseit AC, *et al*. Multiple lifestyle behaviours and mortality, findings from a large population-based Norwegian cohort study - The HUNT Study. *BMC Public Health* 2017;17:58.
13 Verplaetse TL, McKee SA. An overview of alcohol and tobacco/nicotine interactions in the human laboratory. *Am J Drug Alcohol Abuse* 2017;43:186–96.
14 Linardakis M, Smpokos E, Papadaki A, *et al*. Prevalence of multiple behavioral risk factors for chronic diseases in adults aged 50+, from eleven European countries - the SHARE study (2004). *Prev Med* 2013;57:168–72.
15 Lawder R, Harding O, Stockton D, *et al*. Is the Scottish population living dangerously? prevalence of multiple risk factors: the Scottish health survey 2003. *BMC Public Health* 2010;10:330.
16 Redondo-Bravo L, Fernández-Alvira JM, Górriz J, *et al*. Does socioeconomic status influence the risk of subclinical atherosclerosis?: A mediation model. *J Am Coll Cardiol* 2019;74:526–35.
17 Stringhini S, Zaninotto P, Kumari M, *et al*. Socio-economic trajectories and cardiovascular disease mortality in older people: the English longitudinal study of ageing. *Int J Epidemiol* 2018;47:36–46.
18 Sydén L, Landberg J. The contribution of alcohol use and other lifestyle factors to socioeconomic differences in all-cause mortality in a Swedish cohort. *Drug Alcohol Rev* 2017;36:691–700.
19 Oude Groeniger J, Kamphuis CB, Mackenbach JP, *et al*. Repeatedly measured material and behavioral factors changed the explanation of socioeconomic inequalities in all-cause mortality. *J Clin Epidemiol* 2017;91:137–45.
20 Stringhini S, Sabia S, Shipley M, *et al*. Association of socioeconomic position with health behaviors and mortality. *JAMA* 2010;303:1159–66.
21 Nordahl H, Lange T, Osler M, *et al*. Education and cause-specific mortality. *Epidemiology* 2014;25:389–96.
22 Stringhini S, Carmeli C, Jokela M, *et al*. Socioeconomic status and the 25 × 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1·7 million men and women. *The Lancet* 2017;389:1229–37.
23 Katikireddi SV, Whitley E, Lewsey J, *et al*. Socioeconomic status as an effect modifier of alcohol consumption and harm: analysis of linked cohort data. *Lancet Public Health* 2017;2:e267–76.
24 Pampel FC, Rogers RG, status S. Socioeconomic status, smoking, and health: a test of competing theories of cumulative advantage. *J Health Soc Behav* 2004;45:306–21.
25 Belsky DW, Caspi A, Cohen HJ, *et al*. Impact of early personal-history characteristics on the pace of aging: implications for clinical trials of therapies to slow aging and extend healthspan. *Aging Cell* 2017;16:644–51.
26 Fiorito G, Polidoro S, Dugué P-A, *et al*. Social adversity and epigenetic aging: a multi-cohort study on socioeconomic differences in peripheral blood DNA methylation. *Sci Rep* 2017;7:16266.
27 Lewer D, Meier P, Beard E, *et al*. Unravelling the alcohol harm paradox: a population-based study of social gradients across very heavy drinking thresholds. *BMC Public Health* 2016;16:599.
28 Eguchi E, Iso H, Honjo K, *et al*. No modifying effect of education level on the association between lifestyle behaviors and cardiovascular mortality: the Japan collaborative cohort study. *Sci Rep* 2017;7:39820.
29 Warren Andersen S, Zheng W, Sonderman J, *et al*. Combined impact of health behaviors on mortality in low-income Americans. *Am J Prev Med* 2016;51:344–55.
30 Nordahl H, Lange T, Osler M, *et al*. Education and cause-specific mortality: the mediating role of differential exposure and vulnerability to behavioral risk factors. *Epidemiology* 2014;25:389–96.
31 NIHR. Understanding the impact of socioeconomic status on the association between combined lifestyle factors and adverse health outcomes: a systematic review. PROSPERO 2020 CRD42020172588, 2020. Available: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020172588
32 McKenzie J, Brennan SE, Ryan RE. *Chapter 3: defining the criteria for including studies and how they will be grouped for the synthesis. In: Cochrane Handbook for systematic reviews of interventions version 6.0*. London: Cochrane, 2019.

33 Wells GA, Shea B, O'Connell D, *et al*. *The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses*. Ottawa: Ottawa Hospital Research Institute, 2021. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

34 Moher D, Liberati A, Tetzlaff J. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement [Guideline Research Support, Non-US Gov't]. *Bmj* 2009;339:b2535.

35 Deeks J, Higgins JPT, Altman DG. *Chapter 10: analysing data and undertaking meta-analyses. In: Cochrane Handbook for systematic reviews of interventions version 6.0 (updated July 2019)*. London: Cochrane, 2019. www.training.cochrane.org/handbook

36 Campbell M, McKenzie JE, Sowden A, *et al*. Synthesis without meta-analysis (swim) in systematic reviews: reporting guideline. *BMJ* 2020;72:l6890.

37 Stroup DF, Berlin JA, Morton SC, *et al*. Meta-analysis of observational studies in epidemiology: a proposal for reporting. meta-analysis of observational studies in epidemiology (moose) group. *JAMA* 2000;283:2008–12.

38 The NRS Primary Care PPI Group. Get involved, 2020. Available: https://www.nhsresearchscotland.org.uk/research-areas/primary-care/get-involved

Appendix 2: Ethics application and participant information sheet

University of Glasgow | College of Medical, Veterinary & Life Sciences

College of Medical, Veterinary & Life Sciences Ethics Committee for Non-Clinical Research Involving Human Participants

APPLICATION FORM FOR ETHICAL APPROVAL FOR RESEARCH INVOLVING FACE TO FACE CONTACT WITH PARTICIPANTS

NOTES:
THIS APPLICATION FORM SHOULD BE TYPED NOT HANDWRITTEN.

ALL QUESTIONS MUST BE ANSWERED. "NOT APPLICABLE" IS A SATISFACTORY ANSWER WHERE APPROPRIATE.

The University now requires that all research proposals involving the collection, processing and/or storage of data derived from human participants, that are submitted to a College Research Ethics committee for review, must be accompanied by a Data Protection Impact Assessment (DPIA).

- **Information on DPIAs and Privacy Notices**
- **Information on the General Data Protection Regulation (GDPR)**
- **Information on Research Data Management**
- **University of Glasgow policy on surveys of students for research purposes**

If your research involves participants outside Great Britain, Northern Ireland, the Channel Islands or the Isle of Man, pregnant participants or more than 5000 participants the project MUST be referred to the Research Support Office and insurance coverage confirmed before ethical approval is sought. Please contact Dr Debra Stuart in the University's Research Governance Office: debra.stuart@glasgow.ac.uk

Project Details

1. Project title

Understanding interactions between lifestyle and deprivation to support policy and intervention development.

2. Name and position of Principal Researcher

Dr Hamish Foster, Clinical Academic Fellow, PhD candidate

3. Has this application been previously submitted to this or any other ethics committee? If 'Yes', please state the title and reference number.

No

4. List who is doing the research and their qualifications.

Dr Hamish Foster
HF, the primary investigator, is a Clinical Research Fellow and GP. He is a PhD candidate at University of Glasgow and has had some qualitative research experience prior to his PhD. He works in one of the most socioeconomically deprived practices in Scotland and has developed links with colleagues across Glasgow. He will be supported by his supervisors who have extensive experience in qualitative methods and will attend FG and interview moderation courses prior to data collection.

Prof Kate O'Donnell
KOD is a primary care researcher with over 20 years' experience of evaluating large scale primary care innovation and policy, including Keep Well and the Links Worker Programme in Scotland. Much of her work is mixed methods and informed by theoretical frameworks (http://www.gla.ac.uk/researchinstitutes/healthwellbeing/staff/kateodonnell/). KOD has experience in conducting work with hard-to-reach groups, focusing on their views of their health and wellbeing. She has supervised 15 postgraduate research students to successful completion.

Prof Frances Mair
FM is a general practitioner and Head of General Practice and Primary Care at UoG. FM leads an extensive programme of work using mixed methods research and focussing on optimising the care of patients with chronic illness and multimorbidity. FM's work considers the wider socioeconomic environment and social contexts and the importance of understanding implementation issues to help bridge the translational gap between research and practice/policy.

Professors Jason Gill and Duncan Lee are also part of the supervisory team, but principally involved with the earlier quantitative work.

5. Is this a student research project? If yes, confirm if undergraduate, post graduate research (PGR) or post graduate taught (PGT) and confirm supervisory arrangements.

Yes, this is a PGR project. Primary supervisors: Professors Kate O'Donnell and Frances Mair. Secondary supervisors: Professors Jason Gill, and Duncan Lee.

6. Describe the purpose of the research proposed.

Summary & Importance: Unhealthy lifestyle factors (e.g. smoking, physical inactivity) are the main risk factors for non-communicable diseases which, in turn, are the commonest causes of death worldwide. Socioeconomically deprived populations have the highest mortality and morbidity rates which is partly explained by having the greatest prevalence of unhealthy lifestyle factors.[15]

Deeper explanations for lifestyle-related health inequalities include synergistic interactions between lifestyle factors, where combinations of unhealthy lifestyle factors have a multiplicative adverse effect on health. However, interactions between lifestyle factors and deprivation, where deprived populations experience disproportionate harm from unhealthy lifestyles factors, are also implicated.[24] To date, there has been limited examination of these types of interactions. It is unclear which lifestyle factor combinations incur the greatest risk and whether high-risk combinations vary by deprivation with previous studies only focussing on a few well-recognised lifestyle factors.

This PhD aims to address these evidence gaps and deepen our understanding of the complex interplay between lifestyle, deprivation, and ill-health. It will examine the interactions between combinations of lifestyle factors (including emerging lifestyle factors e.g., sleep duration[56]) and deprivation. This will inform the development of policy and interventions at both individual and population-level that aim to better support healthy lifestyles in deprived populations and reduce the excess lifestyle-related mortality and morbidity.

Background: Our understanding of lifestyle associated risks is increasing and beginning to reflect the more complex reality where multiple lifestyle factors are involved. Single lifestyle factors are strongly associated with mortality, but combinations of factors have been shown to interact synergistically.[374] For example, the mortality associated with smoking and high alcohol intake together is more than the sum of the risks associated with each factor alone. Moreover, 'new' lifestyle factors (e.g., sleep duration[56], or television (TV) viewing times[465]) have been shown to be associated with increased mortality, and combinations of 'new' and 'conventional' factors (e.g., smoking, physical activity, alcohol, and diet) are also associated with increased mortality.[10]

However, our understanding of how the risks associated with these broader combinations of lifestyle factors vary with deprivation remains poor and under-investigated. This is despite a well-known socioeconomic gradient, where deprived groups are more likely to have multiple co-occurring unhealthy lifestyle factors.[15] There is some evidence for interactions between some conventional lifestyle factors and deprivation. For example, deprived populations are at greater risk of harm from lower levels of alcohol intake compared to more affluent populations, even after controlling for drinking patterns.[175] Excess harm in deprived groups has also been associated with smoking and physical inactivity.[176] However, studies in this area: i) rarely assess combinations of lifestyle factors that include 'new' factors; ii) are yet to clarify which combinations pose the greatest risks; and iii) lack outcomes other than all-cause and cardiovascular disease (CVD) mortality.

We recently reported a survival analysis of the UK Biobank cohort in Lancet Public Health,[24] which demonstrated an interaction between a combination of lifestyle factors and deprivation for all-cause and CVD mortality. The combination of factors

incorporated new and conventional factors (smoking, alcohol, physical activity, TV viewing time, sleep duration, and four dietary components). Our results show that a broad combination of lifestyle factors is associated with disproportionate harm in deprived populations. In other words, deprived participants experienced much more lifestyle-related ill-health than their affluent counterparts even when they had a similar lifestyle. This strengthens arguments for developing lifestyle policies and interventions that better support deprived populations and for extending the range of lifestyle factors that interventions target.

This mixed methods PhD is building on this previous work by developing a more detailed understanding of the risks associated with a wider range of lifestyle factors (including factors we have already studied but also other factors such as social participation levels and salt/sugar intake) and their interactions with deprivation. This work will identify combinations of lifestyle factors that are most strongly associated with adverse health outcomes in deprived populations. Outcomes of interest include all-cause mortality, cardiovascular morbidity and mortality and cancer morbidity and mortality. This will be explored by quantitative studies using data from two pre-existing cohorts. Work analysing the first cohort (UK Biobank) is already covered by ethical approval (NHS National Research Ethics Service (16/NW/0274); the second quantitative study will utilise data from a Norwegian cohort (the HUNT dataset), and ethical approval has been granted by the local Norwegian ethics committee which scrutinises studies using HUNT data.

However, while such quantitative analyses are important to describe the situation, it is vital to understand the views of the public, professionals, and policymakers in relation to the findings and their implications. Therefore, the final study of this PhD is a qualitative exploration of key stakeholder views of the barriers and facilitators to incorporating this evidence into daily life, practice, and policy to support healthy living in deprived areas. It is this qualitative work for which this application seeks ethical approval. The new evidence generated from this PhD, including the qualitative component, will inform both government policy and public health interventions as well as the future development of novel interventions that can target high-risk individuals.

Research Plan: A qualitative study involving the public and health care professionals living and working in areas of socioeconomic deprivation, and policymakers working in public health and primary care.

Research questions

RQ1: How do the public, healthcare professionals, and policymakers view high risk lifestyle factor combinations in the context of deprivation?

RQ2: What are the barriers and facilitators to incorporating the new evidence about the association of deprivation and lifestyle with poor health outcomes into daily life, practice, and policy to support healthy living in deprived areas?

7. Does this research involve interaction with NHS colleagues or the use of NHS buildings, facilities or data?

Health care professionals (community nurses, general practitioners, and public health practitioners) will be invited to participate, and they may happen to work in the NHS.

However, the research will not involve NHS facilities or data and all recruitment will take place outwith the NHS.

8. Who are the participants? Describe how potential participants will be *identified and approached*. Justify inclusion of any vulnerable groups. Include any recruitment adverts / materials with your application.

This will be a qualitative study with data collected via audio recording from two main groups of participants: 1) members of public for focus groups (FGs) and 2) health care professionals and policymakers for in-depth interviews.

(i) Identified

a) Focus Group participants (members of the public)

FG participants will be invited through the research team's extensive links with the following community and third sector organisations: The Alliance (www.alliance-scotland.org.uk); The Poverty Alliance; Scottish Community Development Centre; REACH Community Service; the Community Links Worker Network. Key contacts within these organisations will be asked to cascade publicity material about the project to their members. This material will outline the aim of the project, what activities participants will engage in, and how the information generated with be used. Key contacts will also be asked to identify individuals who might wish to participate; we will cascade information about this project through their organisational links; and we will advertise through Twitter and Facebook. Examples of our invitation material including images that will be used both in conjunction with study advertisements and as prompts in FGs are included (see 2022_05_04_Invitations_to_Participate_v1.docx, 1-_Lifestyle.jpg. 2-_Financial.jpg, 3-both.jpg).

b) Interview participants (professionals)

Clinicians will be identified via the research team's longstanding links with primary care clinician groups such as GPs in the Deep End and RCGP. Public health professionals and policy makers will be identified via extensive links with staff from Public Health Scotland and Scottish Government. Study information and invitations will be cascaded via these links and within organisational networks. Examples of our invitation material including images that will be used both in conjunction with study advertisements and as prompts in interviews are included (see 2022_05_04_Invitations_to_Participate_v1.docx, 1-_Lifestyle.jpg. 2-_Financial.jpg, 3-both.jpg).

(ii) Approached

Potential participants who respond to adverts or to contacts within linked organisations will be asked to provide their email address or phone number. They will then be sent a copy of the appropriate Participant Information Sheet and be asked to respond if willing to participate (see 2022_05_03_Participant_Information_focus_groups_v2.docx, 2022_05_03_Participant_Information_interviews_v2.docx). In some cases, for example, people nominated within organisations, professionals and policymakers, information material will be sent to named individuals. In those circumstance, HF will contact them 5 to 10 days after sending study information to see if they are interested in participating. FG times will be pre-arranged; individual interviews will take place at a time suitable for the interviewee.

9. Describe how you will obtain consent.

Prior to FGs or interviews, potential participants will be sent a study information pack, comprising the appropriate Participant Information Sheet, a copy of the Topic Guide with the topics for discussion, and a Consent Form. This will be sent by email or, if preferred, by post. Before the interview or FG begins, written consent, in the form of their name on the consent form, will be obtained. Consent forms will be returned electronically, or in a reply-paid envelope, to HF. At the beginning of each FG/interview, HF will go through the Participant Information Sheet and the Consent Form, to ensure that the written consent is fully understood. At the start of the audio recording, verbal consent will be obtained and audio recorded. The contact details, consent form, and data of potential participants who decline to participate and of those who withdraw from the study will be destroyed.

PLEASE NOTE: Some participants might be recruited because of they are part of patient groups e.g., through the Alliance. Whilst their views as patients with health problems are of great interest, all recruitment will take place outwith the NHS.

10. Include the text you will use to inform participants about the study.

"PARTICIPANT INFORMATION SHEET

1. Study title:

'Understanding the relationships between lifestyle and social/financial circumstances and how they affect people's health'

2. Invitation paragraph

You are being invited to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. If you decide to take part in this study, you will be given a copy of this Participant Information Sheet and the signed consent form to keep

3. What is the purpose of this study?

We know that different things that we do can affect our health e.g., smoking or drinking heavily, but it is not always easy to know how to make changes, especially if people are also dealing with lack of money or other resources. Research has also identified a growing number of things that can affect our health. The purpose of his this study is to explore what people think about the risks to their health from combinations of health behaviours or lifestyle factors (e.g., the amount they smoke or drink alcohol, their diet, and the amount of time they spend being physically active, sitting, sleeping, and engaging with social activities like sports or social clubs) together with their social and financial circumstances and to explore ways of addressing these that don't add extra burden to people's lives.

We know that there are many aspects of people's lives that can affect their risk of having health problems. Health behaviours are only one part of that. The constraints and challenges they people experience due to wider social and financial conditions, such as a lack of money or employment, are also very important.

The aim of this research is to achieve a better understanding of the views of both public and professionals and therefore guide governments, charities, health professionals, and researchers in developing better ways of supporting people to live healthy lives.

This research is different from previous studies because it will ask participants to consider many different lifestyle factors at the same time as their social and financial circumstances.

We want to explore the views held by members of the public and health professionals in terms of:

1) understanding specific combinations of health behaviours that pose very high risks for health (e.g., regular cigarette smoking and drinking alcohol and eating too much processed meat)
2) how the risk from combinations of multiple health behaviours might change according to social and financial conditions
3) how people consider and address combinations of health behaviours considering their social and economic circumstances
4) what could make a difference to support healthy living in those affected by difficult social and financial circumstances.

4. Why have I been invited to participate?

We want to get views and opinions of members of the public from a range of backgrounds. However, we have a particular interest in hearing from people who may have experience of challenging social or economic circumstances.

We aim to recruit 24 members of the public to take part in focus groups (guided discussions with groups of 6-8 people). Once we have agreement from enough participants, we will stop inviting new participants. We are also interested in hearing from health and public health professionals who are involved in working with those affected by socioeconomic adversity. Therefore, as a separate part of this study we will also be recruiting 16 professionals for interviews.

You can only take part in this study if:

- you are over 18 years old
- you speak and understand English; unfortunately, we do not have a budget for translators

5. Do I have to take part?

No, it is up to you to decide whether to take part. If you do decide to take part, you are free to withdraw at any time without giving a reason. Also, if you take part, you will be given this information sheet to keep and you will be asked to sign a consent form.

6. What will happen to me if I take part?

If you say you are willing to take part in the study, you will be contacted by a researcher who will be able to answer any questions you have about the research. If you decide to participate, the researcher will send you a consent form by email (or post if you prefer) which you will be asked to complete and return to the researcher (postage will be paid if you cannot use email). The researcher will then sign the consent form as well and return the completed consent form to you, which you can keep.

The study involves taking part in a single focus group carried out either face-to-face or online with up to 8 other members of the public and one researcher. The focus group will last about one to one and a half hours. The decision to meet face to face and the location of the meeting will be made via discussion with the researcher and may change if needed due to COVID restrictions. A time and location that is mutually convenient will be arranged. The University of Glasgow's General Practice and Primary Care building on Horselethill Road is available to use as a safe and central location. Alternative options include local community or health centres.

We will pay for travel expenses if required, and we will give you a £20 gift voucher to thank you for your time. The focus group, whether face-to-face or via Microsoft Teams/Zoom, will be audio recorded.

What will happen in the focus group:

During the focus group the researcher will welcome everyone to the meeting, they will introduce themselves and explain the purpose of the research and how focus groups work. These focus groups will work by:

1. having a relaxed atmosphere – it could seem daunting to some people, but the researcher will help create a non-judgmental space and encourage everyone to respect each other's views
2. people taking part by discussing their own experiences; the researcher will explain that there are no wrong answers!
3. allowing each person taking part a chance to speak
4. using a pre-made guide (see Topic Guide) to provide questions to stimulate relevant discussion

The recordings from the focus group will be typed up by a professional transcription service and any identifying information (names, addresses) will be removed. The typed-up transcript will be stored on a secure university server. The focus groups are part of a wider study, due to finish in March 2023.

7. What do I have to do?

To take part, we ask that you to attend the focus group and take part in the discussion. You do not need to do anything else for this part of the study.

8. What are the possible disadvantages and risks of taking part?

Different people have different experiences and there can be disagreement among participants during the discussion. The researcher will explain the group rules at the start of the discussion. These rules prioritise respect throughout as well as confidentiality of each other's information, especially outside the focus-group. The researcher will also ensure that everyone is able to give their opinion.

We recognise we will be discussing and collecting potentially sensitive information regarding your health and way of life. We will follow all relevant legislation in managing and processing this data.

9. What are the possible advantages and benefits of taking part?

You will receive no direct benefit from taking part in this study. However, the information that is collected during this study will give us a better understanding of what helps and what prevents people from considering combinations of health behaviours in light of their social and financial circumstances. This information may be used to design services and supports that better meet people's needs.

10. Will my taking part in this study be kept confidential?

All information collected about you will be kept strictly confidential, including the responses you provide during the focus groups. You will be identified by an ID number, and any information about you will have your name and address removed so that you cannot be recognised from any information you provide. Please note that assurances on confidentiality will be strictly adhered to unless evidence of serious harm, or risk of serious harm, is uncovered. In such cases, the University may be obliged to contact relevant statutory bodies/agencies.

Any data in paper form will be stored in locked cabinets in rooms with restricted access at the University of Glasgow. All data in electronic format will be stored on secure password–protected computers. No one outside the research team or appropriate governance staff will be able to find out your name, or any other information which could identify you.

11. What will happen to my data?

University of Glasgow is the sponsor for this study based in the UK. We will be using information from you to undertake this study and will act as the data controller for this study. This means that we are responsible for looking after your information and using it properly.
Researchers from the University of Glasgow collect, store, and process all personal information in accordance with the General Data Protection Regulation (2018).

Your rights to access, change or move your information are limited, as we need to manage your information in specific ways for the research to be reliable and accurate. If you withdraw from the study, we will keep the information about you that we have already obtained. To safeguard your rights, we will use the minimum personally identifiable information possible.

You can find out more about how we use your information here:
https://www.gla.ac.uk/myglasgow/dpfoioffice/guidanceonresearch/

The data will be stored in archiving facilities in line with the University of Glasgow retention policy of up to 10 years. After this period, further retention may be agreed, or your data will be securely destroyed in accordance with the relevant standard procedures.
Your identifiable information might be shared with people who check that the study is done properly and, if you agree, in coded form with other organisations or universities to carry out research to improve scientific understanding. Your data will form part of the study result that will be published in expert journals, presentations, student theses and on the internet for other researchers to use. Your name will not appear in any publication.

12. What will happen to the results of the research study?

The results of the study will be published in academic journals, presented at conferences, and be used in Dr Foster's thesis for his PhD. We also plan to feedback the findings of our research to the community via presentations at public events and by working with our third sector partners (The Poverty Alliance, and ALLIANCE Scotland). Direct quotes from the transcripts of the focus groups may be used in these publications – these will be presented anonymously with no information that would be able to link the quotes back to a participant.

13. Who is organising and funding the research?

This research is being funded by the Medical Research Council. You can find more information on the overall research project here:
https://gtr.ukri.org/projects?ref=MR%2FT001585%2F1

14. Who has reviewed the study?

The project has been reviewed by the Ethics Committee based at the University of Glasgow's College of Medical, Veterinary & Life Sciences.

15. Contact for Further Information

Thank you for reading this information sheet. If you have any further questions, please contact:
Dr Hamish Foster
Email: hamish.foster@glasgow.ac.uk
Tel: 0141 330 8333

If you have complaints about this study, please contact:

Dr David Blane, david.blane@Glasgow.ac.uk
or
Dr Bhautesh Jani, bhautesh.jani@Glasgow.ac.uk"

Please see the included equivalent information sheet for interviewees (2022_05_03_Participant_Information_interviews_v2.docx).

11. Describe the research and what will be expected of participants.
If you are performing a survey, what you will ask participants? You should upload a topic guide / example interview with your application.
If you are performing a behavioural experiment, you should describe this in detail.

We plan to collect data from:

- 3 x FGs with 6-8 members of the public per group (2hrs/FG). If focus groups are problematic for some people, we will offer an individual or group interview.
- 8 interviews with primary care practitioners (60 minutes each)
- 8 interviews with public health professionals and policymakers (60 minutes each)

Data collection procedures

The research procedures and how they affect participants differ according to the mode of data collection (i.e., focus group vs. interview). Topic guides for both FGs and interviews are uploaded.

Study outline for a focus group member:

Initial contact and consent-

a) Participant alerted to research e.g., via social media, via ALLIANCE meeting, Links worker, or other organisation

b) Those interested in participation contact HF by email or telephone and leave their contact details (email or telephone number) which will be recorded in a password protected Excel recruitment file on a UoG server. In some cases, where an individual was nominated by an organisation, HF will contact them 5-10 days after sending study information

c) HF provides participant with study information to permit valid consent

d) HF arranges participant to sign and return the consent form for those who agree to participate. HF permanently deletes contact details for those who fail to be contacted after two attempts or decline participation

e) HF contacts consenting participant to arrange a mutually convenient date and time for FG (or Zoom/MS Teams call). Allergies and dietary requirements will be checked at this point as lunch/meal will be offered after face-to-face FG

f) HF assigns participant a unique study code/identifier in secure Excel recruitment file

g) HF gives participant study information again with reminders of FG process

h) HF sends participant a reminder 2 days prior to FG (via email or telephone as previously agreed by participant)

Focus group-

a) Participant attends local facility (e.g., community centre, health centre, GPPC building) at agreed time and welcomed by HF – or joins video link if not able to meet face-to-face

b) Participant invited to introduce themselves to other members of the group; HF introduces themselves and reminds all participants of the study aims, and objectives; the FG process and confidentiality; how their data will be stored, protected, and used; how study findings will be used (e.g., publication of findings, conference presentations); that they may withdraw at any point; participants invited to ask any questions prior to starting the recording (20 minutes)

c) Participant notified that recording has started and contributes to group discussion with intermittent involvement and guidance from HF (30-60 minutes)

d) FG ends, participant notified that recording has stopped, invited to ask questions, and offered the opportunity to be contacted regarding study findings

e) Participant offered lunch if meeting face-to-face (30 minutes), which will be provided by study budget

f) Participant thanked and reimbursed for travel costs (up to £30, if meeting face-to-face) and offered a £20 gift voucher as a thank you for their contribution (for both face-to-face and video link participants)

Post analysis-

a) Participant who requests to be contacted regarding study findings will be directed to any published material and sent a lay summary of study findings

Study outline for an interviewee:

Initial contact and consent-

a) Participant alerted to research e.g., via social media, Deepend GP meeting, other organisational contact, or via direct contact with research team

b) Those interested in participation contact HF by email or telephone and leave their contact details (email or telephone number) which will be recorded in a password protected Excel file on a UoG server. In some cases, where individual was nominated by an organisation, HF will contact them 5-10 days after sending study information.

c) Participant contacted by HF who provides study information to permit valid consent

d) HF arranges the participant to sign and return the consent form for those who agree to participate (contact details are permanently deleted for those who fail to be contacted after two attempts or decline to participation)

e) HF contacts participant who consents to arrange a mutually convenient date and time for face-to-face, telephone, or video link (Zoom/MS Teams) interview

f) HF assigns participant a unique study code/identifier in secure Excel recruitment file

g) HF gives participant study information again with reminders of interview process

h) HF sends participant reminder 1-2 days prior to interview (email or telephone call as previously agreed by participant)

Interview-

a) Face-to-face interview - participant attends local facility and welcomed by HF at agreed time (e.g., GPPC building); Telephone or video-link interview - participant contacts at agreed time

b) Participant reminded of the study aims, and objectives, reminded of interview process, confidentiality, reminded of how their data will be stored and protected, how study findings will be used (e.g., publication of findings, conference presentations), and invited to ask any questions prior to audio recording (5-10 minutes)

c) Participant notified that interview/audio recording started and contributes to discussion in response to questions from HF (30-60 minutes)

d) Interview ends and participant notified that recording stopped, invited to ask questions, and offered the opportunity to be contacted regarding study findings

e) Participant thanked and reimbursed for travel costs (up to £30) and offered reimbursement in line with NIHR INVOLVE rates (£30.10 for nurses, £80 for GPs/public health professionals) as a thank you for their contribution

Post analysis-

a) Participant who requests to be contacted regarding study findings will be directed to any published material and sent a lay summary of study findings

Interview and focus group methods

Participant numbers and meeting length are based on substantial previous qualitative data collection experience of the supervisory team. The volume of data generated from the above plan will very likely lead to qualitative data saturation and information redundancy,[466] where no new qualitative insights can be gleaned from additional FGs/interviews. To ensure sufficient data collection and to ensure data saturation is reached, we have 'over-budgeted' in the above plan. Therefore, if data saturation (in terms of no new substantive issues being raised by participants) is reached prior to the planned number of FGs and interviews have been recorded, then we will reduce the total number of FGs and interviews accordingly and provide justification.

FGs and interviews will be led by the primary researcher (HF), supported by primary supervisor COD, and guided by topic guides. Topic guides have been developed in discussion within the research team and will be informed by Normalisation Process Theory (NPT).[467,468] NPT is a social theory used to understand the implementation and embedding of complex interventions into daily life and working practices. This highlights the action and work involved in lifestyle, practice, or policy change in light of the new evidence around lifestyle risk. Results from the on-going quantitative analyses, conducted as part of this PhD, will also inform topic guides as to the type of lifestyle factors/health behaviours and socioeconomic indicators to be discussed. Further, HF will attend two qualitative interview and focus group courses and will be supported by his supervisors with extensive experience in generating and analysing qualitative data.

Focus groups will be conducted using Zoom/MS Teams (determined by participant preference), or in-person (If COVID19 restrictions allow and participant chooses this). In-person sessions would take place in a venue local to participants. This will also be the plan for the interviews with primary care professionals and policymakers. Consent will be obtained and recorded from all participants prior to the focus group or interview beginning.

Data Analysis.

FGs and interview audio recordings will be professionally transcribed by a University approved transcription company and analysed in NVivo software using a coding framework underpinned by NPT. Transcribed files will be uploaded to secure UoG

servers. Transcripts will be reviewed against audio files by HF to check for inaccuracies. HF, with support from COD and FM will ensure the transcript is anonymised before analysis. Once transcripts have been reviewed, the audio file will be destroyed.

The process of coding will be informed by the Framework approach.[469,470] This is a five-step process: (1) Familiarisation through reading of the transcripts, (2) Developing a thematic framework, (3) Indexing by applying the framework to all the transcripts, (4) Charting of data across the transcripts, (5) Mapping and interpretation of the data, and comparing across cases.

Initial analysis will be thematic,[471] to identify the main themes and sub-themes in the data. Analysis will also draw on NPT, to explore participants understanding of the interactions between lifestyle factors and deprivation and the work that would be required to develop and implement potential interventions. Data will be coded to identify key barriers and facilitators, while paying attention to codes and themes within the data. This will include consideration of the work required of individuals and communities to enact lifestyle change in view of wider socioeconomic and structural influences on lifestyle.

Analysis will be led by HF and supported by COD. Coding clinics will be held with the supervisory team to discuss the analysis and interpretation of data.

12. Does the research involve any sensitive or potentially upsetting questions? If so, how will you deal with after care of participants? If the interview could lead to disclosure of illegality, professional conduct or at risk status, describe how you will deal with this and make the process clear in the information sheet and consent forms

We consider there to be a low level of risk to participants of this study. As all recruitment is outwith the NHS there is no connection between the study and any medical care a person might be receiving. Reassurance will be given to potential participants if this is raised. FG participants may experience discomfort if personal or sensitive lifestyle/health related topics are raised during the discussion. These participants will be informed, both in the participant information sheet and again prior to recording a FG, that if they feel upset or distressed by the discussion then they are free to ask for the recording to stop and to leave the study. FG participants could also raise concerning symptoms and, in which case, will be signposted to their GP. If interview participants (professionals) reveal serious professional misconduct, then a member of their management team will be informed – this is explained on the consent form. Also included on the consent form for both FG and interview participants is a sentence explaining that criminal acts that come to light during discussion will be reported to relevant authorities.

13. Where will the research take place?
If you will use software for an online interview, please describe the software and how this will be done.

Both FG and interview participants will be offered a choice of online or face-to-face sessions. Face-to-face FGs or interviews will take place at GPPC, 1 Horselethill Road or

another safe and convenient location such as a community centre. If they chose an online session, they will be offered either MS Teams or Zoom. Once a date and time has been agreed, a link to the online meeting will be emailed to participants. At the start of the FG or interview HF will explain the purpose of the research and group rules (if FG) and the consent process. Participants will be notified when recording has started and consent will be verbally recorded prior to discussion commencing.

14. Will you record audio or video of any interviews?
If so, describe how this will do done, how transcription will be performed and when any recordings will be deleted.

Audio data will be recorded in face-to-face sessions by way of a digital audio recorder. Audio data will be recorded from online sessions via the record function on Zoom/MS Teams. After completion of each FG/interview, audio data only will be saved to a secure UoG server and audio data deleted from the audio recorder at the earliest opportunity (done by HF). Any automatically generated video data files from Zoom/MS teams will be deleted immediately. Audio files will be converted to ZIP files and sent via the University's encrypted server to a professional transcription service (Smallbiz) to be converted to Microsoft Word files (organised by HF).

Transcriptions will be depersonalised and linked only to participant identifiable information through a unique study code (done by HF and supervisors COD and FM). Once anonymised transcription files have been checked for accuracy against the corresponding audio file then the audio file will be deleted.

15. Will you obtain and store any personal data? Please detail the information security measures that will be in place.

The research team (HF, COD, FM) will have access to all data collected and no other researchers will have access to personal data. Data will be managed and stored in accordance with UoG policy. Participant identifiable data (e.g., contact details, consent forms), FG/interview audio files, and transcripts will be stored in separate locations on secure university servers only accessible by the research team. All related paper documentation will be stored in locked offices and in a locked cabinet in a secure UoG building (GPPC, House 2 or the Clarice Pears building, home of the Institute of Health & Wellbeing from February 2023).

Transcriptions will be depersonalised and linked only to participant identifiable information through a unique study code. Audio files on encrypted digital recorders will be password protected and deleted at the earliest opportunity. Care will be taken if travelling with the audio recorder. Audio files uploaded to UoG servers will be deleted at the earliest opportunity - as soon as they have been transcribed and transcriptions checked for accuracy. Care will be taken during the analysis and reporting of the data that no individual will be identifiable.

At the end of the study period (March 2023) identifiable participant data (names, email addresses, consent forms) will be stored for a period of 10 years, at which point it will

be deleted. Fully anonymised transcript data will be stored for future ethically approved research by permitted researchers for a period of 10 years.

16. Is there a risk that participants could be identified from their responses?

There is extremely low risk that participants could be identified from their responses. Only members of the research team at UoG will have access to the participant's identifiable and focus group/interview data. These data will be managed and stored in accordance with UoG policy. Participant identifiable data (name, email address), audio recording and transcript files will be stored in separate locations on secure university servers. Only HF and his primary supervisor (COD) will have access to the key file that links participants' details to their audio recording. Care will be taken during the analysis and reporting of the data that no individual will be identifiable.

Focus group/interview audio data: FGs and interviews will be recorded using the recording function of Zoom or MS Teams (only audio files will be used; video files will be deleted immediately). If data collection is in-person, FGs and interviews will be recorded using an encrypted digital audio recorder. All participants taking part will be reminded of the confidentiality of the discussion before the start of the group. The audio recorder will be password protected and care will be taken when travelling with the recording device. Audio files will be uploaded to a secure University sever as soon as possible after focus group/interview and deleted from digital recorder. Audio files will be converted to ZIP files and sent via the University's encrypted server to a professional transcription service (Smallbiz) to be converted to Microsoft Word transcription files. Transcription files will be downloaded and saved on a secure University server. Transcriptions files will be depersonalised and checked for accuracy against the audio files, after which the audio files will be deleted. Transcription files will only be linked to participant identifiable information through a unique study identifier. Care will be taken during the analysis and reporting of the data that no individual will be identifiable.

17. Summarise your data management plan, including plans for dissemination of findings.

Data will be collected at three points. The processes will be the same for both groups of participants (members of the public and professionals).

University of Glasgow is the data controller for this work.

Initial expression of interest

1. People interested in participating will contact the research lead (Dr H Foster) by email, Twitter, or Facebook.
2. Contact details will be left – email and/or telephone number.
3. Contact details will be stored securely on password protected Excel Recruitment file until contacted.
4. Excel file will be stored on UoG server.
5. Those who agree to participate will have a unique ID added to their contact details. All further research material e.g., consent forms; interview transcript will use this unique ID.
6. Those who decline to participate will be removed from the Excel Recruitment file.

Consent forms

1. Prior to the focus group or interview participants will compete the consent form and email it back to the researcher. Consent forms will have the ID number assigned in Step 1. If participants do not consent they will leave the study and we will remove their contact details from the above Excel Recruitment file.
2. Before commencing focus group or interview, participants will be led through the consent form and asked to give verbal consent which will be recorded on the Zoom/MS Team interview.
3. All consent forms will be stored electronically in the study site file, in a separate location from the Excel recruitment file.

Focus group/Interview transcripts

Online focus groups and interviews will be recorded on Zoom/MS Teams, whereas those conducted face-to-face will be recorded using an encrypted audio recorder. Audio files will be uploaded to a UoG approved transcribing company by secure file transfer. Companies used will be Smallbiz or 1st Class secretarial Services. Once transcripts are checked for accuracy and anonymised, the audio file will be destroyed.

Participants will be asked for demographic information, such as gender, age, ethnicity, relationship status, and who they live with. However, we will not retain personally identifiable data beyond the interview transcription. The team are experienced in reporting qualitative data and ensuring that people are not personally identifiable.

The processes described here are processes that we have successfully used in several similar projects. The collection of limited personal data, data that the participant freely gives (e.g., they choose to contact us to express an interest in taking part), and storage of data on UoG servers in password protected files ensure privacy and confidentiality.

Anonymised interview transcripts will be shared amongst the research team for coding and analysis. This will be done using shared folder which only principal researchers (Dr Hamish Foster, Professor Catherine O'Donnell, and Professor Frances Mair) have access**.**

Dissemination of findings

Anonymised quotes from focus groups and interviews may be included in reports, manuscripts, and presentations. Reports will be published on websites, shared with third sector partners (ALLIANCE/The Poverty Alliance, GPs in the Deepend), and posted on social media. Manuscripts will be published in academic journals and presentations will be delivered at public engagement events and national and international conferences.

Data storage:

1. **Initial expression of interest**

Data collected:	Name, email and/or telephone number, as preferred by the contact.
Frequency:	Data will be collected once.
Data storage:	Data kept only if participant takes part and will be retained for 10 years Data deleted if person decides not to participate.

	Data stored in password protected Excel Recruitment file and stored on UoG server.
Number of people:	Up to 50 people for 10 years; data from more people will be stored till prospective participants decline to participate and that number depends on people contacting research lead

2. Consent Forms.

Data collected:	Consent form (Including unique study ID, name, and signature of participant)
Frequency:	Data will be collected once per participant
Data storage:	Stored on UoG server in study site file in a different location from Excel recruitment file. Consent forms retained for 10 years
Number of people:	Public participants: Up to 30 individuals Professional participants: Up to 20 individuals.

3. Audio recording data.

Data collected:	Audio data file
Frequency:	Audio data will be collected once each from approximately 3 focus groups and 16 interviews – i.e., approximately 20 audio data files.
Data storage:	Data recorded face-to-face will initially be stored on an encrypted digital recorder and transferred to UoG server and deleted from recorder at the earliest opportunity. Audio data recorded from MS Teams/Zoom and those transferred from the recorder will be stored on UoG server in study site file in a different location from Excel recruitment file. Access will be limited to the lead researcher and supervisors (HF, COD, FM) Stored for approximately 6 months till they have been transcribed to a word document, anonymised, and checked after which they will be deleted.
Number of people:	Public participants: Up to 30 individuals Professional participants: Up to 20 individuals.

4. Transcribed and anonymised interview/Focus group data.

Data collected:	Anonymised word document file
Frequency:	Approximately 20 Word files
Data storage:	Stored on UoG server in study site file in a different location from Excel recruitment file and consent forms. Retained for 10 years
Number of people:	Public participants: Up to 30 individuals Professional participants: Up to 20 individuals.

7. How is the research being funded?

Via a Medical Research Council Clinical Research Training Fellowship (grant number MR/T001585/1) awarded to Dr H Foster.

8. What is the start and end date of the research?

Proposed starting date: 15/06/2022 Expected completion date: 01/11/2022
9. Describe any potential conflicts of interest. Researcher Name: Dr Hamish Foster conflict of interest - No Researcher Name: Prof Kate O'Donnell conflict of interest - No Researcher Name: Prof Frances Mair conflict of interest - No
INCLUDE A COPY OF THE SURVEY, ANY OTHER PARTICIPANT FACING DOCUMENTS AND THE RESEARCHERS' CVs WITH THE APPLICATION.

Confirmation and Signatures

Please initial box to confirm that all relevant research data generated during and after the study will be collected and held in compliance with the General Data Protection Regulation (May 2018).	H
Please initial box to confirm data will be held securely for a period of ten years after the completion of the research project, or for longer if specified by the research funder or sponsor, in accordance with the University's Code of Good Practice in Research.	H
Please initial box to confirm that you have read the University of Glasgow's Data Protection Policy and the University's mandatory online GDPR and Information Security modules have been successfully completed.	HF
Please initial box to confirm appropriate insurance arrangements are in place.	HF

Name _____Dr Hamish Foster_____ Date _____06/05/2022_____
(Proposer of research)
Please type your name on the line above.

For student projects:

I confirm that I have read and contributed to this submission and believe that the methods proposed, and ethical issues discussed are appropriate.

I confirm that the student will have the time and resources to complete this project.

Name _____Prof Catherine O'Donnell_____ Date ___06/05/2022_____
(Supervisor of student)
Please type your name on the line above.

Please upload the completed and signed form, along with other required documents by logging in to the Research Ethics System at - https://frontdoor.spa.gla.ac.uk/login/

Appendix 3: Consent form - focus groups

University of Glasgow | College of Medical, Veterinary & Life Sciences

Centre Number:

Project Number: PhD project 305234

Participant Identification Number for this study:

Title of Project: 'Understanding the risks associated with combinations of lifestyle factors together with social and financial factors'

Name of Researcher(s): Dr. Hamish Foster
(Supervisors Professors Kate O'Donnell and Frances Mair)

CONSENT FORM

	Please initial box
I confirm that I have read and understood the Participant Information Sheet version 2 dated 03/05/2022.	☐
I confirm that I have read and understood the Privacy Notice version 2 dated 03/05/2022.	☐
I have had the opportunity to think about the information and ask questions and I understand the answers I have been given.	☐
I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my legal rights being affected.	☐
I understand that all data and information I provide will be kept confidential and will be seen only by study researchers and regulators whose job it is to check the work of researchers.	☐
I agree that my name, contact details, and data described in the information sheet will be kept for the purposes of this research project. And I understand that my data will be retained for 10 years or until I decide to withdraw from the study.	☐
I agree to audio recording of the focus group in which I take part.	☐
I understand that the recorded focus group will be transcribed word by word and the anonymised transcription stored for up to 10 years in the University of Glasgow's archiving facilities in accordance with Data Protection policies and regulations.	☐
I am willing to conduct the focus group over an online platform, if not able to do so face to face.	☐
I understand that my information and things that I say in an interview may be quoted in reports and articles that are published about the study, but my name or anything else that could tell people who I am will not be revealed.	☐
I agree to be contacted by the research team about taking part in a focus group.	☐

I agree for the focus group discussion data I provide to be anonymously archived in the UK data archive or other approved archiving facilities, and that other researchers can have access to this data only if they have scientific and ethical approval and agree to preserve the confidentiality of this information as set out in this form. ☐

I agree that should significant concerns regarding my mental or physical health arise during my participation in the study that a member of an appropriate clinical team will be immediately informed. ☐

I understand that any criminal acts which come to light because of my participation in this study may have to be reported appropriately to the relevant authorities by the research team. ☐

Name of participant	Date	Signature
Name of Person taking consent (if different from researcher)	Date	Signature
Researcher	Date	Signature

(1 copy for participant; 1 copy for researcher)

Appendix 4: Consent form - interviews

University of Glasgow | College of Medical, Veterinary & Life Sciences

Centre Number:

Project Number: PhD project 305234

Participant Identification Number for this study:

Title of Project: 'Understanding interactions between lifestyle and deprivation to support policy and intervention development.'

Name of Researcher(s):

Dr. Hamish Foster

(Supervisors Professors Kate O'Donnell and Frances Mair)

CONSENT FORM

	Please initial box
I confirm that I have read and understood the Participant Information Sheet version 2 dated 03/05/2022.	☐
I confirm that I have read and understood the Privacy Notice version 2 dated 03/05/2022.	☐
I have had the opportunity to think about the information and ask questions and I understand the answers I have been given.	☐
I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my legal rights being affected.	☐
I understand that all data and information I provide will be kept confidential and will be seen only by study researchers and regulators whose job it is to check the work of researchers.	☐
I agree that my name, contact details and data described in the information sheet will be kept for the purposes of this research project. And I understand that my data will be stored securely for 10 years o until I withdraw from the study	☐
I agree to my interview being audio-recorded.	☐
I understand that the recorded interview will be transcribed word by word and the anonymised transcription stored for up to 10 years in University archiving facilities in accordance with Data Protection policies and regulations.	☐
I am willing to conduct the interview using telephone or an online platform, if not able to do so face to face.	☐
I understand that my information and things that I say in an interview may be quoted in reports and articles that are published about the study, but my name or anything else that could tell people who I am will not be revealed.	☐

I agree to be contacted by the research team about taking part in an interview. ☐

I agree for the interview discussion data I provide to be anonymously archived in the UK data archive or other approved archiving facilities, and that other researchers can have access to this data only if they have scientific and ethical approval, and agree to preserve the confidentiality of this information as set out in this form. ☐

I agree that should significant concerns regarding my mental or physical health arise during my participation in the study that a member of an appropriate clinical team will be immediately informed. ☐

I agree that should serious concerns regarding my professional conduct arise during my participation in the study that a member of my management team will be informed. ☐

I understand that any criminal acts which come to light as a result of my participation in this study may have to be reported appropriately to the relevant authorities by the research team. ☐

Name of participant	Date	Signature
Name of Person taking consent (if different from researcher)	Date	Signature
Researcher	Date	Signature

(1 copy for participant; 1 copy for researcher)

Appendix 5: Topic guide - focus groups

University of Glasgow | College of Medical, Veterinary & Life Sciences

How do the public and professionals view high risk health behaviour combinations in the context of deprivation? What are the barriers and facilitators to incorporating the new evidence into daily life, practice, and policy to support healthy living in deprived areas?

Introduction

- Please introduce yourself to the group.
- Why have you decided to be part of today's discussion?

Public views of high-risk health behaviour combinations

- Discussion on what is considered to be high risk health behaviour factors and why. E.g., smoking, exercise
- Discussion on having more than one high risk health behaviour and what they are?
- How do these factors affect people's health?
- How does regularly engaging in or having multiple unhealthy behaviours at the same time affect people's health?
- Discuss the possibility that some specific combinations of unhealthy behaviours (e.g., if someone regularly smoked + regularly drank large amounts of alcohol + regularly ate too much processed meat) could have higher risk to people's health than other combinations (e.g., if someone did not sleep very much + did not do much physical activity/exercise + did not eat many fruits or vegetables)
- Which combinations of unhealthy behaviours pose the greatest risk to people's health?

Public views of high-risk lifestyle/behaviour combinations in the context of deprivation

- Discuss how people's wider social and financial circumstances impact on their health? How?
- Discuss how people's wider social and financial circumstances impact on the way in which people live and the combination of lifestyle/health behaviours they might have
- Discuss how the impact of people's social and financial circumstances on their health and the way in which they live/lifestyle is or is not recognised
- Research suggests that combinations of unhealthy health behaviours might have an unequal effect on people's health, where people who live in more disadvantaged areas are affected much more. Discuss what this means for your own views or efforts to live healthily.
- Discuss how you feel this should change how people are supported to live healthy lives in more disadvantaged areas?

Setting the scene for incorporating new evidence into daily life

- Describe how people are currently advised or encouraged to live healthy lifestyles?
- Describe how people are currently supported regarding healthy living/lifestyles?
- How do you think wider social and financial circumstances are taken into account when people are being advised/supported to live healthy lives?

Individual level barriers and facilitators to incorporating new evidence from our work into daily life that could support healthy living in deprived areas.

Considering what you or your family might need to do to live free from high-risk lifestyle combinations:

1. What are the **main barriers** to addressing combinations of health behaviours considering:
    a) very high-risk combinations
    b) additional lifestyle risks felt by people who are more socially/financially disadvantaged

2. What are the main **things that help** (or could help if they were changed) address combinations of health behaviour considering:
    a) very high-risk combinations
    b) additional lifestyle risks felt by people who are more socially/financially disadvantaged

Community level barriers and facilitators to incorporating new evidence from our work into daily life that could support healthy living in deprived areas.

As for 1. a) and b) and 2 a) and b) above, but this time – what are **the current/potential main barriers and facilitators in the neighbourhood/community** needed to live free from high-risk lifestyle combinations?

Society level barriers and facilitators to incorporating new evidence from our work into daily life that could support healthy living in deprived areas.

As for 1. a) and b) and 2 a) and b) above, but this time considering society at large, what would be needed to do to live free from high-risk lifestyle combinations?

Appendix 6: Topic guide - interviews

University of Glasgow | College of Medical, Veterinary & Life Sciences

Interview Topic Guide for Health or Public Health practitioners & Policymakers

How do the public and professionals view high risk health behaviour combinations in the context of deprivation? What are the barriers and facilitators to incorporating the new evidence into daily life, practice, and policy to support healthy living in deprived areas?

Introduction

- Please introduce yourself
- Briefly describe your job/work
- In more detail, can you describe the aspects of your work that involve supporting and advising patients/public to live healthy lives with respect to healthy lifestyle factors/behaviours (e.g., smoking, alcohol, physical activity etc.)?
- Describe if and how that aspect of your work (supporting and advising on healthy living) considers wider social factors?

Professionals' views of high-risk lifestyle/behaviour combinations

- Please explain what you understand of the risks associated with having multiple or a combination of unhealthy lifestyle factors? (prompts - risk of: death, CVD, cancer, mental health problems etc.)
- Describe any combinations of unhealthy lifestyle factors you feel incur the greatest health risks?
- Our work suggests that some specific combinations of lifestyle factors are particularly high risk (e.g., smoking + high processed meat intake + social isolation)
- How does this change your understanding of lifestyle factors and risk?
- Would you, and how would you, incorporate such information into daily work/advice.
- How might patients/public respond to this evidence/information

Professionals' views of high-risk combinations in context of deprivation

- How might the same risk factors, or combinations of factors, affect different groups differently? (Prompts: demographics – age, gender, ethnicity, socioeconomic status)
- How health risks discussed above (combinations of lifestyle/behaviours) are influenced by socioeconomic status/factors?
- Combinations of unhealthy lifestyle factors may have a disproportionate effect on more deprived populations - not just because there are more unhealthy behaviours present but because they are somehow more susceptible. Why would that be?
- How do you feel a disproportionate effect would or should change your work in advising/supporting healthy living?

- How do you feel this would or should change how patients/populations are supported to live healthily in more socioeconomic disadvantaged contexts?

Individual level barriers and facilitators to incorporating new evidence from our work into daily life that could support healthy living in deprived areas.
Considering the work that would be required of patients and populations to live free from high-risk lifestyle combinations:

- Discuss potential barriers that might prevent or hamper patients (practitioners) or populations (public health/policy makers) from addressing combinations of unhealthy lifestyle factors considering:
    a) very high-risk combinations
    b) disproportionate lifestyle risks felt by more deprived groups

- Discuss factors that could help or facilitate patients or populations from addressing combinations of unhealthy lifestyle factors considering:
    a) very high-risk combinations
    b) disproportionate lifestyle risks felt by more deprived groups

Professionals' views of barriers and facilitators to incorporating new evidence from our work into practice and policy to support healthy living in deprived areas?
Considering what would be required to change how you currently work or to change your policy aims:

- Discuss potential barriers that might prevent or hamper changing your work/policy to improve patients' or populations' health outcomes considering:
    a) very high-risk combinations
    b) disproportionate lifestyle risks felt by more deprived groups

- Similarly, please discuss factors that could help or facilitate changing your work/policy to improve patients' or populations' health outcomes considering:
    a) very high-risk combinations
    b) disproportionate lifestyle risks felt by more deprived groups

- Is there anything else you would like to add?

N.B. being a semi-structured interview, participants' answers will determine the interview questions that are asked as follow-ups.

Appendix 7: Supporting information for systematic review (Chapter 4)

Supporting information

The influence of socioeconomic status on the association between unhealthy lifestyle factors and adverse health outcomes: a systematic review.

S1 Table PubMed search strategy

Search	MeSH Terms and keywords	Theme
#1	combination*[tiab] OR combined[tiab] OR composite[tiab] OR integrated[tiab] OR interaction*[tiab] OR joint effect* OR merged effect*[tiab] OR score*[tiab] OR adhere* to[tiab] OR collective[tiab] OR cumulative[tiab] OR multiple[tiab]	combined
#2	life style[MeSH] OR life style*[tiab] OR lifestyle*[tiab] OR risk reduction behavior[MeSH] OR risk reduction behavior*[tiab] OR health behavior[MeSH] OR health behavior[tiab] OR health factor*[tiab] OR low risk*[tiab] OR prevention guideline*[tiab] OR protective factor*[tiab] OR risk reduction behaviour*[tiab] OR health* behaviour*[tiab] OR risk behaviour*[tiab] OR modifiable factors[tiab]	lifestyle factors
#3	healthcare disparities[MeSH] OR healthcare disparities[tiab] OR Health Status Disparities[MeSH] OR disparate[tiab] OR disparit*[tiab] OR inequal*[tiab] OR health inequalities[tiab] OR unequal[tiab] OR health inequities[tiab] OR inequit*[tiab] OR socioeconomic factors[MeSH] OR socioeconomic factors[tiab] OR socio-economic*[tiab] OR socioeconomic*[tiab] OR social-economic[tiab] OR Social Determinants of Health[MeSH] OR social determinant*[tiab] OR poverty[MeSH] OR poverty[tiab] OR depriv*[tiab] OR sociological factors[MeSH] OR sociological factors[tiab] OR social medicine[MeSH] OR social medicine[tiab]	SES
#4	cohort studies[MeSH] OR cohort[tiab] OR incidence[MeSH] OR incidence[tiab] OR survival analysis[MeSH] OR survival[tiab] OR early diagnosis[MeSH] OR early diagnosis[tiab] OR prospective*[tiab] OR follow* up[tiab] OR longitudinal[tiab] OR nested case-control[tiab] OR nested case control[tiab] OR predict*[tiab]	study design
#5	#1 AND #2 AND #3 AND #4	-
#6	Death[MeSH] OR death*[tiab] OR mortality[MeSH] OR mortalit*[tiab] OR fatal*[tiab] OR life expectanc*[tiab] OR surviv*[tiab]	mortality outcome
#7	cardiovascular diseases[MeSH] OR cardiovascular[tiab] OR CVD[tiab] OR heart disease*[tiab] OR myocardial ischaemia[tiab] OR AMI[tiab] OR IHD[tiab] OR CHD[tiab] OR coronary artery disease*[tiab] OR CAD[tiab] OR myocardial infarction[tiab] OR heart infarction[tiab] OR acute coronary syndrome[tiab] OR ACS[tiab] OR heart failure[tiab] OR sudden cardiac death[tiab] OR cerebrovascular disorder*[tiab] OR cerebrovascular accident*[tiab] OR cerebrovascular attack*[tiab] OR CVA[tiab] OR cerebrovascular disease*[tiab] OR CBVD[tiab] OR cerebral arterial disease*[tiab] OR stroke*[tiab] OR apoplex*[tiab]	CVD outcomes
#8	neoplasms[MeSH] OR neoplas*[tiab] OR cancer*[tiab] OR carcinoma*[tiab] OR tumour*[tiab] OR malignanc*[tiab]	cancer outcomes
#9	#6 OR #7 OR #8	all outcomes
#10	#9 AND #5	all terms

MeSH, Medical Subject Heading; tiab, contained in either title or abstract; underlined, both UK and American spellings searched; *any group of letters/characters, including no character;SES, socioeconomic status.

S2 Table EMBASE search strategy

Search	Terms
1	exp lifestyle/
2	exp health behavior/
3	(life style or life styles or health factor or health factors or lifestyle or lifestyles or low risk or low risks or prevention guideline or prevention guidelines or protective factor or protective factors or risk reduction behavior or risk reduction behaviour or risk reduction behaviors or risk reduction behaviours or health behavior or health behaviour or health behaviors or health behaviours or healthy behavior or healthy behaviour or healthy behaviors or healthy behaviours or risk behavior or risk behaviour or risk behaviors or risk behaviours or modifiable factors).ti. or (life style or life styles or health factor or health factors or lifestyle or lifestyles or low risk or low risks or prevention guideline or prevention guidelines or protective factor or protective factors or risk reduction behavior or risk reduction behaviour or risk reduction behaviors or risk reduction behaviours or health behavior or health behaviour or health behaviors or health behaviours or healthy behavior or healthy behaviour or healthy behaviors or healthy behaviours or risk behavior or risk behaviour or risk behaviors or risk behaviours or modifiable factors).ab.
4	1 or 2 or 3
5	(combination or combinations or combined or composite or integrated or interaction or interactions or joint effect or joint effects or merged effect or merged effects or score or scores or adherence to or adhere to or adhered to or collective or cumulative or multiple).ti. or (combination or combinations or combined or composite or integrated or interaction or interactions or joint effect or joint effects or merged effect or merged effects or score or scores or adherence to or adhere to or adhered to or collective or cumulative or multiple).ab.
6	exp social determinants of health/
7	exp socioeconomics/
8	exp social stratification/
9	exp social status/
10	exp health care disparity/
11	exp social medicine/
12	(healthcare disparities or disparate or disparit* or inequal* or health inequalities or unequal or health inequities or inequit* or socioeconomic factors or socio-economic or socioeconomic or socioeconomics or socio-economics or social-economic or Social Determinants of Health or social determinant* or poverty or depriv* or sociological factors or social medicine or Health Status Disparities).ti. or (healthcare disparities or disparate or disparit* or inequal* or health inequalities or unequal or health inequities or inequit* or socioeconomic factors or socio-economic or socioeconomic or socioeconomics or socio-economics or social-economic or Social Determinants of Health or social determinant* or poverty or depriv* or sociological factors or social medicine or Health Status Disparities).ab.
13	6 or 7 or 8 or 9 or 10 or 11 or 12
14	exp cohort analysis/
15	exp incidence/
16	exp survival analysis/
17	exp early diagnosis/
18	(prospective or prospectively or cohort or follow up or followed up or longitudinal or nested case-control or nested case control or incidence or early diagnosis or predict or predicts or predicted or predicting or prediction or predictions or predictive or survival).ti. or (prospective or prospectively or cohort or follow up or followed up or longitudinal or nested case-control or nested case control or incidence or early diagnosis or predict or predicts or predicted or predicting or prediction or predictions or predictive or survival).ab.
19	14 or 15 or 16 or 17 or 18
20	exp death/
21	exp mortality/
22	exp survival/

23	(death or deaths or fatal or fatality or fatalities or life expectancy or life expectancies or mortality or mortalities or survival or survive or survived or survives or surviving).ti. or (death or deaths or fatal or fatality or fatalities or life expectancy or life expectancies or mortality or mortalities or survival or survive or survived or survives or surviving).ab.
24	20 or 21 or 22 or 23
25	exp cardiovascular disease/
26	(cardiovascular or CVD or heart disease or heart diseases or myocardial ischemia or myocardial ischaemia or AMI or IHD or CHD or coronary artery disease or coronary artery diseases or CAD or myocardial infarction or heart infarction or acute coronary syndrome or ACS or heart failure or sudden cardiac death or cerebrovascular disorder or cerebrovascular disorders or cerebrovascular accident or cerebrovascular accidents or cerebrovascular attack or cerebrovascular attacks or CVA or cerebrovascular disease or cerebrovascular diseases or CBVD or cerebral arterial disease or cerebral arterial diseases or stroke or strokes or apoplex).ti. or (cardiovascular or CVD or heart disease or heart diseases or myocardial ischemia or myocardial ischaemia or AMI or IHD or CHD or coronary artery disease or coronary artery diseases or CAD or myocardial infarction or heart infarction or acute coronary syndrome or ACS or heart failure or sudden cardiac death or cerebrovascular disorder or cerebrovascular disorders or cerebrovascular accident or cerebrovascular accidents or cerebrovascular attack or cerebrovascular attacks or CVA or cerebrovascular disease or cerebrovascular diseases or CBVD or cerebral arterial disease or cerebral arterial diseases or stroke or strokes or apoplex).ab.
27	25 or 26
28	exp neoplasm/
29	(neoplasm or neoplasms or neoplasia or cancer or cancers or carcinoma or carcinomas or tumour or tumours or tumor or tumors or malignancy or malignancies).ti. or (neoplasm or neoplasms or neoplasia or cancer or cancers or carcinoma or carcinomas or tumour or tumours or tumor or tumors or malignancy or malignancies).ab.
30	28 or 29
31	24 or 27 or 30
32	4 and 5 and 13 and 19 and 31

exp, explode; ti, Title; ab. Abstract;

S3 Table CINAHL search strategy

No	Terms
S1	TI combination* OR TI combined OR TI composite OR TI integrated OR TI interaction OR TI joint effect* OR TI merged effect* OR TI score* OR TI adhere* OR TI collective OR TI cumulative OR TI multiple
S2	AB combination* OR AB combined OR AB composite OR AB integrated OR AB interaction OR AB joint effect* OR AB merged effect* OR AB score* to OR AB adhere* OR AB collective OR AB cumulative OR AB multiple
S3	S1 OR S2
S4	MH "Life Style+"
S5	TI life style* OR TI lifestyle* OR TI health factor* OR TI low risk* OR TI prevention guideline* OR TI protective factor* OR TI risk reduction behaviour* OR TI risk reduction behavior* OR TI health behaviour* OR TI health behavior* OR TI healthy behaviour* OR TI healthy behavior* OR TI risk behaviour* OR TI risk behaviour OR TI modifiable factor*
S6	AB life style* OR AB lifestyle* OR AB health factor* OR AB low risk* OR AB prevention guideline* OR AB protective factor* OR AB risk reduction behaviour* OR AB risk reduction behavior* OR AB health behaviour* OR AB health behavior* OR AB healthy behaviour* OR AB healthy behavior* OR AB risk behaviour* OR AB risk behaviour OR AB modifiable factor*
S7	S4 OR S5 OR S6
S8	MH “Healthcare Disparities” OR MH “Socioeconomic Factors+” OR MH “Social Determinants of Health” OR MH “Social Environment+” OR MH “Health Status Disparities”
S9	TI healthcare disparities OR TI disparate OR TI disparit* OR TI inequal* OR TI health inequalities OR TI unequal OR TI health inequities OR TI inequit* OR TI socioeconomic factor* OR TI socio-economic* OR TI socioeconomic* OR TI social-economic* OR TI social determinant* OR TI poverty OR TI depriv* OR TI sociological factor OR TI social medicine
S10	AB healthcare disparities OR AB disparate OR AB disparit* OR AB inequal* OR AB health inequalities OR AB unequal OR AB health inequities OR AB inequit* OR AB socioeconomic factor* OR AB socio-economic* OR AB socioeconomic* OR AB social-economic* OR AB social determinant* OR AB poverty OR AB depriv* OR AB sociological factor OR AB social medicine
S11	S8 OR S9 OR S10
S12	MH "Prospective Studies+" OR MH "Incidence" OR MH "Survival Analysis+" OR MH "Early Diagnosis+"
S13	TI prospective OR TI prospectively OR TI cohort OR TI follow up OR TI followed up OR TI longitudinal OR TI nested case-control OR TI nested case control OR TI incidence OR TI early diagnosis OR TI predict OR TI predicts OR TI predicted OR TI predicting OR TI prediction OR TI predictions OR TI predictive OR TI survival
S14	AB prospective OR AB prospectively OR AB cohort OR AB follow up OR AB followed up OR AB longitudinal OR AB nested case-control OR AB nested case control OR AB incidence OR AB early diagnosis OR AB predict OR AB predicts OR AB predicted OR AB predicting OR AB prediction OR AB predictions OR AB predictive OR AB survival
S15	S12 OR S13 OR S14
S16	MH "Death+" OR MH "Mortality+" OR MH "Survival"
S17	TI Death* OR TI fatal* OR TI life expectanc* OR TI mortalit* OR TI surviv*
S18	AB Death* OR AB fatal* OR AB life expectanc* OR AB mortalit* OR AB surviv*
S19	S16 OR S17 OR S18
S20	MH "Cardiovascular Diseases+"
S21	TI cardiovascular OR TI CVD OR TI heart disease* OR TI myocardial isch* OR TI AMI OR TI IHD OR TI CHD OR TI coronary artery disease* OR TI CAD OR TI infarction OR TI acute coronary syndrome OR TI ACS OR TI heart failure OR TI sudden cardiac death OR TI cerebrovascular OR TI CBVD OR TI cerebral arterial disease* OR TI stroke* OR TI strokes OR TI apoplexy*
S22	AB cardiovascular OR AB CVD OR AB heart disease* OR AB myocardial isch* OR AB AMI OR AB IHD OR AB CHD OR AB coronary artery disease* OR AB CAD OR AB infarction OR AB acute coronary syndrome OR AB ACS OR AB heart failure OR AB sudden cardiac death OR AB

	cerebrovascular OR AB CBVD OR AB cerebral arterial disease* OR AB stroke* OR AB strokes OR AB apoplexy*
S23	S20 OR S21 OR S22
S24	MH "Neoplasms+"
S25	TI neoplas* OR TI cancer* OR TI carcinoma* OR TI tumour* OR TI tumor* OR TI malignanc*
S26	AB neoplas* OR AB cancer* OR AB carcinoma* OR AB tumour* OR AB tumor* OR AB malignanc*
S27	S24 OR S25 OR S26
S28	S19 OR S23 OR S27
S29	S3 AND S7 AND S11 AND S15 AND S28

TI, Title; AB, abstract; MH, Major and Minor headings

S4 Table Data extraction form

Article Identifiers	**Exposures, confounders, and outcomes**
Author, publication year	Lifestyle or metabolic factors*
Journal, citation	SES measures*
Source (database/grey literature)	Confounder variables or covariates included
Study funding	Outcome definition and ascertainment
Study characteristics	Number of participants with missing data
Study aims and objectives	**Analysis characteristics**
Cohort name	Type of analysis (statistical methods)
Country	Sensitivity analysis conducted
Type/setting (general population/occupational cohort etc.)	Methods to deal with missing data
Number of participants	Results
Average age	Effect of SES on association between lifestyle and adverse health
Proportion female	Size of effect
Ethnic composition	Study conclusions
Study duration/follow up (start/end dates)	NOS score and justification

* definitions and when and how obtained/measured; NOS, Newcastle Ottawa Scale for assessing quality.[1]

S5 Table Adapted Newcastle-Ottawa Scale[1]

Category	Criteria	Stars available
Selection	1) Representativeness of the cohort a) truly representative of those about whom the study makes conclusions (adult/general/low-income population) * b) somewhat representative * c) selected group of users e.g. nurses, volunteers d) no description of the derivation of the cohort	1
	2) Ascertainment of exposures (lifestyle factors and SES measures) a) secure record (e.g. health care records) * b) structured interview * c) written self-report d) no description	1
	3) Analysis rigour a) impact of >1 socioeconomic measure assessed * b) impact of 1 socioeconomic measure assessed	1
	4) Demonstration that participants were disease free at start of study a) yes/sensitivity analysis excluding those with prior disease * b) no	1
Comparability	Comparability of exposed vs unexposed on the basis of the design or analysis a) analysis includes adjustment for key variables and maximises the use of available data * b) study offers reasonable justification for adjustment variables *	2
Outcome	1) Assessment of outcome a) independent blind assessment * b) record linkage * c) self-report d) no description	1
	2) Was follow-up long enough for outcomes to occur a) yes (≥6months) * b) no	1
	3) Adequacy of follow up of cohorts a) all participants accounted for and missingness <50%, or assessment of impact of missingness given, or description of those with missing data allows assessment of impact of missingness* b) some participants unaccounted for, or missingness >50%, or no assessment of impact of missingness, or insufficient description of those with missing data prevents assessment of impact of missingness	1

S6 Table Distribution of individual unhealthy lifestyle factors in included studies

Study (N)	**Number** and type of unhealthy LFs	N (%) participants with unhealthy LFs
Andersen et al. 2016 (79,101)	**0-1**	**5,122 (6.5)**
	2	**15,121 (19.1)**
	3	**26,730 (33.8)**
	4	**22,881 (28.9)**
	5	**5,562 (7.0)**
	Current/former smoking	50,639 (64.0)
	Heavy alcohol intake	14,075 (17.8)
	Low physical activity levels	28,771 (36.4)
	Sedentary time >5.75h/day	59,194 (74.8)
	Low diet score	59,326 (75.0)
Eguchi et al. 2017 (42,647)	**0-1**	**6,112 (14.3)**
	2	**9,469 (22.2)**
	3	**10,654 (25.0)**
	4	**8,587 (20.1)**
	5-8	**7,825 (18.3)**
	Smoker	11,142 (26.1)
	Alcohol intake >46.0 g/day	6,039 (14.2)
	<0.5h/day +/- <5h/ week walking +/- sports	12,084 (28.3)
	Sleep <5.5 or >7.4 h/day	18,636 (43.7)
	BMI <21 or >25 kg/m^2	21,075 (49.4)
	Fruit <1/day	14,870 (34.8)
	Fish <1/day	25,457 (59.7)
	Milk <almost every day	23,635 (55.4)
Andersen et al. 2018 (77,896)	Heavy alcohol intake	13,899 (17.8)
	Low physical activity levels	61,755 (79.0)
	Low diet score	58,628 (75.0)
	Higher sedentary time	52,044 (66.6)
Foster et al. 2018 (328,594)	**0**	**20,975 (6.4)**
	1	**66,769 (20.3)**
	2	**111,826 (34.0)**
	3	**121,879 (37.1)**
	4	**90,276 (27.5)**
	5	**46,979 (14.3)**
	6	**17,643 (5.4)**
	7	**4,754 (1.4)**
	8	**698 (0.2)**
	9	**72 (0.02)**

	Current smoking	31,515 (9.6)	
	Daily/almost daily alcohol	54,452 (16.6)	
	Low physical activity levels	142,409 (43.3)	
	TV viewing ≥4 h/day	85,036 (25.9)	
	Short/Long sleepers	81,771 (24.9)	
	Fruit and vegetable intake <400 g/day	229,582 (69.9)	
	Oily fish intake <1 portion/week	145,803 (44.4)	
	Red meat intake >3 portions/week	44,337 (13.5)	
	Processed meat intake >1 portion/week	100,928 (30.7)	
Choi et al. 2020 (189,087)	**0**	**64,857 (34.3)**	
	1	**79,417 (42.0)**	
	2	**1,891 (1.0)**	
	3	**6,429 (3.4)**	
	4	**284 (0.2)**	
	Current smoker/recent quitter	69,962 (37.0)	
	Heavy alcohol drinker	12,669 (6.7)	
	Physically inactive	74,878 (39.6)	
	BMI <18.5 or ≥35	18,909 (10.0)	
Zhang et al. 2021 (NHANES 44,462; UK Biobank 399,537)		NHANES (44,462)	UK biobank (399,537)
	0-1	**16,134 (36.3)**	**74,674 (18.7)**
	2	**17,134 (38.5)**	**151,768 (38.0)**
	3-4	**11,194 (25.2)**	**173,095 (43.3)**
	Former/current smoking	22,009 (49.5)	180,562 (45.2)
	Unhealthy alcohol consumption	3,424 (7.7)	147,637 (37.0)
	Lower 2/3 of leisure time physical activity	29,345 (66.0)	266,355 (66.7)
	Unhealthy diet	25,077 (56.4)	342,645 (85.8)

S7 Table Study quality measured by Newcastle-Ottawa Scale (NOS)

Study	Selection				Comparability		Outcome			NOS
	ROC	AOE	AR	FOD	AKV	JFA	AO	FUL	AFU	score
Andersen 2016	1	1	0	0	1	1	1	1	1	7
Eguchi 2017	1	0	0	1	1	0	1	1	0	5
Andersen 2018	1	1	0	0	1	0	1	1	1	6
Foster 2018	1	0	1	1	1	1	1	1	1	8
Choi 2020	1	1	0	0	1	1	1	1	0	6
Zhang 2021	1	1	1	1	1	1	1	1	1	9

ROC, representativeness of cohort; AOE, ascertainment of exposure; AR, Analysis rigour; FOD, demonstration participants free of disease at study start; AKV, adjustment of key variables; JFA, justification for adjustment; AO, assessment of outcome; FUL, follow-up length; AFU, adequacy of follow up

S8 Tables Study estimates for combinations of unhealthy compared with healthy LFs in the most and least advantaged SES groups (main comparator)

A. All-cause mortality – Main analyses

Study	**Andersen 2016**				**Eguchi 2017**			**Andersen 2018**		**Foster 2018**	**Choi 2020**	**Zhang 2021 (2 cohorts)**	
SES	Income				Education			NDI		TDI	Income	latent class	
Categories	</≥$15,000 p.a.				</≥16 years old at last formal education			quartiles		quintiles	</≥200% FPL	3 categories	
Population	A.A. women	A.A. men	W. women	W. men	women	men	total	women	men	total	total	NHANES	UKB
Most advantaged	2.08 (1.32-3.33)	**3.70 (2.22-6.25)**	2.86 (1.28 - 6.67)	**4.17 (2.27-7.69)**	**1.82 (1.49-2.22)**	1.82 (1.56-2.17)	**1.79 (1.61-2.00)**	0.68 (0.32-1.45)	**1.43 (0.94-2.16)**	1.65 (1.25-2.19)	**4.39 (CI not reported)**	**1.76 (1.43-2.17)**	1.17 (1.04-1.32)
Least advantaged	**2.22 (1.67-3.03)**	2.00 (1.39-2.94)	**4.00 (2.22-7.14)**	1.72 (1.00-3.03)	1.49 (1.25-1.79)	**1.89 (1.52-2.38)**	1.67 (1.47-1.89)	**1.59 (1.32-1.91)**	1.42 (1.23-1.65)	**2.47 (2.04-3.00)**	3.45 (CI not reported)	1.30 (1.13-1.50)	**1.80 (1.69-1.93)**

B. All-cause mortality – Sensitivity analyses

Study	**Eguchi 2017**		**Andersen 2018**		**Foster 2018**				
SES	Education		NDI		TDI			Income	Education
Categories	</≥16 years old at last formal education		quartiles	quartiles	quintiles			5 categories	5 categories
Population	total		women	men	total			total	
Analysis	a	b	c	c	d	e	f	g	h
Most advantaged	**1.85 (1.64-2.04)**	**2.13 (1.79-2.50)**	**1.62 (1.18-2.21)**	**1.45 (1.09-1.93)**	1.39 (1.08-1.79) *	1.57 (1.25-1.96) *	1.92 (1.38-2.68) *	1.29 (0.51-3.24) *	1.88 (1.47-2.42) *
Least advantaged	1.64 (1.45-1.82)	1.92 (1.61-2.27)	1.61 (1.42-1.83)	1.30 (1.15-1.48)	**2.00 (1.66-2.40)** *	**2.86 (2.28-3.59)** *	**2.44 (1.91-3.11)** *	**2.36 (1.97-2.83)** *	**2.29 1.90-2.75) ***

C. All-cause mortality – Sensitivity analyses contd.

Study	**Zhang 2021 (2 cohorts)**														
SES	latent class analysis of income, education level, occupation/employment, & health insurance (NHANES only)														
Categories	3 categories														
Population	NHANES							UKB							
Analysis	i	j	k	l	m	n	o	i	j	k	l	n	o	p	q
Most advantaged	**1.74 (1.43-2.11)**	**1.61 (1.27-2.04)**	**1.69 (1.26-2.26)**	**1.82 (1.46-2.28)**	**1.75 (1.42-2.17)**	**1.73 (1.45-2.07)**	1.00 (0.78-1.28)	1.33 (1.19-1.48)	1.32 (1.12-1.55)	1.19 (1.03-1.39)	1.17 (1.03-1.32)	1.31 (1.19-1.45)	1.17 (1.04-1.32)	1.17 (1.04-1.31)	1.31 (1.20-1.43)
Least advantaged	1.25 (1.08-1.43)	1.27 (1.09-1.47)	1.32 (1.06-1.66)	1.23 (1.05-1.45)	1.36 (1.17-1.58)	1.35 (1.20-1.52)	**1.30 (1.13-1.51)**	**1.92 (1.81-2.04)**	**1.81 (1.64-2.00)**	**1.80 (1.63-1.99)**	**1.75 (1.63-1.88)**	**1.79 (1.69-1.89)**	**1.81 (1.69-1.93)**	**1.75 (1.64-1.87)**	**1.75 (1.58-1.92)**

D. CVD mortality – main and sensitivity analyses

Study	Eguchi 2017					Foster 2018						Zhang 2021 (2 cohorts)							
SES	Education					TDI				Income	Education	UK Biobank: Latent class analysis of income, education level, occupation/employment							
Categories	</≥16 years old at last formal education					quintiles				5 categories		3 categories							
Population	women	men	total	total	total	total			f	total		total			k	l	total		
Analysis	main			a	b	main	d	e	f	g	h	main	i	j	k	l	n	o	p
Most advantaged	**2.56 (1.75-3.70)**	2.17 (1.56-3.03)	1.96 (1.92-3.03)	**2.44 (1.96-3.03)**	2.70 (2.00-3.70)	1.93 (1.16-3.20)*	1.49 (0.93-2.40)*	1.44 (0.95-2.19)*	2.26 (1.18-4.33)*	1.66 (0.84-3.27)†*	1.91 (1.20-3.02)*	0.97 (0.74-1.29)	1.16 (0.90-1.51)	1.18 (0.80-1.73)	1.09 (0.76-1.55)	0.93 (0.69-1.25)	1.14 (0.93-1.40)	0.97 (0.74-1.29)	0.97 (0.73-1.28)
Least advantaged	2.27 (1.64-3.23)	**4.00 (2.27-7.14)**	**2.78 (2.13-3.03)**	2.27 (1.82-2.78)	**3.13 (2.27-4.35)**	**3.36 (2.36-4.76)***	**2.11 (1.51-2.93)***	**4.16 (2.55-6.77)***	**3.56 (2.24-5.64)***	**3.23 (2.37-4.41)***	**2.79 (2.01-3.87)***	**1.76 (1.53-2.04)**	**1.91 (1.67-2.18)**	**2.08 (1.63-2.67)**	**1.58 (1.27-1.97)**	**1.64 (1.41-1.91)**	**1.75 (1.55-1.97)**	**1.76 (1.53-2.04)**	**1.71 (1.48-1.97)**

E. CVD incidence – main and sensitivity analyses

Study	Foster 2018						Zhang 2021 (2 cohorts)								
SES	TDI				Income	Education	UK Biobank: Latent class analysis of income, education level, occupation/employment								
Categories	quintiles				5 categories		3 categories								
Population	total			f	total		Total	Total	Total	k	l	Total	Total	Total	Total
Analysis	Main	d	e	f	g	h	Main	i	j	k	l	n	o	p	q
Most advantaged	1.29 (1.10-1.52)	1.15 (0.99-1.33)	1.32 (1.16-1.50)	**1.42 (1.17-1.73)**	1.09 (0.07-1.69)	1.29 (1.11-1.50)	1.15 (0.96-1.38)	1.31 (1.11-1.56)	1.32 (1.04-1.69)	1.15 (0.94-1.39)	1.11 (0.90-1.36)	1.14 (1.01-1.30)	1.15 (0.96-1.37)	1.15 (0.96-1.38)	1.10 (0.95-1.27)
Least advantaged	**1.41 (1.25-1.60)**	**1.24 (1.09-1.40)**	**1.33 (1.18-1.50)**	1.39 (1.18-1.63)	**1.46 (1.29-1.65)**	**1.39 (1.23-1.57)**	**1.58 (1.39-1.79)**	**1.73 (1.54-1.95)**	**1.93 (1.59-2.35)**	**1.52 (1.32-1.74)**	**1.59 (1.37-1.84)**	**1.61 (1.44-1.79)**	**1.57 (1.39-1.79)**	**1.54 (1.36-1.75)**	**1.67 (1.39-2.01)**

F. All other outcomes – main analyses

Outcome	Study	Eguchi 2017			Zhang 2021 (2 cohorts)	
	SES measure	Education			Latent class	
	Categories	</≥16 years old at last formal education			3 categories	
	Population	women	men	total	NHANES	UKB
	Analysis	Main			Main	
CHD mortality	M	2.27 (0.92-5.88)	3.23 (1.47-7.14)	2.70 (1.61-4.55)	-	0.91 (0.62-1.35)
	L	**2.63 (1.22-5.88)**	3.23 (0.99-10.00)	**2.94 (1.56-5.56)**	-	**1.85 (1.51-2.26)**
Heart disease mortality	M	-	-	-	**1.42 (0.89-2.27)**	-
	L	-	-	-	1.32 (0.95-1.81)	-
MI incidence	M	-	-	-	-	1.11 (0.88-1.40)
	L	-	-	-		**1.58 (1.35-1.86)**
Stroke mortality	M	**2.44 (1.35-4.55)**	2.13 (1.30-3.45)	**2.44 (1.69-3.45)**	-	0.73 (0.38-1.43)
	L	2.00 (1.28-3.13)	**3.70 (1.61-8.33)**	2.33 (1.61-3.45)		**2.09 (1.46-2.99)**
Stroke incidence	M	-	-	-	-	1.26 (0.94-1.67)
	L	-	-	-	-	**1.54 (1.27-1.87)**
Cancer mortality	M	**1.28 (1.14-1.82)**	**1.67 (1.28-2.13)**	**2.00 (1.67-2.33)**	-	-
	L	0.63 (0.43-0.91)	1.28 (0.92-1.82)	1.61 (1.32-2.00)	-	-

LFs, lifestyle factors; SES, socioeconomic status; p.a., per annum; CVD, cardiovascular disease; CHD, coronary heart disease; MI, myocardial infarction; M, most advantaged; L, least advantaged; FPL, federal poverty level; MI, myocardial infarction; NHANES, National Health and Nutrition Examination Survey; UKB, UK Biobank; A.A., African American, W., White
Bold font indicates SES group with higher HR. * p-value <0.05 for interaction between SES measure and lifestyle factor combination
Analyses: [a] unhealthy sleep duration extended to <5.5 or >8.4 hours; [b] unhealthy diet extended to include ≤once/day intake of vegetable and bean and unhealthy sleep duration extended to <5.5 or >8.4 hours; [c] never-smokers only; [d] 'traditional' lifestyle factors only i.e. sleep and TV viewing time excluded; [e] only lifestyle factors with significant and independent associations with all-cause and CVD-mortality, and CVD incidence included; [f] participants with hypertension excluded; [g] annual household income as SES measure; [h] maximum education attainment as SES measure; [i] weighted lifestyle score; [j] lifestyle score including body mass index; [k] excluding individuals with prevalent diabetes, CVD, cancer, chronic bronchitis, emphysema, or COPD; [l] excluding events that occurred within the first three years of follow-up; [m] individuals aged 40 years or older; [n] multiple imputation of all missing independent variables; [o] additional inclusion of quadratic terms of age in model; [p] adjusted for area-level SES; [q] by area level SES adjusted for individual-level SES
* no events in most advantaged SES category therefore estimates given for next most advantaged category

S9 Table Study estimates for combined associations between combinations of unhealthy LFs, SES, and all-cause mortality (single reference group analyses*)

Study	Population	Lifestyle score	SES categories (More advantaged → Less advantaged)				
			High education	**Low education**	-	-	-
Eguchi et al. 2017	Women	7-8 (healthy)	1.00 (Ref)	1.14 (0.82-1.59)			
		6	1.11 (0.81-1.73)	1.28 (0.94-1.73)			
		5	1.26 (0.91-1.88)	1.40 (1.03-1.88)			
		4	1.47 (1.06-2.20)	1.63 (1.19-2.20)			
		0-3 (unhealthy)	***1.81 (1.25-2.04)***	1.75 (1.49-2.04)			
	Men	7-8 (healthy)	1.00 (Ref)	1.06 (0.71-1.53)			
		6	1.10 (0.83-1.47)	1.42 (1.03-1.91)			
		5	1.36 (1.03-1.77)	1.54 (1.15-2.00)			
		4	1.68 (1.29-2.16)	1.90 (1.44-2.47)			
		0-3 (unhealthy)	1.84 (1.41-2.35)	***2.00 (1.69-2.33)***			
	Total†	7-8 (healthy)	1.00 (Ref)	1.13			
		6	1.09	1.34			
		5	1.28	1.45			
		4	1.57	1.77			
		0-3 (unhealthy)	1.77	***1.89***			
Andersen et al. 2018			**Quartile 1**	**Quartile 2**	**Quartile 3**	**Quartile 4**	-
	Women	3-4 (healthy)	1.00 (Ref)	1.07 (0.77-1.47)	1.15 (0.85-1.56)	1.25 (0.94-1.65)	
		2	1.17 (0.86-1.60)	1.54 (1.16-2.04)	1.37 (1.04-1.80)	1.62 (1.24-2.12)	
		0-1 (unhealthy)	1.50 (1.12-2.03)	1.65 (1.25-2.18)	1.60 (1.22-2.11)	***1.79 (1.37-2.34)***	
	Men	3-4 (healthy)	1.00 (Ref)	0.92 (0.66-1.28)	0.93 (0.68-1.28)	1.18 (0.90-1.54)	
		2	1.15 (0.86-1.52)	1.21 (0.92-1.59)	1.27 (0.98-1.65)	1.35 (1.05-1.73)	
		0-1 (unhealthy)	1.39 (1.06-1.83)	1.39 (1.07-1.80)	1.41 (1.09-1.83)	***1.43 (1.11-1.84)***	
Foster et al. 2018	Total		**Quintile 1**	**Quintile 2**	**Quintile 3**	**Quintile 4**	**Quintile 5**
		0-2 (healthy)	1.00 (ref)	1.07 (0.94-1.23)	1.11 (0.97-1.28)	1.20 (1.04-1.38)	1.28 (1.10-1.49)
		3-5	1.22 (1.08-1.39)	1.16 (1.02-1.31)	1.29 (1.14-1.46)	1.60 (1.41-1.80)	1.83 (1.62-2.07)
		6-9 (unhealthy)	1.66 (1.25-2.19)	2.02 (1.56-2.61)	2.22 (1.76-2.81)	2.46 (1.97-3.06)	***3.34 (2.79-3.99)***
Zhang et al. 2021			**High SES**	**Medium SES**	**Low SES**		
	NHANES	3-4 (healthy)	1.00 (1.00-1.00)	1.67 (1.46-1.90)	2.45 (2.11-2.84)		
		2	1.33 (1.13-1.57)	2.04 (1.79-2.32)	2.57 (2.17-3.06)		
		0-1 (unhealthy)	1.67 (1.38-2.03)	2.51 (2.19-2.88)	***3.53 (3.01-4.14)***		
	UKB	3-4 (healthy)	1.00 (1.00-1.00)	1.14 (1.02-1.28)	1.45 (1.29-1.63)		
		2	0.99 (0.88-1.12)	1.28 (1.15-1.42)	1.83 (1.65-2.04)		
		0-1 (unhealthy)	1.20 (1.07-1.35)	1.63 (1.47-1.81)	***2.65 (2.39-2.94)***		

*all groups compared to the reference group of those with the healthiest combination of LFs from the most advantaged SES category. Results were transformed to match this where studies reported the reference group as those with the least healthy combination of LFs that were also in the least advantaged SES category. Sensitivity analyses from Foster et al. 2018 and Zhang et al. 2021 are not shown but results were the same, whereby the highest HRs were seen in the unhealthiest LF-least advantaged group.
†Confidence intervals not reported
Bold italics indicate the highest HR from each analysis

S10 Table Study estimates for combined associations between combinations of unhealthy LFs, SES, and CVD outcomes (single reference group analyses *)

Study (outcome)	Population	Lifestyle score	SES categories (More advantaged → Less advantaged)				
			High education	**Low education**	-	-	-
Eguchi 2017 (CVD mortality)	Women	7-8 (healthy)	1.00 (Ref)	1.19 (0.63-2.22)			
		6	1.44 (0.82-2.48)	1.92 (1.10-3.26)			
		5	1.58 (0.90-2.74)	2.14 (1.24-3.59)			
		4	1.75 (0.96-3.15)	2.47 (1.43-4.15)			
		0-3 (unhealthy)	2.53 (1.31-4.81)	***2.78 (2.04-3.70)***			
	Men	7-8 (healthy)	1.00 (Ref)	0.58 (0.24-1.42)			
		6	1.14 (0.64-2.03)	1.53 (0.85-2.77)			
		5	1.42 (0.85-2.42)	1.60 (0.93-2.77)		-	
		4	1.72 (1.03-2.90)	1.81 (1.07-3.10)			
		0-3 (unhealthy)	2.14 (1.31-3.58)	***2.33 (1.69-3.23)***			
	Total*	7-8 (healthy)	1.00	1.00			
		6	1.32	1.81			
		5	1.57	1.97			
		4	1.89	2.27			
		0-3 (unhealthy)	2.54	***2.70***			
Foster 2018 (CVD mortality)	Total		**Quintile 1**	**Quintile 2**	**Quintile 3**	**Quintile 4**	**Quintile 5**
		3-4 (healthy)	1.00 (ref)	1.17 (0.89-1.54)	1.16 (0.87-1.53)	1.30 (0.98-1.73)	1.27 (0.93-1.74)
		2	1.26 (0.98-1.62)	1.07 (0.82-1.39)	1.67 (1.31-2.13)	1.88 (1.47-2.40)	2.38 (1.87-3.02)
		0-1 (unhealthy)	2.01 (1.21-3.33)	2.33 (1.45-3.74)	1.47 (0.85-2.53)	2.81 (1.87-4.21)	***4.59 (3.33-6.32)***
Foster 2018 (CVD incidence)	Total		**Quintile 1**	**Quintile 2**	**Quintile 3**	**Quintile 4**	**Quintile 5**
		3-4 (healthy)	1.00 (ref)	1.10 (1.02-1.18)	1.08 (1.00-1.16)	1.14 (1.05-1.23)	1.24 (1.15-1.35)
		2	1.08 (1.01-1.16)	1.11 (1.04-1.19)	1.16 (1.08-1.24)	1.20 (1.12-1.28)	1.36 (1.27-1.45)
		0-1 (unhealthy)	1.30 (1.10-1.53)	1.37 (1.17-1.60)	1.44 (1.24-1.67)	1.42 (1.23-1.64)	***1.75 (1.55-1.97)***
Zhang 2021 (CVD mortality, main analysis)	UKB		**High SES**	**Medium SES**	**Low SES**		
		3-4 (healthy)	1.00 (1.00-1.00)	1.13 (0.87-1.46)	1.52 (1.16-1.99)		
		2	0.86 (0.64-1.15)	1.16 (0.91-1.49)	1.92 (1.50-2.46)		
		0-1 (unhealthy)	1.06 (0.80-1.39)	1.45 (1.14-1.85)	***2.65 (2.09-3.38)***		
Zhang 2021 (CVD incidence, main analysis)	UKB		**High SES**	**Medium SES**	**Low SES**		
		3-4 (healthy)	1.00 (1.00-1.00)	1.25 (1.05-1.48)	1.36 (1.12-1.64)		
		2	0.95 (0.79-1.15)	1.29 (1.10-1.52)	1.54 (1.31-1.83)		
		0-1 (unhealthy)		1.47 (1.25-1.72)	***2.09 (1.78-2.46)***		

* all groups compared to the reference group of those with the healthiest combination of LFs from the most advantaged SES category - results were transformed to match this where studies reported the reference group as those with the least healthy combination of LFs and were also in the low SES category. Sensitivity analyses from Foster et al. 2018 and Zhang et al. 2021 are not shown but results were the same, whereby the highest HRs were seen in the unhealthiest LF-least advantaged group.
*Confidence intervals not reported
Bold italics indicate the highest HR from each analysis

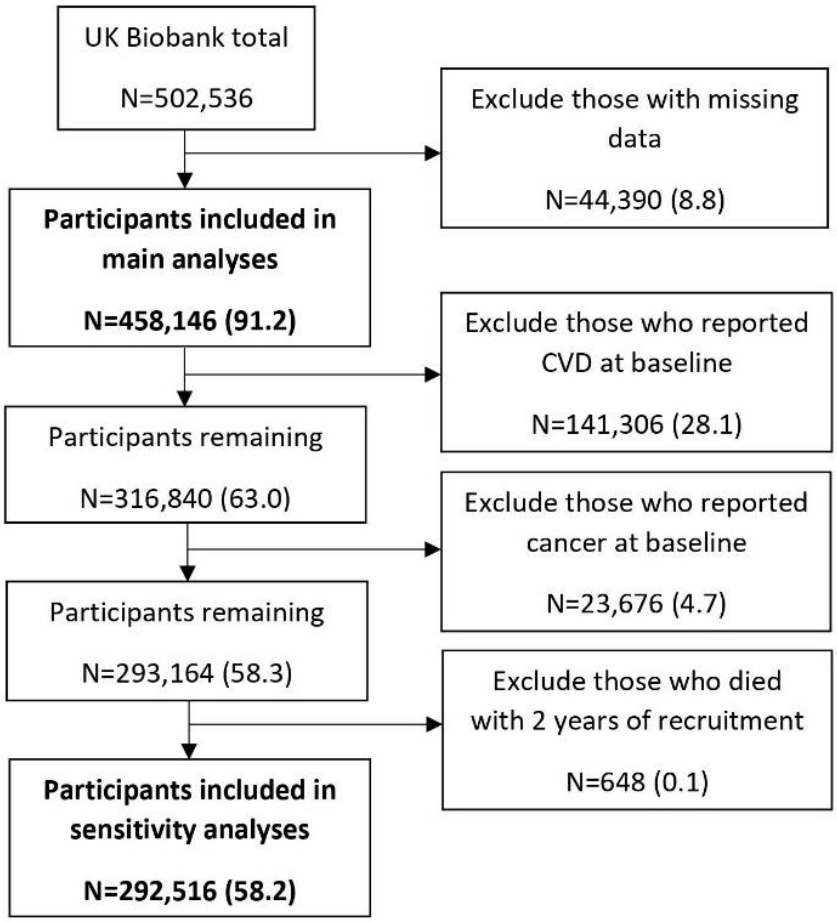
REFERENCES

1. Wells G, Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses: Ottawa Hospital Research Institute [Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp], 2019.

Appendix 8: Social connection analysis - Additional file 1 (Chapter 5)

S1 Figure

Flowchart of study participants (% total UK Biobank)



S1 Table

Table of analyses

	Analysis of association between adverse health outcomes and:	Corresponding research question
1	Frequency of ability to confide in someone close	RQ1
2	Often feeling lonely	RQ1
3	Frequency of ability to confide in someone close and often feeling lonely combined (and their interaction)	RQ1
4	Functional isolation (never able to confide OR often feeling lonely)	RQ1
5	Frequency of friends and family visits	RQ2
6	Weekly group activity	RQ2
7	Living alone	RQ2
8	Frequency of friends and family visits and weekly group activity combined (and their interaction)	RQ2
9	Frequency of friends and family visits and living alone combined (and their interaction)	RQ2
10	Frequency of friends and family visits stratified by living alone	RQ2
11	Weekly group activity and living alone combined (and their interaction)	RQ2
12	Weekly group activity stratified by living alone	RQ2
13	Structural isolation (friends and family visits <monthly OR no weekly group activity OR living alone)	RQ2
14	Frequency of ability confide, often feeling lonely, and structural isolation combined	RQ3
15	Frequency of friends and family visits, weekly group activity, living alone, and functional isolation combined	RQ3
16	Functional and structural isolation (and their interaction)	RQ4
17	Sensitivity analyses: analyses 1-16 repeated with participants with baseline self-reported CVD or cancer and those who died within two years of recruitment excluded	RQ1-4

S2 Table

Comparison of characteristics of participants with missing and complete data.

	Missing	Complete data
N	44,390	458,146
Female	22,641 (51.0%)	250,761 (54.7%)
Mean age (SD)	56.8 (8.27)	56.5 (8.08)
Ethnicity		
White	35,263 (79.4%)	437,462 (95.5%)
Mixed	312 (0.7%)	2,646 (0.6%)
Asian or Asian British	2,951 (6.6%)	6,931 (1.5%)
Black or Black British	1,562 (3.5%)	6,499 (1.4%)
Chinese	426 (1.0%)	1,148 (0.3%)
Other ethnic group	1,099 (2.5%)	3,460 (0.8%)
Missing	2,777 (6.3%)	
Month of assessment		
Jan	2,875 (6.5%)	32,468 (7.1%)
Feb	3,185 (7.2%)	37,992 (8.3%)
Mar	4,298 (9.7%)	45,314 (9.9%)
Apr	4,428 (10.0%)	39,690 (8.7%)
May	5,895 (13.3%)	46,858 (10.2%)
Jun	5,328 (12.0%)	46,677 (10.2%)
Jul	3,578 (8.1%)	38,956 (8.5%)
Aug	3,088 (7.0%)	34,372 (7.5%)

Sep	2,700 (6.1%)	32,942 (7.2%)
Oct	3,348 (7.5%)	38,783 (8.5%)
Nov	3,416 (7.7%)	38,202 (8.3%)
Dec	2,251 (5.1%)	25,892 (5.7%)
Winter assessment	9,015 (20.3%)	218,651 (47.7%)
Townsend index		
Mean (SD)	-0.30 (3.49)	-1.39 (3.04)
Missing	623 (1.4%)	
Smoking		
Current	5,745 (12.9%)	47,234 (10.3%)
Missing	2,950 (6.6%)	
Alcohol intake		
High	2,988 (6.7%)	41,125 (9.0%)
Missing	1,502 (3.4%)	
Physical activity level		
Low	15,082 (34.0%)	89,942 (19.6%)
Missing	5,041 (11.4%)	
BMI, kg/m²		
Mean (SD)	27.8 (5.04)	27.4 (4.78)
Missing	3,105 (7.0%)	
Mean number of long-term conditions (SD)	1.28 (1.33)	1.20 (1.23)

Figures given are N (column %) unless stated otherwise. Higher Townsend index equates to higher levels of deprivation. High alcohol intake, > 35 (females) and > 50 (males) weekly units of alcohol. Low physical activity, <450 MET minutes per week. BMI, body mass index.

S3 Table

Generalised variance inflation factors (GVIF) for all variables included in Cox models

Variable	GVIF	Df	Adjusted GVIF
Sex	1.08	1	1.04
Ethnicity	1.09	19	1.00
Month of assessment	1.01	11	1.00
Townsend index	1.14	1	1.07
Smoking status	1.05	1	1.03
Alcohol intake	1.03	1	1.01
Physical activity level	1.04	1	1.02
Body mass index	1.10	1	1.05
Number of long-term conditions	1.09	1	1.04
Frequency of ability to confide in someone close	1.16	5	1.01
Often feeling lonely	1.13	1	1.06
Friends and family visit frequency	1.10	5	1.01
Weekly group activity	1.04	1	1.02
Living alone	1.13	1	1.06

GVIF, generalised variance inflation factors. Df, degrees of freedom. Adjusted GVIF, adjusted for degrees of freedom via the formula: GVIF^(1/(2*Df))

S4 Table

Fully adjusted models of the combined association between frequency of ability to confide in someone close, feeling lonely and adverse health outcomes.

Outcome	Often feels lonely	Frequency of ability to confide in someone close	n	Deaths (%)	HR	lci	uci

All-cause mortality	No	Daily	221,955	14,715 (6.6%)	1 (ref)	-	-
		2-4 times a week	34,788	2,097 (6.0%)	0.98	0.94	1.03
		Weekly	36,762	2,492 (6.8%)	1.00	0.96	1.04
		Monthly	17,009	1,159 (6.8%)	0.98	0.93	1.04
		Once every 3 months	18,798	1,368 (7.3%)	0.99	0.94	1.05
		Never	44,919	4,351 (9.7%)	1.08	1.04	1.12
	Yes	Daily	24,896	1,873 (7.5%)	1.07	1.02	1.12
		2-4 times a week	9,479	690 (7.3%)	1.06	0.99	1.15
		Weekly	13,558	1,064 (7.8%)	1.07	1.00	1.14
		Monthly	7,394	607 (8.2%)	1.11	1.03	1.21
		Once every 3 months	6,869	525 (7.6%)	1.03	0.94	1.12
		Never	21,719	2,194 (10.1%)	1.10	1.05	1.15
CVD mortality	No	Daily	221,955	2,135 (1.0%)	1 (ref)	-	-
		2-4 times a week	34,788	284 (0.8%)	0.97	0.86	1.10
		Weekly	36,762	345 (0.9%)	1.00	0.89	1.12
		Monthly	17,009	165 (1.0%)	0.98	0.84	1.15
		Once every 3 months	18,798	211 (1.1%)	1.05	0.91	1.21
		Never	44,919	792 (1.8%)	1.19	1.09	1.29
	Yes	Daily	24,896	290 (1.2%)	1.09	0.96	1.23
		2-4 times a week	9,479	96 (1.0%)	1.00	0.81	1.23
		Weekly	13,558	159 (1.2%)	1.06	0.89	1.25
		Monthly	7,394	107 (1.4%)	1.31	1.07	1.59
		Once every 3 months	6,869	89 (1.3%)	1.12	0.91	1.39
		Never	21,719	439 (2.0%)	1.23	1.10	1.37

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, social isolation measures [frequency of friend and family visits, weekly group activity, living alone]. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S5 Table

Interaction estimates for adverse health outcomes for binary exposures of frequency of ability to confide in someone close (≥Once every 3 months vs never) and often feels lonely (no vs yes).

	Interaction	Estimate
All-cause mortality	Multiplicative scale	0.96 [0.90, 1.02]
	RERI	-0.04 [-0.11, 0.03]
	AP	-0.04 [-0.10, 0.02]
	SI	0.77 [0.50, 1.18]
CVD mortality	Multiplicative scale	0.94 [0.82, 1.08]
	RERI	-0.06 [-0.23, 0.11]
	AP	-0.05 [-0.20, 0.08]
	SI	0.80 [0.42, 1.52]

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, frequency of friend and family visits, weekly group activity, living alone. Estimates given as Hazard ratios [95% confidence intervals]).

RERI, relative excess risk for interaction; AP, attributable portion, SI, synergy index. A RERI or AP of zero means no interaction or perfect additivity. A RERI or AP of greater than zero means positive interaction or more than additivity. A RERI or AP of less than zero means negative interaction or less than additivity. An SI of greater than one means positive interaction or more than additivity. An SI of less than one means negative interaction or less than additivity.

S6 Table

Fully adjusted models of joint association between frequency of friends and family visits, engaging in weekly group activity, and adverse health outcomes.

Outcome	Weekly group activity	Frequency of friends and family visits	N	Deaths (%)	HR	lci	uci
All-cause mortality	Yes	Daily	38,297	3,090 (8.1%)	1 (ref)	-	-
		2-4 times a week	104,838	7,359 (7.0%)	0.98	0.94	1.03
		Weekly	114,374	7,115 (6.2%)	0.99	0.95	1.03
		Monthly	40,183	2,532 (6.3%)	1.05	1.00	1.11
		Once every 3 months	18,646	1,387 (7.4%)	1.14	1.07	1.21
		Never	4,007	564 (14.1%)	1.50	1.37	1.64
	No	Daily	15,284	1,458 (9.5%)	1.17	1.10	1.25
		2-4 times a week	37,043	3,132 (8.5%)	1.14	1.09	1.20
		Weekly	49,346	3,578 (7.3%)	1.12	1.07	1.18
		Monthly	21,201	1,489 (7.0%)	1.14	1.07	1.21
		Once every 3 months	11,380	940 (8.3%)	1.24	1.15	1.34
		Never	3,547	491 (13.8%)	1.49	1.36	1.64
CVD mortality	Yes	Daily	38,297	471 (1.2%)	1 (ref)	-	-
		2-4 times a week	104,838	1,077 (1.0%)	0.96	0.86	1.07
		Weekly	114,374	1,065 (0.9%)	0.96	0.86	1.07
		Monthly	40,183	385 (1.0%)	1.01	0.88	1.15
		Once every 3 months	18,646	243 (1.3%)	1.18	1.01	1.38
		Never	4,007	126 (3.1%)	1.66	1.36	2.02
	No	Daily	15,284	223 (1.5%)	1.15	0.98	1.35
		2-4 times a week	37,043	447 (1.2%)	1.05	0.92	1.20
		Weekly	49,346	549 (1.1%)	1.07	0.94	1.21
		Monthly	21,201	242 (1.1%)	1.10	0.94	1.29
		Once every 3 months	11,380	175 (1.5%)	1.30	1.09	1.55
		Never			1.60	1.29	1.97

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, living alone, and functional isolation. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S7 Table

Interaction estimates for adverse health outcomes for binary exposures of frequency of friends and family visits (≥ monthly/< monthly) and weekly group activity (yes vs no).

	Interaction	Estimate
All-cause mortality	Multiplicative scale	0.94 [0.88, 1.01]
	RERI	-0.05 [-0.14, 0.04]
	AP	-0.04 [-0.11, 0.03]
	SI	0.86 [0.66, 1.13]
CVD mortality	Multiplicative scale	0.96 [0.81, 1.14]
	RERI	-0.02 [-0.24, 0.21]
	AP	-0.01 [-0.19, 0.13
	SI	0.96 [0.58, 1.60]

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, living alone, and functional isolation. Estimates given as Hazard ratios [95% confidence intervals]). RERI, relative excess risk for interaction; AP, attributable portion, SI, synergy index. RERI or AP of zero means no interaction or perfect additivity. A RERI or AP of greater than zero means positive interaction or more than

additivity. A RERI or AP of less than zero means negative interaction or less than additivity. An SI of greater than one means positive interaction or more than additivity. An SI of less than one means negative interaction or less than additivity.

S8 Table

Fully adjusted models of joint association between frequency of friends and family visits, living alone, and adverse health outcomes.

Outcome	Lives alone	Frequency of friends and family visits	N	Deaths (%)	HR	lci	uci
All-cause mortality	No	Daily	40,244	3,043 (7.6%)	1 (ref)	-	-
		2-4 times a week	113,223	7,726 (6.8%)	0.98	0.94	1.02
		Weekly	137,647	8,113 (5.9%)	0.95	0.91	0.99
		Monthly	52,781	3,137 (5.9%)	1.00	0.95	1.05
		Once every 3 months	24,725	1,662 (6.7%)	1.07	1.01	1.14
		Never	5,054	547 (10.8%)	1.33	1.22	1.46
	Yes	Daily	13,337	1,505 (11.3%)	1.19	1.12	1.26
		2-4 times a week	28,658	2,765 (9.6%)	1.16	1.10	1.22
		Weekly	26,073	2,580 (9.9%)	1.26	1.19	1.33
		Monthly	8,603	884 (10.3%)	1.30	1.20	1.40
		Once every 3 months	5,301	665 (12.5%)	1.43	1.31	1.56
		Never	2,500	508 (20.3%)	1.77	1.61	1.95
CVD mortality	No	Daily	40,244	441 (1.1%)	1 (ref)	-	-
		2-4 times a week	113,223	1,089 (1.0%)	0.95	0.85	1.06
		Weekly	137,647	1,165 (0.8%)	0.91	0.81	1.01
		Monthly	52,781	462 (0.9%)	0.95	0.84	1.09
		Once every 3 months	24,725	281 (1.1%)	1.13	0.97	1.32
		Never	5,054	109 (2.2%)	1.49	1.21	1.84
	Yes	Daily	13,337	253 (1.9%)	1.39	1.19	1.63
		2-4 times a week	28,658	435 (1.5%)	1.29	1.13	1.47
		Weekly	26,073	449 (1.7%)	1.45	1.27	1.66
		Monthly	8,603	165 (1.9%)	1.50	1.25	1.80
		Once every 3 months	5,301	137 (2.6%)	1.70	1.40	2.07
		Never	2,500	126 (5.0%)	2.23	1.82	2.73

Adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, weekly group activity, and functional isolation. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S9 Table

Interaction estimates for adverse health outcomes for binary exposures of frequency of friends and family visits (≥ monthly/< monthly) and living alone (yes/no).

	Interaction	Estimate
All-cause mortality	Multiplicative scale	1.11 [1.03, 1.20]
	RERI	0.20 [0.09, 0.31]
	AP	0.13 [0.06, 0.18]
	SI	1.15 [1.22, 1.87]
CVD mortality	Multiplicative scale	1.07 [0.90, 1.27]
	RERI	0.27 [-0.01, 0.57]
	AP	0.13 [-0.01, 0.25]
	SI	1.35 [0.99, 1.86]

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, weekly group activity, and functional isolation. Estimates given as Hazard ratios [95%

confidence intervals]). RERI, relative excess risk for interaction; AP, attributable portion, SI, synergy index. RERI or AP of zero means no interaction or perfect additivity. A RERI or AP of greater than zero means positive interaction or more than additivity. A RERI or AP of less than zero means negative interaction or less than additivity. An SI of greater than one means positive interaction or more than additivity. An SI of less than one means negative interaction or less than additivity.

S10 Table

Fully adjusted models of association between frequency of friends and family visits and adverse health outcomes stratified by living alone.

Outcome	Lives alone	Frequency of friends and family visits	HR	lci	uci
All-cause mortality	No	Daily	1 (ref)	-	-
		2-4 times a week	0.98	0.94	1.02
		Weekly	0.96	0.92	1.00
		Monthly	1.01	0.96	1.07
		Once every 3 months	1.09	1.02	1.16
		Never	1.36	1.24	1.50
	Yes	Daily	1 (ref)	-	-
		2-4 times a week	0.97	0.91	1.03
		Weekly	1.03	0.97	1.10
		Monthly	1.05	0.97	1.15
		Once every 3 months	1.15	1.05	1.26
		Never	1.40	1.26	1.55
CVD mortality	No	Daily	1 (ref)	-	-
		2-4 times a week	0.96	0.85	1.07
		Weekly	0.93	0.83	1.03
		Monthly	0.98	0.86	1.12
		Once every 3 months	1.16	1.00	1.35
		Never	1.53	1.24	1.89
	Yes	Daily	1 (ref)	-	-
		2-4 times a week	0.91	0.78	1.07
		Weekly	1.01	0.86	1.18
		Monthly	1.03	0.85	1.26
		Once every 3 months	1.16	0.94	1.43
		Never	1.52	1.22	1.89

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, weekly group activity, and functional isolation. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S11 Table

Fully adjusted models of joint association between weekly group activity, living alone, and adverse health outcomes.

Outcome	Lives alone	Weekly group activity	N	Deaths (%)	HR	lci	uci
All-cause mortality	No	Yes	260,452	16,238 (6.2)	1 (ref)		
		No	113,222	7,990 (7.1)	1.11	1.08	1.14
	Yes	Yes	59,893	5,809 (9.7)	1.23	1.19	1.26
		No	24,579	3,098 (12.6)	1.46	1.40	1.52

CVD mortality	No	Yes	260,452	2,358 (0.9)	1 (ref)		
		No	113,222	1,189 (1.1)	1.08	1.01	1.16
	Yes	Yes	59,893	1,009 (1.7)	1.45	1.34	1.57
		No	24,579	556 (2.3)	1.66	1.50	1.83

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, frequency of friends and family visits, and functional isolation. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S12 Table

Interaction estimates for adverse health outcomes for binary exposures of weekly group activity (yes/no) and living alone (yes/no).

	Interaction	**Estimate**
All-cause mortality	Multiplicative scale	1.07 [1.02, 1.13]
	RERI	0.12 [0.06, 0.19]
	AP	0.08 [0.04, 0.13]
	SI	1.37 [1.15, 1.63]
CVD mortality	Multiplicative scale	1.05 [0.93, 1.19]
	RERI	0.12 [-0.06, 0.3]
	AP	0.07 [-0.04, 0.17]
	SI	1.23 [0.90, 1.67]

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, frequency of friends and family visits, and functional isolation. Estimates given as Hazard ratios [95% confidence intervals]). RERI, relative excess risk for interaction; AP, attributable portion, SI, synergy index. RERI or AP of zero means no interaction or perfect additivity. A RERI or AP of greater than zero means positive interaction or more than additivity. A RERI or AP of less than zero means negative interaction or less than additivity. An SI of greater than one means positive interaction or more than additivity. An SI of less than one means negative interaction or less than additivity.

S13 Table

Fully adjusted models of association between weekly group activity and adverse health outcomes stratified by living alone.

Outcome	Living alone	Weekly group activity	N	deaths (%)	HR	lci	uci
All-cause mortality	No	Yes	260,452	16,238 (6.2)	1 (ref)		
		No	113,222	7,990 (7.1)	1.11	1.08	1.14
	Yes	Yes	59,893	5,809 (9.7)	1 (ref)		
		No	24,579	3,098 (12.6)	1.19	1.14	1.25
CVD mortality	No	Yes	260,452	2,358 (0.9)	1 (ref)		
		No	113,222	1,189 (1.1)	1.08	1.00	1.16
	Yes	Yes	59,893	1,009 (1.7)	1 (ref)		
		No	24,579	556 (2.3)	1.16	1.04	1.29

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, frequency of friends and family visits, and functional isolation. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S14 Table

Fully adjusted models of joint associations between frequency of ability to confide in someone close, often feeling lonely, and structural isolation (defined as <monthly friends and family visits or not engaging in weekly group activity or living alone), and adverse health outcomes. Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count. HR, hazard ratio; LCI, lower confidence interval; UCI, upper confidence interval.

All-cause mortality

Weekly group activity, living alone, and structural isolation group	Frequency of ability to confide	N	deaths (%)	HR	lci	uci
• not often lonely • no structural isolation	Daily	142,049	8,693 (6.1%)	1.00		
	2-4 times a week	19,876	1,035 (5.2%)	0.98	0.92	1.05
	Weekly	20,142	1,162 (5.8%)	0.99	0.94	1.06
	Monthly	9,460	579 (6.1%)	1.04	0.96	1.13
	Once every 3 months	10,725	667 (6.2%)	0.98	0.90	1.06
	Never	22,971	1,869 (8.1%)	1.07	1.02	1.12
• not often lonely • structural isolation	Daily	79,906	6,022 (7.5%)	1.18	1.14	1.22
	2-4 times a week	14,912	1,062 (7.1%)	1.19	1.12	1.27
	Weekly	16,620	1,330 (8.0%)	1.23	1.16	1.30
	Monthly	7,549	580 (7.7%)	1.14	1.05	1.24
	Once every 3 months	8,073	701 (8.7%)	1.24	1.15	1.34
	Never	21,948	2,482 (11.3%)	1.38	1.32	1.44
• often lonely • no structural isolation	Daily	12,002	752 (6.3%)	1.15	1.06	1.24
	2-4 times a week	3,927	228 (5.8%)	1.23	1.07	1.40
	Weekly	5,331	315 (5.9%)	1.14	1.02	1.28
	Monthly	2,966	170 (5.7%)	1.13	0.97	1.31
	Once every 3 months	2,810	156 (5.6%)	1.05	0.89	1.23
	Never	8,193	612 (7.5%)	1.16	1.07	1.26
• often lonely • structural isolation	Daily	12,894	1,121 (8.7%)	1.27	1.19	1.35
	2-4 times a week	5,552	462 (8.3%)	1.25	1.14	1.38
	Weekly	8,227	749 (9.1%)	1.30	1.21	1.40
	Monthly	4,428	437 (9.9%)	1.40	1.27	1.54
	Once every 3 months	4,059	369 (9.1%)	1.29	1.17	1.44
	Never	13,526	1,582 (11.7%)	1.41	1.34	1.49

CVD mortality

Weekly group activity, living alone, and functional isolation group	Frequency of ability to confide	N	deaths (%)	HR	lci	uci
• not often lonely • no structural isolation	Daily	142,049	1,248 (0.9%)	1.00		
	2-4 times a week	19,876	128 (0.6%)	0.92	0.76	1.10
	Weekly	20,142	151 (0.7%)	0.97	0.82	1.15
	Monthly	9,460	81 (0.9%)	1.09	0.87	1.37
	Once every 3 months	10,725	99 (0.9%)	1.03	0.84	1.27
	Never	22,971	311 (1.4%)	1.15	1.01	1.30
• not often lonely • structural isolation	Daily	79,906	887 (1.1%)	1.20	1.10	1.31
	2-4 times a week	14,912	156 (1.0%)	1.33	1.12	1.57
	Weekly	16,620	194 (1.2%)	1.34	1.15	1.56
	Monthly	7,549	84 (1.1%)	1.17	0.94	1.47
	Once every 3 months	8,073	112 (1.4%)	1.40	1.16	1.70
	Never	21,948	481 (2.2%)	1.67	1.50	1.86
• often lonely • no structural isolation	Daily	12,002	117 (1.0%)	1.23	1.02	1.49
	2-4 times a week	3,927	22 (0.6%)	0.87	0.57	1.33
	Weekly	5,331	41 (0.8%)	1.14	0.83	1.55

	Monthly	2,966	23 (0.8%)	1.15	0.76	1.73
	Once every 3 months	2,810	31 (1.1%)	1.53	1.07	2.19
	Never	8,193	106 (1.3%)	1.31	1.07	1.60
• often lonely • structural isolation	Daily	12,894	173 (1.3%)	1.34	1.14	1.58
	2-4 times a week	5,552	74 (1.3%)	1.46	1.15	1.85
	Weekly	8,227	118 (1.4%)	1.43	1.19	1.74
	Monthly	4,428	84 (1.9%)	1.91	1.53	2.38
	Once every 3 months	4,059	58 (1.4%)	1.39	1.07	1.82
	Never	13,526	333 (2.5%)	1.78	1.57	2.01

S2 Figure

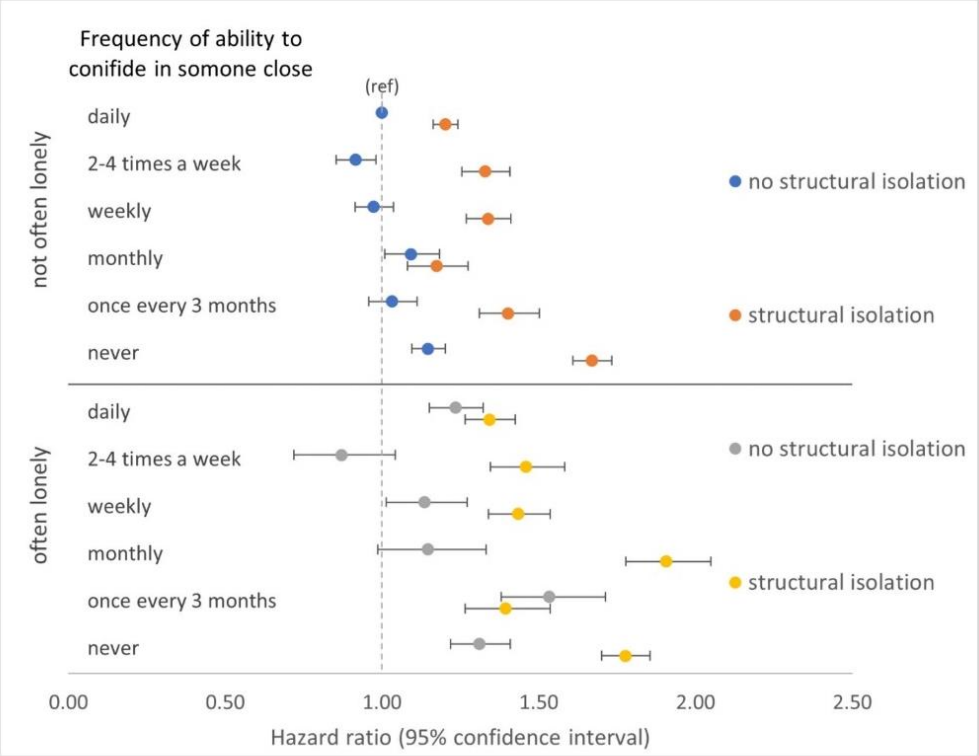


Figure S2 Models of joint association between frequency of ability to confide in someone close, often feeling lonely, structural isolation, and CVD mortality.

S15 Table

Fully adjusted models of joint associations between frequency of friends and family visits, weekly group activity, living alone, functional isolation (defined as either never able to confide in someone close or often feeling lonely), and adverse health outcomes. Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count. HR, hazard ratio; LCI, lower confidence interval; UCI, upper confidence interval.

All-cause mortality

Weekly group activity, living alone, and functional isolation category	Frequency of friends and family visits	N	deaths (%)	HR	lci	uci
• weekly group activity • not living alone • no functional isolation	Daily	23,052	1,626 (7.1%)	1 (ref)	-	-
	2-4 times a week	67,443	4,280 (6.3%)	0.98	0.93	1.04
	Weekly	74,356	4,018 (5.4%)	0.95	0.90	1.01
	Monthly	25,741	1,404 (5.5%)	1.01	0.94	1.09
	Once every 3 months	10,386	660 (6.4%)	1.11	1.01	1.22

	Never	1,274	148 (11.6%)	1.57	1.32	1.85
• weekly group activity • living alone • no functional isolation	Daily	6,167	584 (9.5%)	1.14	1.04	1.25
	2-4 times a week	12,963	1,048 (8.1%)	1.16	1.07	1.25
	Weekly	9,728	791 (8.1%)	1.25	1.15	1.36
	Monthly	2,790	219 (7.8%)	1.24	1.08	1.43
	Once every 3 months	1,362	141 (10.4%)	1.53	1.28	1.81
	Never	279	54 (19.4%)	2.21	1.68	2.90
• weekly group activity • not living alone • functional isolation	Daily	5,492	459 (8.4%)	1.14	1.03	1.26
	2-4 times a week	15,932	1,165 (7.3%)	1.05	0.97	1.13
	Weekly	21,779	1,414 (6.5%)	1.07	0.99	1.14
	Monthly	8,775	586 (6.7%)	1.17	1.06	1.28
	Once every 3 months	4,863	343 (7.1%)	1.21	1.08	1.36
	Never	1,359	135 (9.9%)	1.43	1.20	1.70
• weekly group activity • living alone • functional isolation	Daily	3,586	421 (11.7%)	1.26	1.13	1.41
	2-4 times a week	8,500	866 (10.2%)	1.23	1.13	1.33
	Weekly	8,511	892 (10.5%)	1.37	1.26	1.49
	Monthly	2,877	323 (11.2%)	1.46	1.30	1.65
	Once every 3 months	2,035	243 (11.9%)	1.45	1.27	1.66
	Never	1,095	227 (20.7%)	1.98	1.72	2.27
• no weekly group activity • not living alone • no functional isolation	Daily	8,718	699 (8.0%)	1.14	1.04	1.24
	2-4 times a week	22,718	1,695 (7.5%)	1.14	1.07	1.22
	Weekly	30,465	1,853 (6.1%)	1.07	1.00	1.14
	Monthly	12,859	779 (6.1%)	1.11	1.02	1.21
	Once every 3 months	6,008	395 (6.6%)	1.16	1.04	1.29
	Never	1,021	118 (11.6%)	1.55	1.29	1.87
• no weekly group activity • living alone • no functional isolation	Daily	1,977	249 (12.6%)	1.49	1.31	1.71
	2-4 times a week	3,943	406 (10.3%)	1.40	1.25	1.56
	Weekly	3,892	398 (10.2%)	1.50	1.34	1.67
	Monthly	1,292	138 (10.7%)	1.61	1.35	1.91
	Once every 3 months	708	95 (13.4%)	1.73	1.41	2.13
	Never	170	33 (19.4%)	2.34	1.65	3.30
• no weekly group activity • not living alone • functional isolation	Daily	2,982	259 (8.7%)	1.22	1.07	1.40
	2-4 times a week	7,130	586 (8.2%)	1.17	1.06	1.28
	Weekly	11,047	828 (7.5%)	1.20	1.10	1.31
	Monthly	5,406	368 (6.8%)	1.16	1.03	1.29
	Once every 3 months	3,468	264 (7.6%)	1.27	1.11	1.45
	Never	1,400	146 (10.4%)	1.44	1.21	1.71
• no weekly group activity • living alone • functional isolation	Daily	1,607	251 (15.6%)	1.58	1.38	1.80
	2-4 times a week	3,252	445 (13.7%)	1.49	1.34	1.65
	Weekly	3,942	499 (12.7%)	1.55	1.40	1.71
	Monthly	1,644	204 (12.4%)	1.55	1.34	1.79
	Once every 3 months	1,196	186 (15.6%)	1.82	1.56	2.12
	Never	956	194 (20.3%)	1.99	1.71	2.31

CVD mortality

Weekly group activity, living alone, and functional isolation category	Frequency of friends and family visits	N	deaths (%)	HR	LCI	UCI
• weekly group activity • not living alone • no functional isolation	Daily	23,052	226 (1.0%)	1.00		
	2-4 times a week	67,443	596 (0.9%)	0.98	0.84	1.14
	Weekly	74,356	565 (0.8%)	0.93	0.80	1.09
	Monthly	25,741	190 (0.7%)	0.93	0.77	1.13
	Once every 3 months	10,386	103 (1.0%)	1.14	0.90	1.44
	Never	1,274	27 (2.1%)	1.70	1.14	2.54
	Daily	6,167	95 (1.5%)	1.43	1.13	1.82

• weekly group activity • lives alone • no functional isolation	2-4 times a week	12,963	144 (1.1%)	1.25	1.01	1.54
	Weekly	9,728	119 (1.2%)	1.39	1.11	1.74
	Monthly	2,790	37 (1.3%)	1.48	1.04	2.09
	Once every 3 months	1,362	24 (1.8%)	1.74	1.14	2.65
	Never	279	11 (3.9%)	2.72	1.48	4.99
• weekly group activity • not living alone • functional isolation	Daily	5,492	74 (1.3%)	1.22	0.94	1.58
	2-4 times a week	15,932	191 (1.2%)	1.17	0.97	1.42
	Weekly	21,779	208 (1.0%)	1.05	0.87	1.27
	Monthly	8,775	89 (1.0%)	1.18	0.92	1.50
	Once every 3 months	4,863	61 (1.3%)	1.40	1.05	1.86
	Never	1,359	28 (2.1%)	1.76	1.19	2.60
• weekly group activity • living alone • functional isolation	Daily	3,586	76 (2.1%)	1.56	1.20	2.02
	2-4 times a week	8,500	146 (1.7%)	1.45	1.18	1.79
	Weekly	8,511	173 (2.0%)	1.79	1.46	2.18
	Monthly	2,877	69 (2.4%)	1.98	1.51	2.59
	Once every 3 months	2,035	55 (2.7%)	1.98	1.47	2.66
	Never	1,095	60 (5.5%)	2.80	2.10	3.74
• no weekly group activity • not living alone • no functional isolation	Daily	8,718	96 (1.1%)	1.11	0.87	1.40
	2-4 times a week	22,718	228 (1.0%)	1.07	0.89	1.29
	Weekly	30,465	263 (0.9%)	1.02	0.85	1.22
	Monthly	12,859	113 (0.9%)	1.04	0.83	1.31
	Once every 3 months	6,008	71 (1.2%)	1.32	1.01	1.73
	Never	1,021	30 (2.9%)	2.26	1.54	3.31
• no weekly group activity • living alone • no functional isolation	Daily	1,977	37 (1.9%)	1.72	1.22	2.44
	2-4 times a week	3,943	64 (1.6%)	1.69	1.28	2.23
	Weekly	3,892	61 (1.6%)	1.62	1.22	2.15
	Monthly	1,292	17 (1.3%)	1.34	0.81	2.19
	Once every 3 months	708	18 (2.5%)	2.04	1.26	3.30
	Never	170	5 (2.9%)	2.12	0.87	5.14
• no weekly group activity • not living alone • functional isolation	Daily	2,982	45 (1.5%)	1.45	1.05	2.00
	2-4 times a week	7,130	74 (1.0%)	1.00	0.77	1.30
	Weekly	11,047	129 (1.2%)	1.21	0.97	1.50
	Monthly	5,406	70 (1.3%)	1.40	1.07	1.83
	Once every 3 months	3,468	46 (1.3%)	1.37	0.99	1.88
	Never	1,400	24 (1.7%)	1.36	0.89	2.08
• no weekly group activity • living alone • functional isolation	Daily	1,607	45 (2.8%)	1.83	1.32	2.52
	2-4 times a week	3,252	81 (2.5%)	1.84	1.42	2.37
	Weekly	3,942	96 (2.4%)	1.92	1.51	2.45
	Monthly	1,644	42 (2.6%)	1.92	1.38	2.68
	Once every 3 months	1,196	40 (3.3%)	2.23	1.59	3.12
	Never	956	50 (5.2%)	2.67	1.96	3.65

S3 Figure

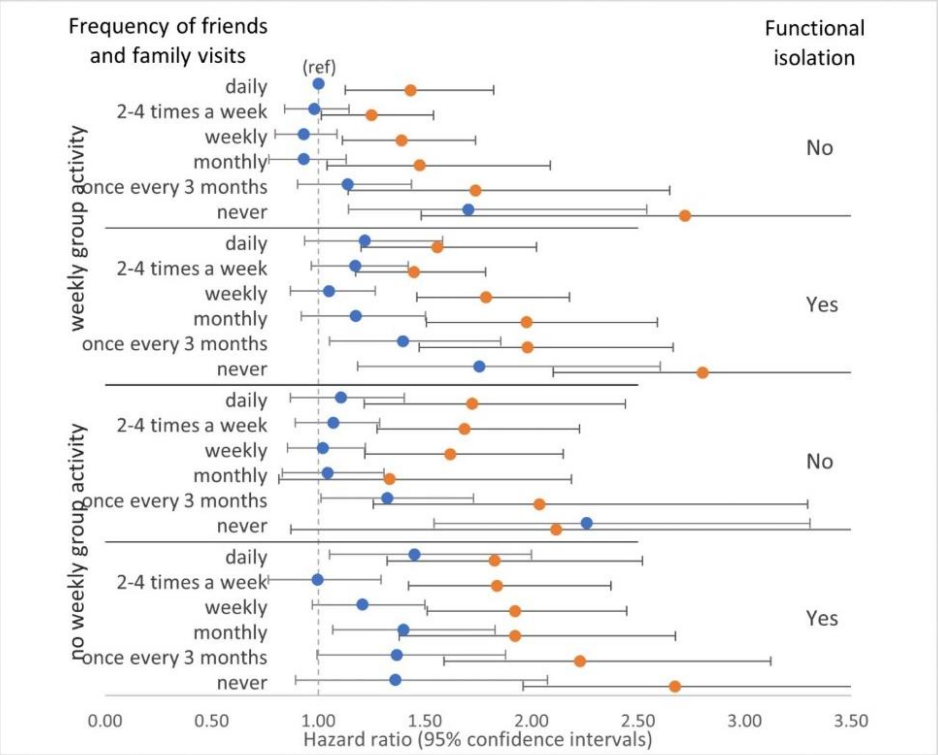


Figure S3 Models of joint association between frequency of friends and family visits, weekly group activity, living alone, functional isolation, and CVD mortality.

S16 Table

Fully adjusted models of joint association between functional and structural isolation and adverse health outcomes.

Outcome	Functional isolation	Structural isolation	N	Deaths (%)	HR	lci	uci
All-cause mortality	No	No	190,592	11,328 (5.9%)	1 (ref)	-	-
	Yes	No	51,978	3,624 (7.0%)	1.11	1.06	1.15
	No	Yes	138,720	10,503 (7.6%)	1.21	1.17	1.24
	Yes	Yes	76,856	7,680 (10.0%)	1.36	1.32	1.40
CVD mortality	No	No	190,592	1,577 (0.8%)	1 (ref)	-	-
	Yes	No	51,978	562 (1.1%)	1.17	1.06	1.29
	No	Yes	138,720	1,563 (1.1%)	1.27	1.18	1.36
	Yes	Yes	76,856	1,410 (1.8%)	1.63	1.51	1.76

S17 Table

Interaction estimates for adverse health outcomes for binary exposures of functional and structural isolation.

	Interaction	Estimate
All-cause mortality	Multiplicative scale	1.02 [0.97, 1.07]
	RERI	0.05 [-0.01, 0.1]
	AP	0.03 [-0.01, 0.07]
	SI	1.15 [0.97, 1.37]

CVD mortality	Multiplicative scale	1.10 [0.97, 1.24]
	RERI	0.19 [0.04, 0.34]
	AP	0.12 [0.02, 0.2]
	SI	1.43 [1.03, 1.99]

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count. Estimates given as Hazard ratios [95% confidence intervals]). RERI, relative excess risk for interaction; AP, attributable portion, SI, synergy index. RERI or AP of zero means no interaction or perfect additivity. A RERI or AP of greater than zero means positive interaction or more than additivity. A RERI or AP of less than zero means negative interaction or less than additivity. An SI of greater than one means positive interaction or more than additivity. An SI of less than one means negative interaction or less than additivity.

Appendix 9: Social connection analysis - Additional file 2 (Chapter 5)

Additional File 2

Sensitivity analyses

Similar results were seen across all sensitivity analyses where we excluded those with prior CVD or cancer or who died within 2 years of recruitment (Tables S18-S35). For example, the associations between functional isolation and all-cause and CVD mortality were 1.12 (1.07-1.16) and 1.18 (1.05-1.32), respectively (Table S22). And compared to those with daily friends and family visits, those who never had friends or family visits hazards for all-cause and CVD mortality were 1.43 (1.27-1.61) and 1.58 (1.17-2.12), respectively (Table S23). A significant interaction on a multiplicative scale between frequency of friends and family visits and living alone for all-cause mortality was observed once again (Table S27) with markedly stronger associations with each of the adverse health outcomes at every level of friend and family visit frequency in those who lived alone (Table S26). For example, all-cause mortality HRs for those who reported never having friends and family visits were 1.20 (1.02-1.42) in those not living alone and 2.09 (1.77-2.47) in those living alone. As per main analyses, results were suggestive of an additive interaction between frequency of friends and family visits and living alone for CVD-mortality, but the estimates had wide confidence intervals (Tables S26 & S27). Also consistent with the main results were the results from sensitivity analyses examining the interaction between weekly group activity and living alone (a multiplicative interaction for all-cause mortality and strongly suggestive of an additive interaction for CVD-mortality) interaction. Results were also consistent with an additive interaction between overall functional and structural isolation (Tables S28 and S29).

S18 Table

Fully adjusted models of association between frequency of ability to confide in someone close and adverse health outcomes. Excluding those with self-reported prior CVD or cancer or who died with 2 years of recruitment. Participants included in analysis n = 292,516. After median follow up of 12.6 years (IQR 11.9-13.3) there were 12,574 (4.3%) deaths including 1,462 (0.5%) CVD deaths).

Outcome	Frequency of ability to confide in someone close	N	Deaths (%)	HR	lci	uci
All-cause mortality	Daily	159,873	6,326 (4.0%)	1 (ref)		
	2-4 times a week	29,207	1,099 (3.8%)	1.00	0.94	1.07
	Weekly	32,436	1,428 (4.4%)	1.05	0.99	1.12
	Monthly	15,586	686 (4.4%)	1.05	0.97	1.13
	Once every 3 months	16,036	706 (4.4%)	0.99	0.91	1.07
	Never	39,378	2,329 (5.9%)	1.11	1.06	1.17
CVD mortality	Daily	159,873	727 (0.5%)	1 (ref)		
	2-4 times a week	29,207	112 (0.4%)	0.91	0.74	1.12
	Weekly	32,436	159 (0.5%)	1.02	0.85	1.22
	Monthly	15,586	70 (0.4%)	0.92	0.71	1.17
	Once every 3 months	16,036	70 (0.4%)	0.81	0.63	1.04
	Never	39,378	324 (0.8%)	1.13	0.98	1.30

Model adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, frequency of friend and family visits, weekly group activity, living alone, and often feeling lonely. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S19 Table

Fully adjusted models of association between often feeling lonely and adverse health outcomes. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

Outcome	Often feels lonely	N	Deaths (%)	HR	lci	uci
All-cause mortality	No	241,292	10,187 (4.2%)	1 (ref)		
	Yes	51,224	2,387 (4.7%)	1.08	1.03	1.13

CVD mortality	No	241,292	1,157 (0.5%)	1 (ref)		
	Yes	51,224	305 (0.6%)	1.17	1.02	1.35

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, frequency of friend and family visits, weekly group activity, living alone, and frequency of ability to confide. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S20 Table

Fully adjusted models of combined associations between frequency of ability to confide in someone close, often feeling lonely and adverse health outcomes. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

Outcome	Often feels lonely	Frequency of ability to confide in someone close	HR	lci	uci
All-cause mortality	No	Daily	1 (ref)	-	-
		2-4 times a week	0.99	0.92	1.06
		Weekly	1.04	0.98	1.12
		Monthly	1.08	0.98	1.18
		Once every 3 months	0.98	0.90	1.08
		Never	1.13	1.07	1.19
	Yes	Daily	1.09	1.01	1.19
		2-4 times a week	1.14	1.00	1.29
		Weekly	1.16	1.05	1.29
		Monthly	1.05	0.91	1.22
		Once every 3 months	1.09	0.94	1.26
		Never	1.16	1.07	1.26
CVD mortality	No	Daily	1 (ref)	-	-
		2-4 times a week	0.88	0.70	1.11
		Weekly	1.03	0.84	1.26
		Monthly	0.93	0.69	1.24
		Once every 3 months	0.81	0.60	1.08
		Never	1.11	0.94	1.30
	Yes	Daily	1.12	0.88	1.43
		2-4 times a week	1.16	0.79	1.69
		Weekly	1.16	0.85	1.57
		Monthly	1.03	0.67	1.59
		Once every 3 months	0.95	0.60	1.49
		Never	1.35	1.09	1.67

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, frequency of friend and family visits, weekly group activity, living alone. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S21 Table

Interaction estimates for adverse health outcomes for binary exposures of never able to confide in someone close and often feeling lonely. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

	Interaction	**Estimate**
All-cause mortality	Multiplicative scale	0.95 [0.86, 1.05]
	RERI	-0.05 [-0.17, 0.07]
	AP	-0.04 [-0.15, 0.05]
	SI	0.80 [0.47, 1.38]
	Multiplicative scale	1.11 [0.84, 1.47]

CVD mortality	RERI	0.17 [-0.18, 0.55]
	AP	0.12 [-0.16, 0.30]
	SI	1.55 [0.60, 1.05]

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, frequency of friend and family visits, weekly group activity, living alone. Estimates given as Hazard ratios [95% confidence intervals]). RERI, relative excess risk for interaction; AP, attributable portion, SI, synergy index. RERI or AP of zero means no interaction or perfect additivity. A RERI or AP of greater than zero means positive interaction or more than additivity. A RERI or AP of less than zero means negative interaction or less than additivity. An SI of greater than one means positive interaction or more than additivity. An SI of less than one means negative interaction or less than additivity.

S22 Table

Fully adjusted models of association between functional isolation and all-cause and CVD mortality. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

Outcome	Measure	N	Deaths (%)	HR	lci	uci
All-cause mortality	*Functional isolation**					
	No	214,342	8,566 (4.0%)	1 (ref)	-	-
	Yes	78,174	4,008 (5.1%)	1.12	1.07	1.16
CVD mortality	*Functional isolation**					
	No	329,312	3,140 (1.0%)	1 (ref)	-	-
	Yes	128,834	1,972 (1.5%)	1.18	1.05	1.32

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, frequency of friend and family visits, weekly group activity, and living alone. * Functional isolation defined as never able to confide in someone close or often feels lonely. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S23 Table

Fully adjusted models of association between structural component measures and all-cause and CVD mortality. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

Outcome	**Measure**	**N**	**Deaths (%)**	**HR**	**lci**	**uci**
All-cause mortality	**Frequency of friends and family visits**					
	Daily	31,417	1,600 (5.1%)	1 (ref)	-	-
	2-4 times a week	88,426	3,913 (4.4%)	0.97	0.91	1.02
	Weekly	107,893	4,183 (3.9%)	0.97	0.92	1.03
	Monthly	41,278	1,652 (4.0%)	1.06	0.98	1.13
	Once every 3 months	19,342	892 (4.6%)	1.14	1.05	1.24
	Never	4,160	334 (8.0%)	1.43	1.27	1.61
	Engages in weekly group activity					
	Yes	206,757	8,608 (4.2%)	1 (ref)	-	-
	No	85,759	3,966 (4.6%)	1.13	1.09	1.17
	Lives alone					
	No	242,725	9,431 (3.9%)	1 (ref)	-	-
	Yes	49,791	3,143 (6.3%)	1.23	1.18	1.28
	Structural isolation*					
	No	159,142	6,024 (3.8%)	1 (ref)		
	Yes	133,374	6,550 (4.9%)	1.15	1.11	1.20
CVD mortality	**Frequency of friends and family visits**					
	Daily	31,417	194 (0.6%)	1 (ref)		
	2-4 times a week	88,426	413 (0.5%)	0.83	0.70	0.99
	Weekly	107,893	486 (0.5%)	0.87	0.74	1.03

	Monthly	41,278	187 (0.5%)	0.89	0.73	1.09
	Once every 3 months	19,342	122 (0.6%)	1.10	0.88	1.39
	Never	4,160	60 (1.4%)	1.58	1.17	2.12
	Engages in weekly group activity					
	No	206,757	982 (0.5%)	1 (ref)		
	Yes	85,759	480 (0.6%)	1.17	1.05	1.31
	Lives alone					
	No	242,725	1,025 (0.4%)	1 (ref)		
	Yes	49,791	437 (0.9%)	1.61	1.43	1.82
	Structural isolation*					
	No	159,142	633 (0.4%)	1 (ref)		
	Yes	133,374	829 (0.6%)	1.34	1.20	1.50

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, new dichotomous functional isolation variable – never able to confide in someone close OR often feeling lonely, and mutually for frequency of friend and family visits, weekly group activity, and living alone. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S24 Table

Fully adjusted models of joint association between frequency of friends and family visits, engaging in weekly group activity, and adverse health outcomes. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

Outcome	Weekly group activity	Frequency of friends and family visits	HR	lci	uci
All-cause mortality	Yes	Daily	1 (ref)		
		2-4 times a week	0.99	0.93	1.07
		Weekly	1.01	0.94	1.08
		Monthly	1.09	1.00	1.19
		Once every 3 months	1.16	1.04	1.29
		Never	1.64	1.40	1.92
	No	Daily	1.24	1.11	1.38
		2-4 times a week	1.12	1.03	1.22
		Weekly	1.12	1.03	1.21
		Monthly	1.22	1.10	1.35
		Once every 3 months	1.34	1.19	1.52
		Never	1.46	1.23	1.74
CVD mortality	Yes	Daily	1 (ref)		
		2-4 times a week	0.82	0.67	1.00
		Weekly	0.81	0.66	0.99
		Monthly	0.92	0.72	1.18
		Once every 3 months	1.06	0.80	1.41
		Never	1.72	1.17	2.51
	No	Daily	1.07	0.78	1.46
		2-4 times a week	0.92	0.72	1.18
		Weekly	1.10	0.88	1.38
		Monthly	0.91	0.67	1.23
		Once every 3 months	1.28	0.93	1.78
		Never	1.58	1.03	2.41

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, living alone, and functional isolation. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S25 Table

Interaction estimates for adverse health outcomes for binary exposures of frequency of friends and family visits (≥ monthly vs < monthly) and weekly group activity (yes vs no). Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

All-cause mortality	Multiplicative scale	0.96 [0.85, 1.09]
	RERI	-0.02 [-0.17, 0.14]
	AP	-0.01 [-0.14, 0.09]
	SI	0.95 [0.62, 1.46]
CVD mortality	Multiplicative scale	0.96 [0.70, 1.32]
	RERI	0.01 [-0.46, 0.50]
	AP	0.00 [-0.36, 0.23]
	SI	1.01 [0.46, 2.21]

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, living alone, and functional isolation. Estimates given as Hazard ratios [95% confidence intervals]). RERI, relative excess risk for interaction; AP, attributable portion, SI, synergy index. RERI or AP of zero means no interaction or perfect additivity. A RERI or AP of greater than zero means positive interaction or more than additivity. A RERI or AP of less than zero means negative interaction or less than additivity. An SI of greater than one means positive interaction or more than additivity. An SI of less than one means negative interaction or less than additivity.

S26 Table

Fully adjusted models of joint association between frequency of friends and family visits, living alone, and adverse health outcomes. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

Outcome	Lives alone	Frequency of friends and family visits	N	Deaths (%)	HR	lci	uci
All-cause mortality	No	Daily	24,157	1,098 (4.5%)	1 (ref)		
		2-4 times a week	71,437	2,933 (4.1%)	0.97	0.91	1.04
		Weekly	92,002	3,250 (3.5%)	0.95	0.89	1.02
		Monthly	35,957	1,313 (3.7%)	1.03	0.95	1.11
		Once every 3 months	16,233	671 (4.1%)	1.13	1.02	1.24
		Never	2,939	166 (5.6%)	1.20	1.02	1.42
	Yes	Daily	7,260	502 (6.9%)	1.17	1.05	1.30
		2-4 times a week	16,989	980 (5.8%)	1.11	1.01	1.21
		Weekly	15,891	933 (5.9%)	1.22	1.11	1.33
		Monthly	5,321	339 (6.4%)	1.35	1.19	1.52
		Once every 3 months	3,109	221 (7.1%)	1.37	1.19	1.59
		Never	1,221	168 (13.8%)	2.09	1.77	2.47
CVD mortality	No	Daily	24,157	121 (0.5%)	1 (ref)		
		2-4 times a week	71,437	297 (0.4%)	0.86	0.70	1.07
		Weekly	92,002	350 (0.4%)	0.85	0.69	1.04
		Monthly	35,957	144 (0.4%)	0.90	0.71	1.15
		Once every 3 months	16,233	88 (0.5%)	1.15	0.87	1.51
		Never	2,939	25 (0.9%)	1.31	0.85	2.02
	Yes	Daily	7,260	73 (1.0%)	1.62	1.21	2.17
		2-4 times a week	16,989	116 (0.7%)	1.22	0.95	1.58
		Weekly	15,891	136 (0.9%)	1.54	1.20	1.97
		Monthly	5,321	43 (0.8%)	1.39	0.98	1.98
		Once every 3 months	3,109	34 (1.1%)	1.63	1.11	2.40
		Never	1,221	35 (2.9%)	2.99	2.03	4.42

Adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, weekly group activity, and functional isolation. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S27 Table

Interaction estimates for adverse health outcomes for binary exposures of friends and family visits less than monthly and living alone. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

	Interaction	**Estimate**
All-cause mortality	Multiplicative scale	1.17 [1.03, 1.33]
	RERI	0.27 [0.09, 0.47]
	AP	0.16 [0.05, 0.25]
	SI	1.72 [1.20, 2.47]
CVD mortality	Multiplicative scale	1.11 [0.80, 1.54]
	RERI	0.44 [-0.16, 1.15]
	AP	0.19 [-0.11, 0.37]
	SI	1.47 [0.87, 2.48]

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, weekly group activity, and functional isolation. Estimates given as Hazard ratios [95% confidence intervals]). RERI, relative excess risk for interaction; AP, attributable portion, SI, synergy index. RERI or AP of zero means no interaction or perfect additivity. A RERI or AP of greater than zero means positive interaction or more than additivity. A RERI or AP of less than zero means negative interaction or less than additivity. An SI of greater than one means positive interaction or more than additivity. An SI of less than one means negative interaction or less than additivity.

S28 Table

Fully adjusted models of association between frequency of friends and family visits and adverse health outcomes stratified by living alone. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

Outcome	Lives alone	Frequency of friends and family visits	HR	lci	uci
All-cause mortality	No	Daily	1 (ref)		
		2-4 times a week	0.97	0.91	1.04
		Weekly	0.96	0.89	1.03
		Monthly	1.04	0.96	1.13
		Once every 3 months	1.14	1.04	1.26
		Never	1.24	1.05	1.47
	Yes	Daily	1 (ref)		
		2-4 times a week	0.94	0.85	1.05
		Weekly	1.01	0.91	1.13
		Monthly	1.11	0.97	1.28
		Once every 3 months	1.12	0.95	1.31
		Never	1.64	1.37	1.97
CVD mortality	No	Daily	1 (ref)		
		2-4 times a week	0.86	0.70	1.07
		Weekly	0.85	0.69	1.05
		Monthly	0.91	0.71	1.17
		Once every 3 months	1.17	0.88	1.54
		Never	1.37	0.89	2.12

	Yes	Daily	1 (ref)		
		2-4 times a week	0.76	0.57	1.02
		Weekly	0.93	0.70	1.24
		Monthly	0.83	0.57	1.22
		Once every 3 months	0.96	0.63	1.45
		Never	1.72	1.13	2.62

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, weekly group activity, and functional isolation. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S29 Table

Fully adjusted models of joint association between weekly group activity, living alone, and adverse health outcomes. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

Outcome	Lives alone	Weekly group activity	N	Deaths (%)	HR	lci	uci
All-cause mortality	No	Yes	170,828	6,531 (3.8)	1 (ref)		
		No	71,897	2,900 (4.0)	1.08	1.04	1.13
	Yes	Yes	35,929	2,077 (5.8)	1.17	1.11	1.23
		No	13,862	1,066 (7.7)	1.49	1.39	1.59
CVD mortality	No	Yes	170,828	700 (0.4)	1 (ref)		
		No	71,897	325 (0.5)	1.11	0.97	1.27
	Yes	Yes	35,929	282 (0.8)	1.52	1.31	1.76
		No	13,862	155 (1.1)	2.03	1.69	2.44

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, frequency of friends and family visits, and functional isolation. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S30 Table

Interaction estimates for adverse health outcomes for binary exposures of weekly group activity (yes/no) and living alone (yes/no). Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

	Interaction	**Estimate**
All-cause mortality	Multiplicative scale	1.18 [1.08, 1.29]
	RERI	0.24 [0.13, 0.35]
	AP	0.16 [0.09, 0.23]
	SI	1.97 [1.39, 2.79]
CVD mortality	Multiplicative scale	1.21 [0.95, 1.53]
	RERI	0.40 [0.02, 0.81]
	AP	0.20 [0.00, 0.34]
	SI	1.64 [1.00, 2.69]

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, frequency of friends and family visits, and functional isolation. Estimates given as Hazard ratios [95% confidence intervals]). RERI, relative excess risk for interaction; AP, attributable portion, SI, synergy index. RERI or AP of zero means no interaction or perfect additivity. A RERI or AP of greater than zero means positive interaction or more than additivity. A RERI or AP of less than zero means negative interaction or less than additivity. An SI of greater than one means positive interaction or more than additivity. An SI of less than one means negative interaction or less than additivity.

S31 Table

Fully adjusted models of association between weekly group activity and adverse health outcomes stratified by living alone. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

Outcome	Living alone	Weekly group activity	N	deaths (%)	HR	lci	uci
All-cause mortality	No	Yes	170,828	6,531 (3.8)	1 (ref)		
		No	71,897	2,900 (4.0)	1.09	1.04	1.14
	Yes	Yes	35,929	2,077 (5.8)	1 (ref)		
		No	13,862	1,066 (7.7)	1.26	1.17	1.36
CVD mortality	No	Yes	170,828	700 (0.4)	1 (ref)		
		No	71,897	325 (0.5)	1.12	0.98	1.28
	Yes	Yes	35,929	282 (0.8)	1 (ref)		
		No	13,862	155 (1.1)	1.31	1.07	1.61

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count, frequency of friends and family visits, and functional isolation. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S32 Table

Fully adjusted models of joint associations between frequency of ability to confide in someone close, often feeling lonely, and structural isolation (defined as <monthly friends and family visits or not engaging in weekly group activity or living alone), and adverse health outcomes. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded. Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count. HR, hazard ratio; LCI, lower confidence interval; UCI, upper confidence interval.

All-cause mortality

Weekly group activity, living alone, and functional isolation group	Frequency of ability to confide	N	deaths (%)	HR	lci	uci
• not often lonely • no structural isolation	Daily	88,438	3,279 (3.7%)	1 (ref)	-	-
	2-4 times a week	12,871	412 (3.2%)	0.99	0.89	1.10
	Weekly	12,654	461 (3.6%)	1.01	0.92	1.11
	Monthly	5,920	251 (4.2%)	1.14	1.00	1.30
	Once every 3 months	6,283	245 (3.9%)	0.98	0.86	1.11
	Never	12,613	644 (5.1%)	1.09	1.00	1.19
• not often lonely • structural isolation	Daily	7,086	248 (3.5%)	1.15	1.01	1.31
	2-4 times a week	2,466	84 (3.4%)	1.28	1.03	1.59
	Weekly	3,232	115 (3.6%)	1.22	1.01	1.47
	Monthly	1,788	56 (3.1%)	1.07	0.82	1.40
	Once every 3 months	1,594	50 (3.1%)	1.00	0.76	1.33
	Never	4,197	179 (4.3%)	1.16	1.00	1.35
• often lonely • no structural isolation	Daily	55,917	2,406 (4.3%)	1.15	1.09	1.21
	2-4 times a week	10,326	432 (4.2%)	1.17	1.05	1.29
	Weekly	11,265	564 (5.0%)	1.27	1.16	1.39
	Monthly	5,099	243 (4.8%)	1.22	1.07	1.39
	Once every 3 months	5,569	273 (4.9%)	1.18	1.04	1.33
	Never	14,337	977 (6.8%)	1.41	1.31	1.51
• often lonely	Daily	8,432	393 (4.7%)	1.27	1.15	1.42
	2-4 times a week	3,544	171 (4.8%)	1.31	1.12	1.53

• structural isolation	Weekly	5,285	288 (5.4%)	1.39	1.23	1.57
	Monthly	2,779	136 (4.9%)	1.27	1.07	1.51
	Once every 3 months	2,590	138 (5.3%)	1.37	1.15	1.63
	Never	8,231	529 (6.4%)	1.48	1.35	1.62

CVD mortality

Weekly group activity, living alone, and functional isolation group	Frequency of ability to confide	N	deaths (%)	HR	lci	uci
• not often lonely • no structural isolation	Daily	88,438	356 (0.4%)	1.00	-	-
	2-4 times a week	12,871	38 (0.3%)	0.91	0.65	1.27
	Weekly	12,654	47 (0.4%)	1.02	0.75	1.38
	Monthly	5,920	26 (0.4%)	1.16	0.78	1.73
	Once every 3 months	6,283	19 (0.3%)	0.72	0.45	1.14
	Never	12,613	66 (0.5%)	0.95	0.73	1.24
• not often lonely • structural isolation	Daily	7,086	29 (0.4%)	1.31	0.90	1.92
	2-4 times a week	2,466	8 (0.3%)	1.31	0.65	2.64
	Weekly	3,232	13 (0.4%)	1.45	0.83	2.53
	Monthly	1,788	5 (0.3%)	1.02	0.42	2.46
	Once every 3 months	1,594	3 (0.2%)	0.62	0.20	1.93
	Never	4,197	23 (0.5%)	1.36	0.89	2.08
• often lonely • no structural isolation	Daily	55,917	296 (0.5%)	1.31	1.12	1.52
	2-4 times a week	10,326	45 (0.4%)	1.22	0.89	1.66
	Weekly	11,265	66 (0.6%)	1.48	1.13	1.93
	Monthly	5,099	22 (0.4%)	1.06	0.69	1.64
	Once every 3 months	5,569	31 (0.6%)	1.24	0.86	1.79
	Never	14,337	145 (1.0%)	1.77	1.45	2.15
• often lonely • structural isolation	Daily	8,432	46 (0.5%)	1.49	1.09	2.04
	2-4 times a week	3,544	21 (0.6%)	1.68	1.08	2.62
	Weekly	5,285	33 (0.6%)	1.65	1.15	2.36
	Monthly	2,779	17 (0.6%)	1.60	0.98	2.62
	Once every 3 months	2,590	17 (0.7%)	1.61	0.99	2.63
	Never	8,231	90 (1.1%)	2.19	1.73	2.77

S33 Table

Fully adjusted models of joint associations between frequency of friends and family visits, weekly group activity, living alone, functional isolation (defined as either never able to confide in someone close or often feeling lonely), and adverse health outcomes. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded. Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count. HR, hazard ratio; LCI, lower confidence interval; UCI, upper confidence interval.

All-cause mortality

Weekly group activity, living alone, and functional isolation category	Frequency of friends and family visits	N	deaths (%)	HR	LCI	UCI
• weekly group activity • not living alone • no functional isolation	Daily	14,161	618 (4.4%)	1 (ref)		
	2-4 times a week	43,448	1,731 (4.0%)	1.00	0.91	1.10
	Weekly	50,687	1,699 (3.4%)	0.96	0.88	1.06
	Monthly	17,870	600 (3.4%)	1.02	0.91	1.15
	Once every 3 months	6,974	286 (4.1%)	1.20	1.04	1.38
	Never	717	47 (6.6%)	1.59	1.18	2.13
• weekly group activity • lives alone • no functional isolation	Daily	3,558	205 (5.8%)	1.08	0.92	1.27
	2-4 times a week	8,061	390 (4.8%)	1.07	0.94	1.21
	Weekly	6,165	292 (4.7%)	1.16	1.01	1.34
	Monthly	1,864	100 (5.4%)	1.35	1.09	1.66

	Once every 3 months	846	50 (5.9%)	1.45	1.08	1.93
	Never	158	21 (13.3%)	2.38	1.54	3.67
• weekly group activity • not living alone • functional isolation	Daily	3,175	150 (4.7%)	1.13	0.94	1.35
	2-4 times a week	9,688	410 (4.2%)	1.04	0.92	1.18
	Weekly	14,210	578 (4.1%)	1.13	1.01	1.26
	Monthly	5,903	238 (4.0%)	1.19	1.02	1.38
	Once every 3 months	3,189	126 (4.0%)	1.17	0.97	1.42
	Never	806	48 (6.0%)	1.51	1.12	2.02
• weekly group activity • living alone • functional isolation	Daily	1,851	128 (6.9%)	1.25	1.04	1.52
	2-4 times a week	4,934	312 (6.3%)	1.25	1.09	1.44
	Weekly	5,085	309 (6.1%)	1.31	1.14	1.50
	Monthly	1,703	121 (7.1%)	1.55	1.28	1.89
	Once every 3 months	1,181	76 (6.4%)	1.32	1.04	1.67
	Never	523	73 (14.0%)	2.32	1.81	2.96
• no weekly group activity • not living alone • no functional isolation	Daily	5,139	249 (4.8%)	1.18	1.01	1.36
	2-4 times a week	14,091	592 (4.2%)	1.09	0.97	1.22
	Weekly	20,215	683 (3.4%)	1.02	0.92	1.14
	Monthly	8,692	316 (3.6%)	1.14	0.99	1.30
	Once every 3 months	3,901	163 (4.2%)	1.28	1.08	1.52
	Never	616	34 (5.5%)	1.25	0.88	1.76
• no weekly group activity • living alone • no functional isolation	Daily	1,079	93 (8.6%)	1.59	1.28	1.98
	2-4 times a week	2,328	139 (6.0%)	1.28	1.06	1.54
	Weekly	2,437	161 (6.6%)	1.55	1.30	1.85
	Monthly	815	51 (6.3%)	1.50	1.13	2.00
	Once every 3 months	439	35 (8.0%)	1.76	1.25	2.48
	Never	81	11 (13.6%)	2.87	1.58	5.22
• no weekly group activity • not living alone • functional isolation	Daily	1,682	81 (4.8%)	1.22	0.97	1.54
	2-4 times a week	4,210	200 (4.8%)	1.19	1.02	1.40
	Weekly	6,890	290 (4.2%)	1.20	1.04	1.37
	Monthly	3,492	159 (4.6%)	1.37	1.15	1.63
	Once every 3 months	2,169	96 (4.4%)	1.36	1.09	1.68
	Never	800	37 (4.6%)	1.14	0.82	1.59
• no weekly group activity • living alone • functional isolation	Daily	772	76 (9.8%)	1.74	1.37	2.20
	2-4 times a week	1,666	139 (8.3%)	1.54	1.28	1.85
	Weekly	2,204	171 (7.8%)	1.65	1.39	1.96
	Monthly	939	67 (7.1%)	1.68	1.30	2.16
	Once every 3 months	643	60 (9.3%)	1.97	1.51	2.57
	Never	459	63 (13.7%)	2.53	1.95	3.28

CVD mortality

Weekly group activity, living alone, and functional isolation category	Frequency of F&FVs	N	deaths (%)	HR	LCI	UCI
• weekly group activity • not living alone • no functional isolation	Daily	14,161	71 (0.5%)	1.00		
	2-4 times a week	43,448	173 (0.4%)	0.84	0.64	1.11
	Weekly	50,687	175 (0.3%)	0.79	0.60	1.05
	Monthly	17,870	67 (0.4%)	0.89	0.64	1.24
	Once every 3 months	6,974	34 (0.5%)	1.08	0.71	1.62
	Never	717	5 (0.7%)	1.20	0.49	2.99
• weekly group activity • lives alone • no functional isolation	Daily	3,558	35 (1.0%)	1.74	1.16	2.62
	2-4 times a week	8,061	42 (0.5%)	1.06	0.73	1.56
	Weekly	6,165	31 (0.5%)	1.07	0.70	1.64
	Monthly	1,864	12 (0.6%)	1.34	0.73	2.48
	Once every 3 months	846	7 (0.8%)	1.59	0.73	3.47

	Never	158	5 (3.2%)	4.18	1.68	10.37
• weekly group activity • not living alone • functional isolation	Daily	3,175	19 (0.6%)	1.18	0.71	1.96
	2-4 times a week	9,688	44 (0.5%)	0.93	0.64	1.35
	Weekly	14,210	59 (0.4%)	0.91	0.64	1.28
	Monthly	5,903	25 (0.4%)	0.98	0.62	1.54
	Once every 3 months	3,189	20 (0.6%)	1.42	0.86	2.33
	Never	806	8 (1.0%)	1.82	0.88	3.79
• weekly group activity • living alone • functional isolation	Daily	1,851	15 (0.8%)	1.29	0.74	2.26
	2-4 times a week	4,934	42 (0.9%)	1.47	1.00	2.15
	Weekly	5,085	46 (0.9%)	1.61	1.11	2.33
	Monthly	1,703	20 (1.2%)	2.00	1.21	3.29
	Once every 3 months	1,181	11 (0.9%)	1.41	0.75	2.67
	Never	523	16 (3.1%)	3.32	1.92	5.75
• no weekly group activity • not living alone • no functional isolation	Daily	5,139	19 (0.4%)	0.80	0.48	1.32
	2-4 times a week	14,091	62 (0.4%)	0.96	0.68	1.35
	Weekly	20,215	82 (0.4%)	0.96	0.70	1.33
	Monthly	8,692	31 (0.4%)	0.84	0.55	1.28
	Once every 3 months	3,901	21 (0.5%)	1.21	0.74	1.97
	Never	616	9 (1.5%)	2.27	1.13	4.56
• no weekly group activity • living alone • no functional isolation	Daily	1,079	13 (1.2%)	2.12	1.17	3.83
	2-4 times a week	2,328	17 (0.7%)	1.46	0.86	2.49
	Weekly	2,437	25 (1.0%)	2.05	1.30	3.24
	Monthly	815	4 (0.5%)	0.96	0.35	2.63
	Once every 3 months	439	6 (1.4%)	2.32	1.01	5.36
	Never	81	0	-	-	-
• no weekly group activity • not living alone • functional isolation	Daily	1,682	12 (0.7%)	1.57	0.85	2.90
	2-4 times a week	4,210	18 (0.4%)	0.92	0.55	1.55
	Weekly	6,890	34 (0.5%)	1.12	0.74	1.68
	Monthly	3,492	21 (0.6%)	1.39	0.85	2.27
	Once every 3 months	2,169	13 (0.6%)	1.39	0.77	2.52
	Never	800	3 (0.4%)	0.66	0.21	2.09
• no weekly group activity • living alone • functional isolation	Daily	772	10 (1.3%)	1.99	1.02	3.87
	2-4 times a week	1,666	15 (0.9%)	1.50	0.86	2.63
	Weekly	2,204	34 (1.5%)	2.73	1.81	4.12
	Monthly	939	7 (0.7%)	1.35	0.62	2.93
	Once every 3 months	643	10 (1.6%)	2.47	1.27	4.81
	Never	459	14 (3.1%)	3.91	2.19	6.98

S34 Table

Fully adjusted models of joint association between functional and structural isolation and adverse health outcomes. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

Outcome	Functional isolation	Structural isolation	N	Deaths (%)	HR	lci	uci
All-cause mortality	No	No	126,166	4,648 (3.7%)	1 (ref)	-	-
	Yes	No	32,976	1,376 (4.2%)	1.12	1.05	1.19
	No	Yes	88,176	3,918 (4.4%)	1.17	1.12	1.22
	Yes	Yes	45,198	2,632 (5.8%)	1.37	1.31	1.44
CVD mortality	No	No	126,166	486 (0.4%)	1 (ref)	-	-
	Yes	No	32,976	147 (0.4%)	1.12	0.93	1.34
	No	Yes	88,176	460 (0.5%)	1.32	1.16	1.50

	Yes	Yes	45,198	369 (0.8%)	1.79	1.56	2.06

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count. HR, hazard ratio; lci, lower confidence interval; uci, upper confidence interval.

S35 Table

Interaction estimates for adverse health outcomes for binary exposures of functional and structural isolation. Those with self-reported prior CVD or cancer or who died with 2 years of recruitment excluded.

	Interaction	**Estimate**
All-cause mortality	Multiplicative scale	1.05 [0.97, 1.14]
	RERI	0.09 [-0.01, 0.18]
	AP	0.06 [-0.01, 0.13]
	SI	1.30 [0.96, 1.77]
CVD mortality	Multiplicative scale	1.22 [0.97, 1.54]
	RERI	0.36 [0.06, 0.65]
	AP	0.20 [0.03, 0.34]
	SI	1.84 [0.98, 3.45]

Models adjusted for sex, ethnicity, Townsend, month of assessment, smoking, alcohol, physical activity, BMI, long-term condition count. Estimates given as Hazard ratios [95% confidence intervals]). RERI, relative excess risk for interaction; AP, attributable portion, SI, synergy index. RERI or AP of zero means no interaction or perfect additivity. A RERI or AP of greater than zero means positive interaction or more than additivity. A RERI or AP of less than zero means negative interaction or less than additivity. An SI of greater than one means positive interaction or more than additivity. An SI of less than one means negative interaction or less than additivity.

Appendix 10: Weighted lifestyle score analysis - Supplementary material 1 (Chapter 6)

Table S1 UK Biobank questionnaire items for lifestyle factors and classification

Lifestyle factor	Questionnaire	'Healthy'	'Unhealthy'
Smoking status	"Do you smoke tobacco now?" and "In the past, how often have you smoked tobacco?"	past or never smoker (non-smoker)	current (smoker)
Alcohol intake	"About how often do you drink alcohol?"	*≤35 [females]/ ≤50 [males] units/week*	*>35 [females]/ >50 [males] units/week*
Physical activity*	IPAQ short form[2]– total time walking or moderate and vigorous-intensity PA in previous week	≥450 MET minutes/ week	<450 MET minutes/ week
TV viewing time ǂ	"In a typical day, how many hours do you spend watching TV?"	< 4 h/day	≥ 4 h/day
Sleep duration	"About how many hours sleep do you get in every 24 hours?"	>7 or <9h/day	<7 or >9h/day
Fruit and vegetable intake	"About how many of …. would you eat per day?" Separate questions for pieces of fresh and dried fruit, tablespoons of salad or cooked/raw vegetables. Combined and converted to g/day (1 portion = 80 g)	≥ 400 g/day	<400 g/ day
Oily fish intake	"How often do you eat oily fish? (e.g. sardines, salmon, mackerel, herring)"	≥1 portion/week	<1 portion/week
Red meat intake	"How often do you eat…?" Separate questions for Beef / lamb or mutton / pork (excluding processed meats such as ham or bacon).	≤3 portion/week	>3 portion/week
Processed meat intake	"How often do you eat processed meats (such as bacon, ham, sausages, meat pies, kebabs, burgers, chicken nuggets)?"	≤1 portion/week	>1 portion/week
Added salt	"Do you add salt to your food? (Do not include salt used in cooking)"	never, rarely, or sometimes	usually or always
Social participation	"How often do you visit friends or family or have them visit you?" **And** "Which of the following [leisure/social activities] do you engage in once a # week or more often?"	Friends and family visit frequency: Almost daily, 2-4 times a week, About once a week, or About once a month **And** Weekly engagement in leisure/social activities	Friends and family visit frequency: Once every few months, Never or almost never, or no friends/family outside household **Or** No weekly engagement in leisure/social activities < weekly

Table S1 shows the relevant UK Biobank questionnaire items and response options used to classify participants into healthy/unhealthy categories for each lifestyle factor. Alcohol cut offs were based on high-risk drinking levels identified for England and in UK Biobank.[1,2] UK Biobank physical activity (PA) data were analysed in accordance with the International Physical Activity Questionnaire (IPAQ)[3] scoring protocol with physical activity. To calculate MET [metabolic equivalent of task] minutes/week, MET weightings (in parentheses) were applied to summed time spent walking (3.3), in moderate (4.0), and in vigorous (8.0) PA. Classification of unhealthy/healthy PA was informed by UK physical activity guielines.[4] TV time classification was based on previous evidence from UK Biobank.[5] UK Biobank dietary information was collected via the Oxford WebQ; a web-based 24-hour recall questionnaire developed specifically for use in large population studies.[6] National dietary guidelines were used for classifying the dietary components.[7,8] Social participation was based on previous UK Biobank work and using only those measures of social connection that are potentially more modifiable by individuals and families.[9]

Table S2 Characteristics of those with missing and those with complete data.

	Missing	**Complete**
N	40,224	462,235
Sex		
Female	20,677 (51.4%)	252,676 (54.7%)
Mean age (SD)	56.8 (8.24)	56.5 (8.08)
Ethnicity		
White	32,053 (79.7%)	440,603 (95.3%)
Mixed	317 (0.8%)	2,639 (0.6%)
Asian or Asian British	2,166 (5.4%)	7,714 (1.7%)
Black or Black British	1,726 (4.3%)	6,334 (1.4%)
Chinese	238 (0.6%)	1,335 (0.3%)
Other	948 (2.4%)	3,610 (0.8%)
Missing	2,776 (6.9%)	
Townsend index		
Mean (SD)	-0.0216 (3.54)	-1.40 (3.03)
Missing	623 (1.5%)	
Summer assessment	22,827 (56.7%)	241,636 (52.3%)
Blood pressure at assessment		
Mean (SD)	135 (18.8)	134 (18.3)
Missing	1,325 (3.3%)	
Body mass index (kg/m^2)		
Mean (SD)	28.3 (5.36)	27.4 (4.75)
Missing	3,104 (7.7%)	
CVD medication		
yes	12,019 (29.9%)	125,410 (27.1%)
Long term condition count		
Mean (SD)	1.38 (1.39)	1.19 (1.23)
Smoking status		
Current smoker	5,874 (14.6%)	47,097 (10.2%)
Missing	2,948 (7.3%)	
Alcohol intake		
High	2,506 (6.2%)	41,604 (9.0%)
Physical activity level		
<450 METminutes/week	16,134 (40.1%)	88,859 (19.2%)
Missing	5,039 (12.5%)	
TV watching time		

≥4 hours	13,371 (33.2%)	13,2131 (28.6%)
Missing	5,413 (13.5%)	
Sleep duration		
<7 or ≥9 hours/day	14,548 (36.2%)	14,7034 (31.8%)
Missing	4,214 (10.5%)	
Fruit and vegetable intake		
<400 g/day	6,038 (15.0%)	89,856 (19.4%)
Missing	15,398 (38.3%)	
Oily fish intake		
<1 portion/week	17,026 (42.3%)	202,681 (43.8%)
Missing	4,079 (10.1%)	
Red meat intake		
>3 portions/week	6,543 (16.3%)	68,740 (14.9%)
Missing	6,063 (15.1%)	
Processed meat intake		
>1 portion/week	12,381 (30.8%)	142,706 (30.9%)
Missing	2,229 (5.5%)	
Added salt		
usually or always	8,200 (20.4%)	74,621 (16.1%)
Missing	1,125 (2.8%)	
Social participation level		
Low	13,690 (34.0%)	161,249 (34.9%)
Missing	9,424 (23.4%)	252,676 (54.7%)

Figure S1 Flow diagram to show numbers and % of all UK Biobank participants included in analyses.

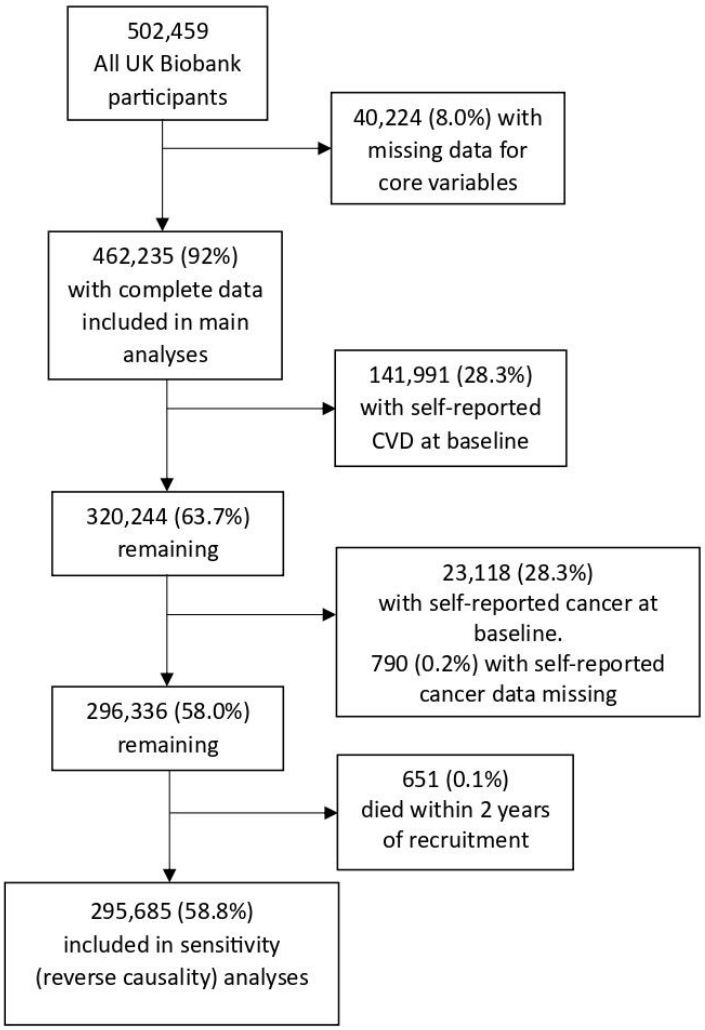


Figure S2 Distribution of weighted lifestyle score among UK Biobank participants

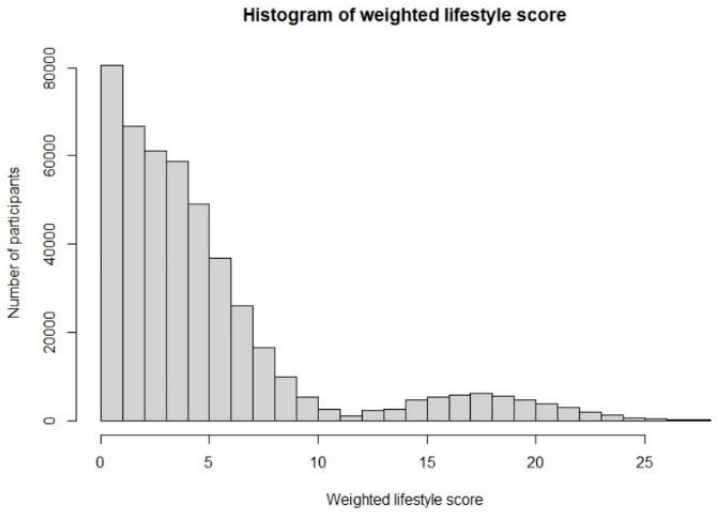


Table S3 (Interaction deprivation) Interaction between binary weighted lifestyle score and binary Townsend index

	All-cause mortality			CVD mortality		
	Weighted lifestyle score		Effect of less healthy stratified by deprivation	Weighted lifestyle score		Effect of less healthy stratified by deprivation
	More healthy (≤9)	Less healthy (≥10)		More healthy (≤9)	Less healthy (≥10)	
	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]
Less deprived	1 [Reference]	1.9 [1.83, 1.97]	1.9 [1.83, 1.97]	1 [Reference]	2.02 [1.84, 2.22]	2.02 [1.84, 2.22]
More deprived	1.23 [1.2, 1.26]	2.86 [2.77, 2.96]	2.33 [2.24, 2.41]	1.28 [1.2, 1.38]	3.34 [3.08, 3.63]	2.6 [2.38, 2.84]
Effect of deprivation stratified by weighted score	1.23 [1.2, 1.26]	1.51 [1.44, 1.58]		1.28 [1.2, 1.38]	1.65 [1.48, 1.85]	
Multiplicative scale	1.22 [1.16, 1.29]			1.29 [1.13, 1.47]		
RERI	0.73 [0.62, 0.84]			1.04 [0.74, 1.34]		
AP	0.26 [0.22, 0.29]			0.31 [0.23, 0.38]		
SI	1.65 [1.52, 1.78]			1.8 [1.5, 2.15]		

RERI, relative excess risk for interaction; AP, attributable portion; SI, synergy index.[10] A RERI or AP of zero can be interpreted as no interaction; greater than zero as positive interaction; and less than zero as negative interaction. SIs greater than one indicate positive interaction and those less than one indicate negative interaction.

Figure S3 Association between weighted lifestyle score and mortality stratified by deprivation.

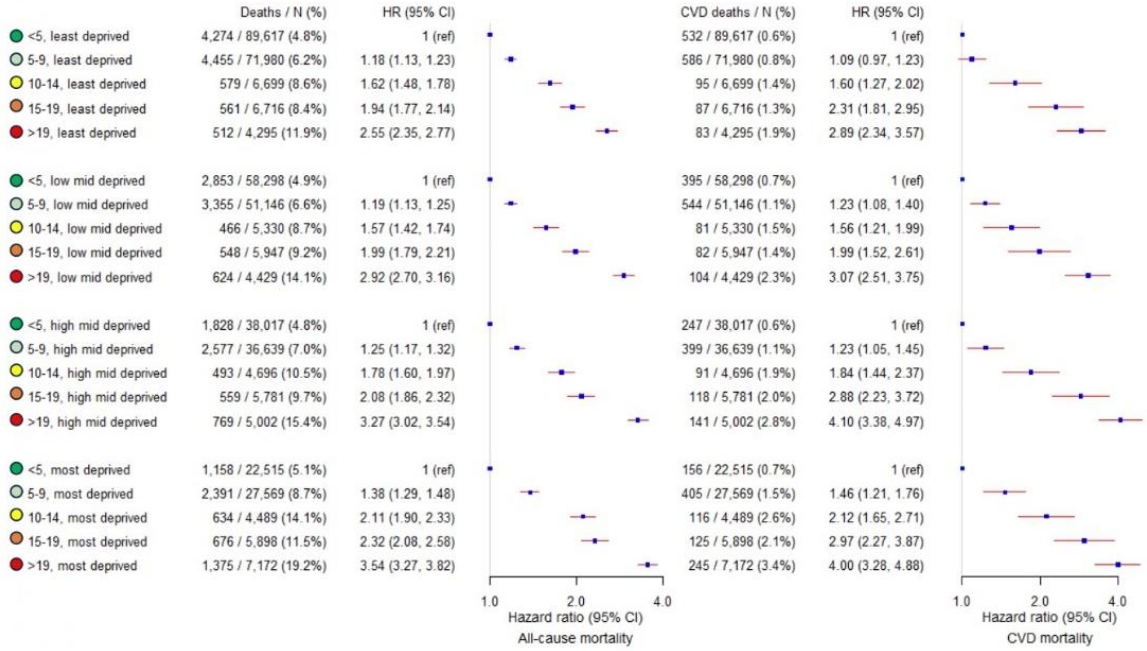


Table S4 (Interaction sex) Interaction between binary weighted lifestyle score and sex

	All-cause mortality			CVD mortality		
	Weighted lifestyle score		Effect of less healthy stratified by sex	Weighted lifestyle score		Effect of less healthy stratified by sex
	More healthy (≤9)	Less healthy (≥10)		More healthy (≤9)	Less healthy (≥10)	
	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]
Women	1 [Reference]	2.08 [1.99, 2.17]	2.08 [1.99, 2.17]	1 [Reference]	2.36 [2.06, 2.71]	2.36 [2.06, 2.71]
Men	1.64 [1.6, 1.69]	3.36 [3.25, 3.48]	2.05 [1.98, 2.12]	3.41 [3.15, 3.69]	7.56 [6.9, 8.28]	2.22 [2.06, 2.39]
Effect of male sex stratified by weighted score	1.64 [1.6, 1.69]	1.62 [1.54, 1.7]		3.41 [3.15, 3.69]	3.2 [2.81, 3.65]	
Multiplicative scale	0.99 [0.93, 1.04]			0.94 [0.81, 1.09]		
RERI	0.64 [0.52, 0.77]			2.79 [2.24, 3.37]		
AP	0.19 [0.16, 0.22]			0.37 [0.31, 0.42]		
SI	1.37 [1.29, 1.47]			1.74 [1.56, 1.94]		

Figure S4 (Combined sex) Combined association between weighted lifestyle score category, sex, and outcomes.

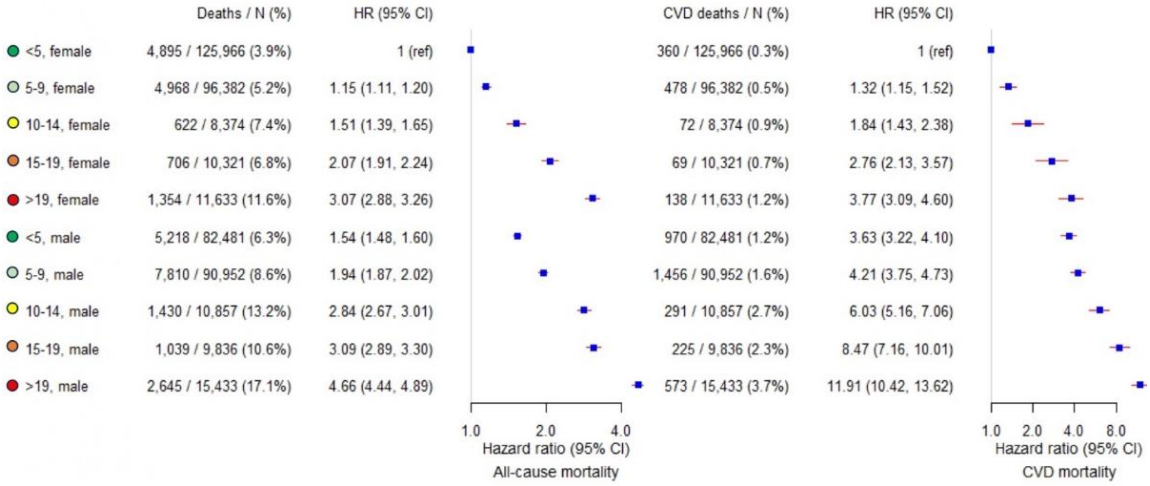


Figure S5 (Stratified sex) Association between weighted lifestyle score category and outcomes stratified by sex

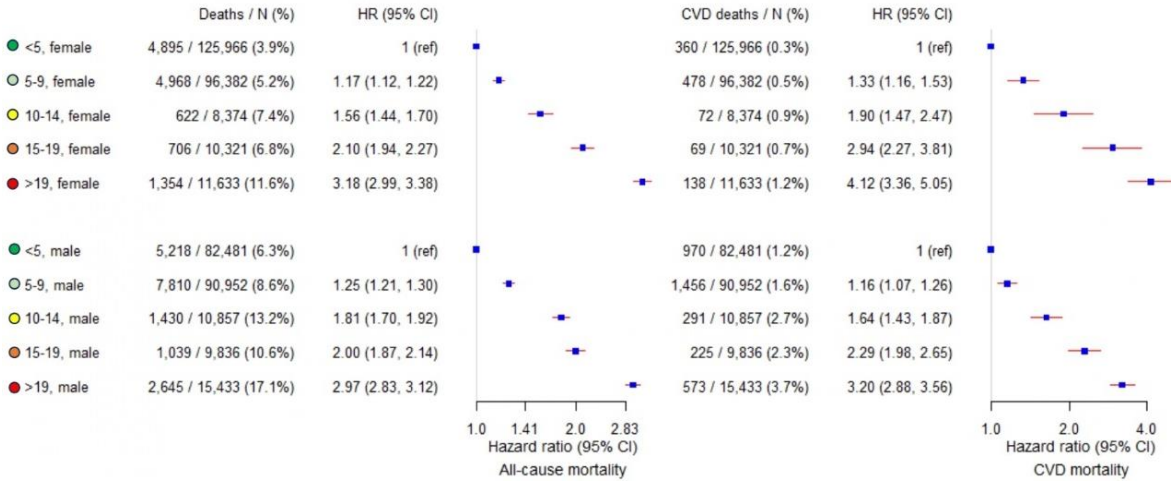


Figure S6 (Combined ethnicity) Combined association between weighted lifestyle score, ethnicity, and outcomes

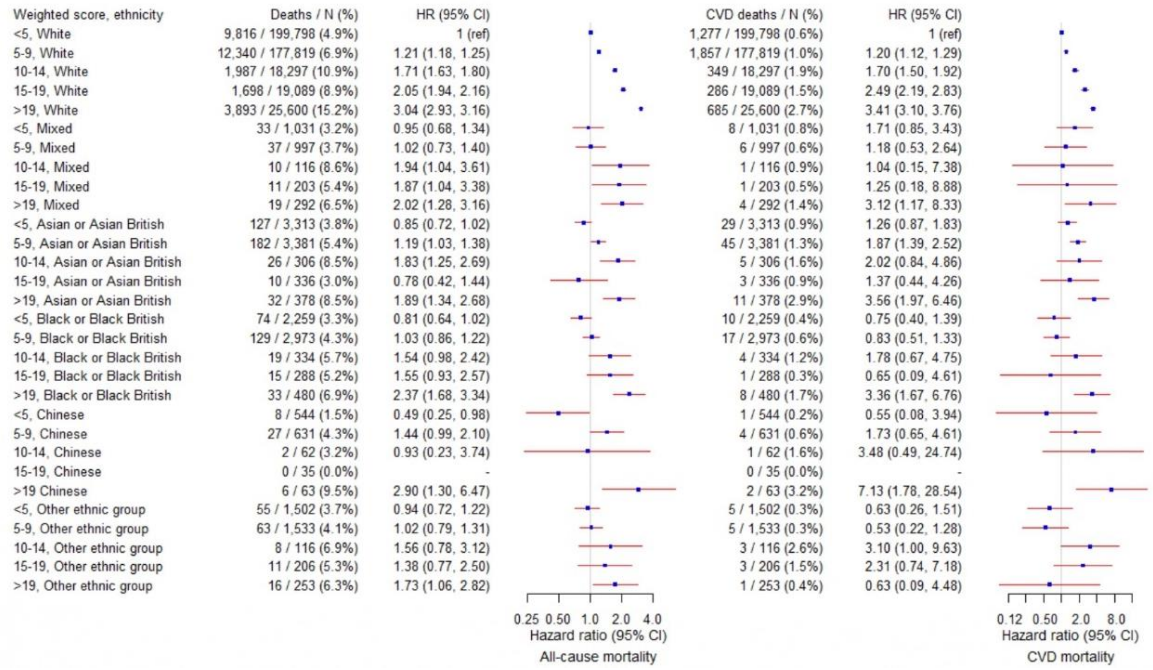


Figure S7 (Stratified ethnicity) Association between weighted lifestyle score category and outcomes stratified by ethnicity.

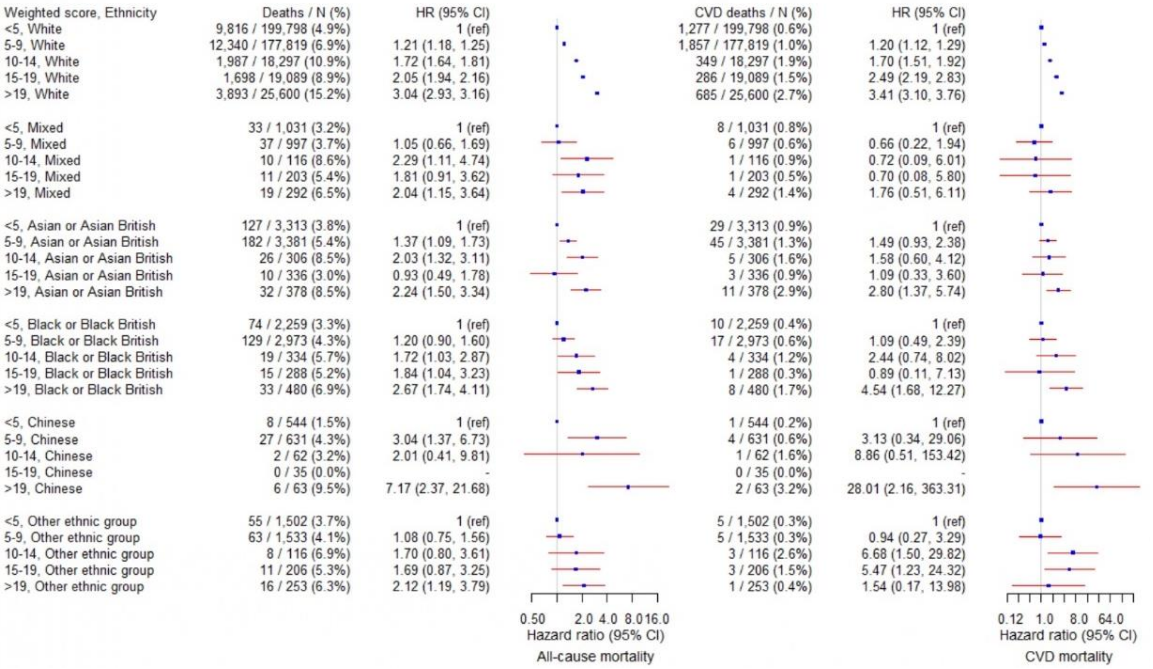


Table S5 (Interaction age) Interaction between binary weighted lifestyle score and age (>55 or ≤55) for all-cause mortality.

	All-cause mortality			CVD mortality		
	Weighted lifestyle score		Effect of less healthy stratified by age	Weighted lifestyle score		Effect of less healthy stratified by age
	More healthy (≤9)	Less healthy (≥10)		More healthy (≤9)	Less healthy (≥10)	
	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]
Age ≤55	1 [Reference]	2.13 [2.01, 2.25]	2.13 [2.01, 2.25]	1 [Reference]	2.55 [2.2, 2.96]	2.55 [2.2, 2.96]
Age >55	3.34 [3.22, 3.47]	6.29 [6.02, 6.57]	1.88 [1.83, 1.94]	3.11 [2.8, 3.45]	6.29 [5.6, 7.06]	2.02 [1.88, 2.17]
Effect of age >55 stratified by weighted score	3.34 [3.22, 3.47]	2.96 [2.8, 3.12]		3.11 [2.8, 3.45]	2.46 [2.16, 2.8]	
Multiplicative scale	0.88 [0.83, 0.94]			0.79 [0.67, 0.93]		
RERI	1.82 [1.62, 2.02]			1.63 [1.13, 2.15]		
AP	0.29 [0.26, 0.31]			0.26 [0.19, 0.32]		
SI	1.52 [1.46, 1.6]			1.44 [1.29, 1.62]		

Figure S8 (Combined age) Combined association between weighted lifestyle score category, age category, and outcomes.

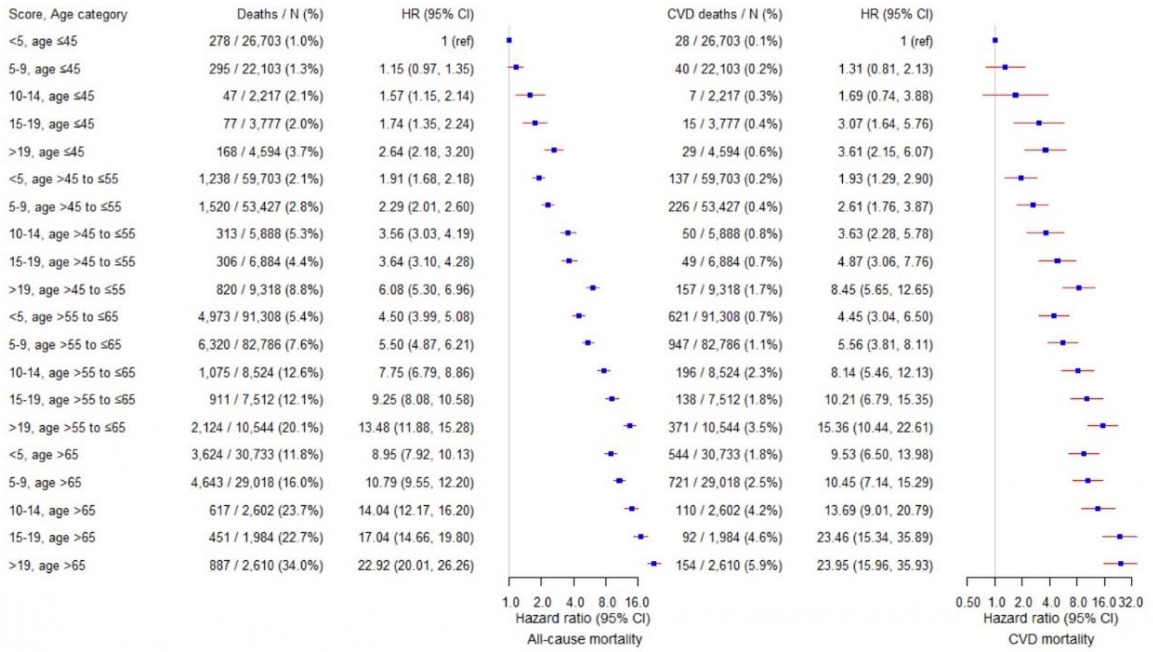
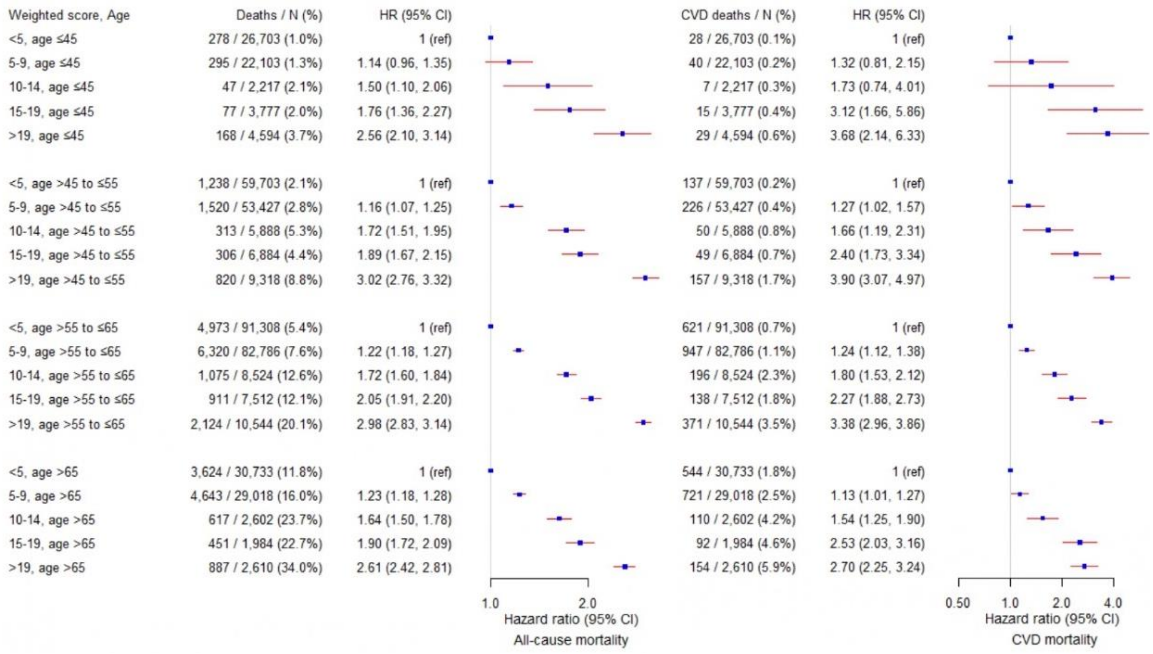


Figure S9 (Stratified age) Association between weighted lifestyle score category and outcomes stratified by age category.



References

1. Jani BD, McQueenie R, Nicholl BI, et al. Association between patterns of alcohol consumption (beverage type, frequency and consumption with food) and risk of adverse health outcomes: a prospective cohort study. *BMC Med* 2021; **19**(1): 8.
2. Osborne B, and Cooper, V. . Health Survey for England 2017 Adult health related behaviours - NHS Digital 2018.
3. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ). November 2005. https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/ipaq_analysis.pdf (accessed 13/01/2024.
4. UK Chief Medical Officers' physical activity guidelines 2020, 2020.
5. Foster HME, Ho FK, Sattar N, et al. Understanding How Much TV is Too Much: A Nonlinear Analysis of the Association Between Television Viewing Time and Adverse Health Outcomes. *Mayo Clin Proc* 2020; **95**(11): 2429-41.
6. Galante J, Adamska L, Young A, et al. The acceptability of repeat Internet-based hybrid diet assessment of previous 24-h dietary intake: administration of the Oxford WebQ in UK Biobank. *Br J Nutr* 2016; **115**(4): 681-6.
7. Food Standards Scotland - The five food groups. 2024. https://www.foodstandards.gov.scot/consumers/healthy-eating/nutrition/the-five-food-groups.
8. The Eatwell Guide (England, Wales, & Northern Ireland's recommendations on eating healthily), 2016.
9. Foster HME, Gill JMR, Mair FS, et al. Social connection and mortality in UK Biobank: a prospective cohort analysis. *BMC Med* 2023; **21**(1): 384.
10. Knol MJ, VanderWeele TJ. Recommendations for presenting analyses of effect modification and interaction. *Int J Epidemiol* 2012; **41**(2): 514-20.

Appendix 11: Weighted lifestyle score analysis - Supplementary material 2 (Chapter 6)

Sensitivity analysis - CVD mortality derived score

Summary

Associations between individual LFs and CVD mortality were weak for high alcohol intake, low oily fish intake and high added salt (Table S7), and these factors therefore did not contribute to the alternative, CVD mortality-derived weighted score. Whereas associations between CVD mortality and smoking, compared with non-smoking, were stronger than in the main analysis: HR (95% CI) 2.44 (2.27, 2.63). Using the same scale to create a simple score, smoking contributed 19 out of a total 30 points, with short/long sleep, low fruit and vegetable, high processed meat, and low social participation contributing 2 points each, and low physical activity, high TV time, and high red meat intake 1 point each (Fig S10). This again resulted a bimodal distribution of the weighted score comprising non-smokers (scores <19) and smokers (scores ⩾19) (Fig S11). There was a dose-response association between this weighted score and CVD mortality and with larger increments among smokers (Table S8).

Table S7 (CVD score) Association between individual lifestyle factors and CVD mortality.

Lifestyle factor	Healthy/Unhealthy level	N	CVD deaths (%)	HR	95% confidence intervals
Smoking status	Non-smoker	415,138	3,630 (0.9)	1(ref)	-
	Current	47,097	1,002 (2.1)	2.44	(2.27 - 2.63)
Alcohol intake	Low	420,631	4,074 (1.0)	1(ref)	-
	High	41,604	558 (1.3)	1.07	(0.97 - 1.17)
Physical activity level	High	373,376	3,477 (0.9)	1(ref)	-
	Low	88,859	1,155 (1.3)	1.11	(1.03 - 1.19)
TV time	<4 h/day	330,104	2,585 (0.8)	1(ref)	-
	≥4 hours/day	132,131	2,047 (1.5)	1.10	(1.04 - 1.17)
Sleep duration	>7 or <9h/day	315,201	2,734 (0.9)	1(ref)	-
	<7 or ≥9 hours/day	147,034	1,898 (1.3)	1.18	(1.11 - 1.25)
Fruit and vegetable intake	≥ 400 g/day	372,379	3,472 (0.9)	1(ref)	-
	< 400 g/day	89,856	1,160 (1.3)	1.14	(1.06 - 1.22)
Oily fish intake	<1 portion/week	259,554	2,641 (1.0)	1(ref)	-
	<1 portion/week	202,681	1,991 (1.0)	1.04	(0.98 - 1.10)
Red meat intake	≤3 portion/week	393,495	3,703 (0.9)	1(ref)	-
	>3 portions/week	68,740	929 (1.4)	1.09	(1.01 - 1.17)
Processed meat intake	≤1 portion/week	319,529	2,655 (0.8)	1(ref)	-
	>1 portion/week	142,706	1,977 (1.4)	1.14	(1.07 - 1.21)
Added salt	never, rarely, or sometimes	387,614	3,672 (0.9)	1(ref)	-
	usually or always	74,621	960 (1.3)	1.01	(0.93 - 1.08)
Social participation level	High	300,986	2,743 (0.9)	1(ref)	-
	Low	161,249	1,889 (1.2)	1.17	(1.10 - 1.24)
	Total	462,235	4,632 (1.0)		

Figure S10 (CVD score) Associations between individual lifestyle factors and CVD mortality to create alternative score.

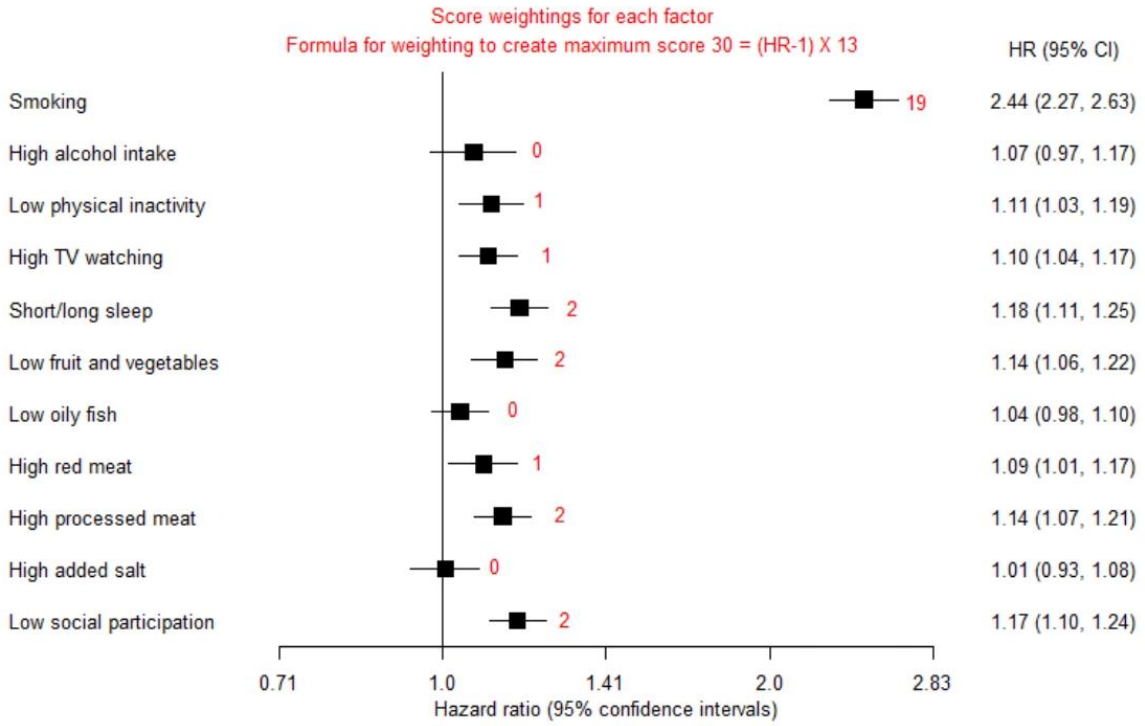


Figure S11 (CVD score) Distribution of alternative weighted score.

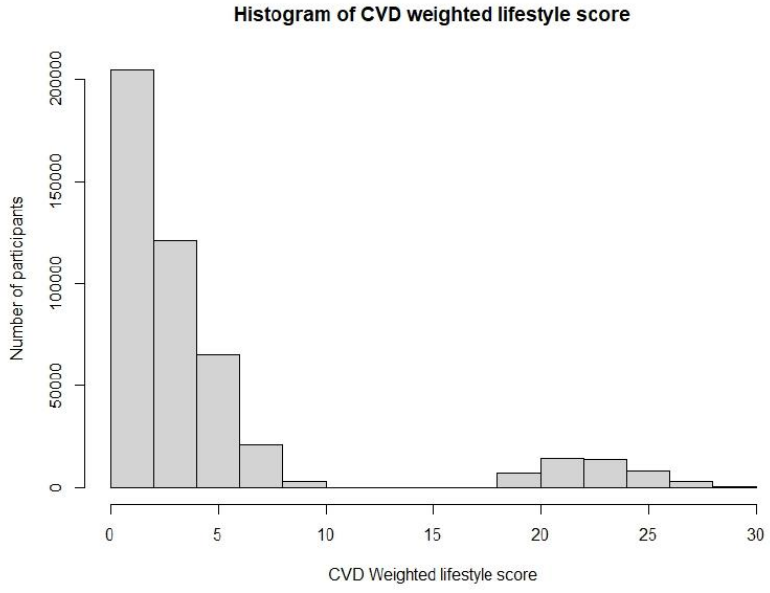


Table S8 (CVD score) Association between alternative weighted lifestyle score and CVD mortality.

Alternative (CVD mortality-derived) weighted lifestyle score	N	CVD deaths (%)	HR	Lower CI	Upper CI
0	72,672	352 (0.5)			
1	36,717	267 (0.7)	1.07	0.91	1.25
2	95,026	634 (0.7)	1.13	0.99	1.29
3	60,623	579 (1.0)	1.24	1.08	1.42
4	60,226	558 (0.9)	1.27	1.11	1.45
5	40,493	475 (1.2)	1.39	1.21	1.60
6	24,793	328 (1.3)	1.50	1.28	1.74
7	14,291	226 (1.6)	1.73	1.46	2.05
8	6,726	143 (2.1)	2.16	1.77	2.63
9	2,493	42 (1.7)	1.56	1.13	2.15
10	937	22 (2.3)	2.22	1.44	3.43
11	141	4 (2.8)	2.64	0.99	7.09
19	4,480	48 (1.1)	2.62	1.94	3.54
20	2,591	27 (1.0)	1.74	1.18	2.58
21	8,102	116 (1.4)	2.83	2.29	3.49
22	6,171	135 (2.2)	3.28	2.68	4.00
23	7,400	154 (2.1)	3.38	2.79	4.09
24	6,362	137 (2.2)	3.00	2.46	3.66
25	4,963	138 (2.8)	3.85	3.16	4.70
26	3,413	107 (3.1)	4.41	3.54	5.49
27	2,110	69 (3.3)	4.21	3.24	5.46
28	947	45 (4.8)	6.07	4.44	8.31
29	476	21 (4.4)	4.62	2.97	7.19
30	82	5 (6.1)	6.73	2.78	16.31
total	462,235	4,632 (1.0)			

Sensitivity analysis - those with baseline CVD/cancer or who died within 2 years excluded.

Summary

After excluding those with prior CVD or cancer or who died within two years of recruitment there were 295,685 (58.8%) participants remaining for reverse causality mitigation analyses (Fig S1). Broadly results were very similar to the main analyses and evidence for interactions and the patterns of combined and stratified associations remained the same between weighted score and deprivation/sex/age (Tables S9-S11 and Figs S12-S22) with minor differences presented below. Modelling the associations between individual lifestyle factors and all-cause mortality among these participants showed a very similar pattern with slightly stronger associations for smoking compared with non-smoking than in the main analysis (Figure S12). This resulted in similar score weightings but with smoking contributing 17 points and high alcohol intake 3 points and a score total of 34. Patterns of association between this weighted score and outcomes were similar to the main analyses (Figure S13, Figure S14). There was again evidence for interaction on multiplicative and additive scales between weighted score and deprivation for both outcomes (Table S9). Both combined associations between weighted score category, deprivation, and outcomes and those between weighted score category and outcomes stratified by deprivation were like the main analyses with stronger associations seen with increasing deprivation (Figure S15 & Figure S16). The effect of sex on the association between weighted score and outcomes was very similar to main analyses with evidence for an additive interaction and a large effect for male sex alone for CVD mortality (Table S10, Figure S17, Figure S18). With few events, both the combined association between weighted score category, ethnicity, and outcomes and the association between weighted score and outcomes stratified by ethnicity were difficult to interpret (Figure S19, Figure S20). With time-on-study as the timescale for models, evidence remained for additive and negative multiplicative interactions between weighted score (≤9/≥10) and age (≤55/>55) for all-cause mortality but only on the additive scale for CVD mortality (Table S11). Models of combined associations between weighted score, age, and outcomes showed markedly stronger associations in those from older age categories with unhealthier scores (Figure S21). However, as per main analyses stratified by age category, the relative increase in strength of association with higher score categories was less in older age categories (Figure S22).

Figure S12 (All-cause individual LFs) Association between individual lifestyle factors and all-cause mortality (those with baseline CVD/cancer or who died within 2 years excluded).

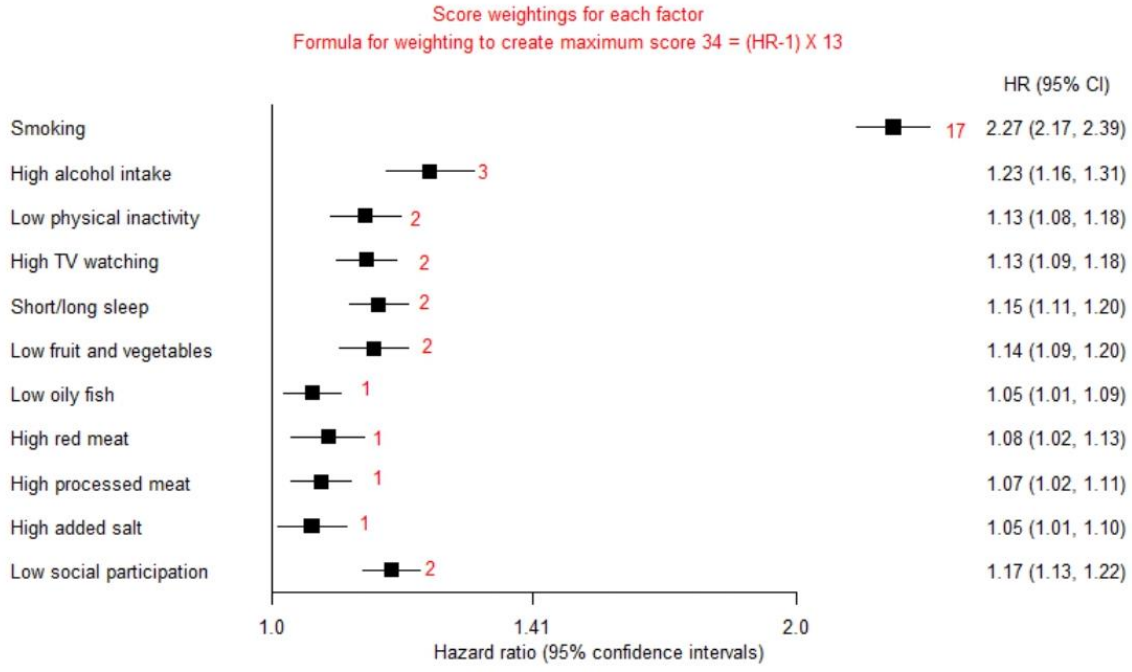


Figure S13 (All-cause weighted score) Association between alternative weighted lifestyle score and all-cause mortality (those with baseline CVD/cancer or who died within 2 years excluded).

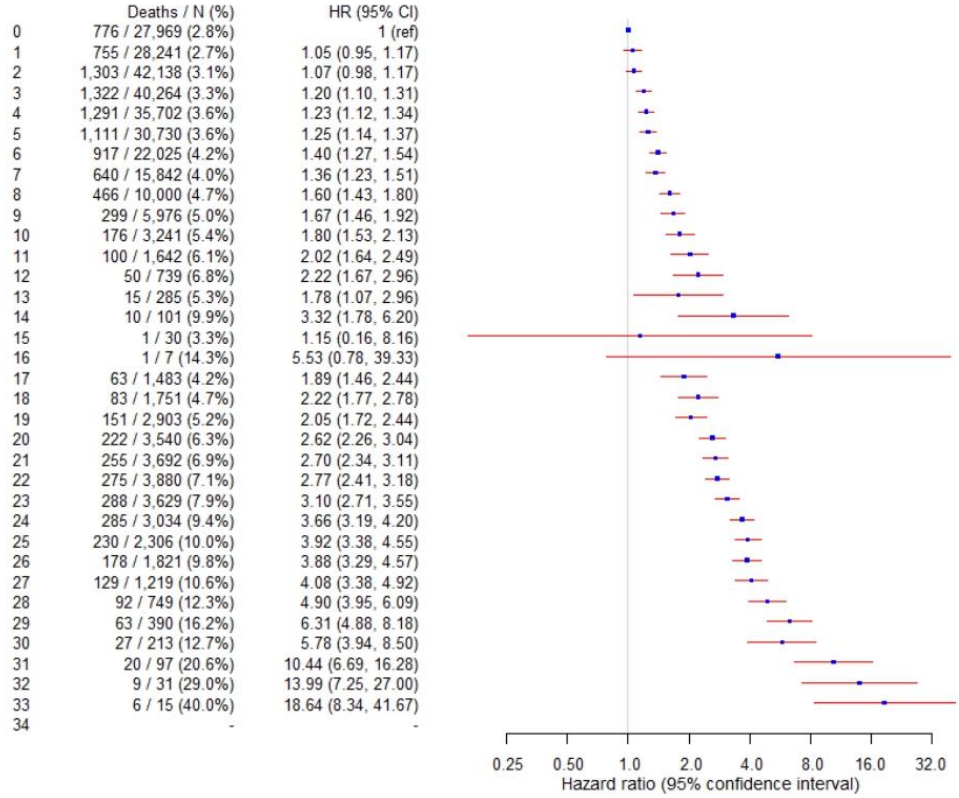


Figure S14 (CVD weighted score) Association between alternative weighted lifestyle score and CVD m (those with baseline CVD/cancer or who died within 2 years excluded).

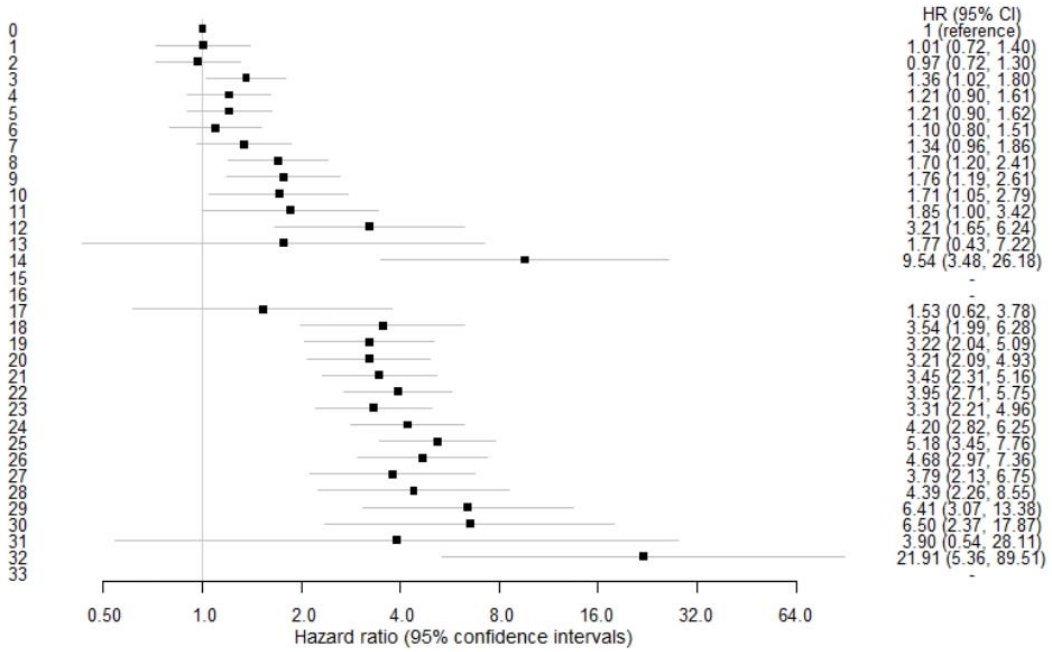


Table S9 (Interaction deprivation) Interaction between binary weighted lifestyle score and binary deprivation for outcomes (those with baseline CVD/cancer or who died within 2 years excluded)

	All-cause mortality			CVD mortality		
	Weighted lifestyle score		Effect of less healthy stratified by deprivation	Weighted lifestyle score		Effect of less healthy stratified by deprivation
	More healthy (≤9)	Less healthy (≥10)		More healthy (≤9)	Less healthy (≥10)	
	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]
Less deprived	1 [Reference]	2.07 [1.95, 2.19]	2.07 [1.95, 2.19]	1 [Reference]	2.3 [1.95, 2.72]	2.3 [1.95, 2.72]
More deprived	1.22 [1.17, 1.28]	3.06 [2.9, 3.24]	2.5 [2.35, 2.66]	1.27 [1.1, 1.46]	4.05 [3.48, 4.72]	3.19 [2.69, 3.8]
Effect of deprivation stratified by weighted score	1.22 [1.17, 1.28]	1.48 [1.38, 1.59]		1.27 [1.1, 1.46]	1.76 [1.45, 2.13]	
Multiplicative scale	1.21 [1.11, 1.32]			1.39 [1.09, 1.76]		
RERI	0.77 [0.58, 0.96]			1.48 [0.85, 2.14]		
AP	0.25 [0.2, 0.3]			0.36 [0.23, 0.47]		
SI	1.6 [1.42, 1.8]			1.94 [1.44, 2.61]		

RERI, relative excess risk for interaction; AP, attributable portion; SI, synergy index.

Figure S15 (Deprivation combined) Combined association between weighted score, deprivation quartile, and all-cause mortality (those with baseline CVD/cancer or who died within 2 years excluded)

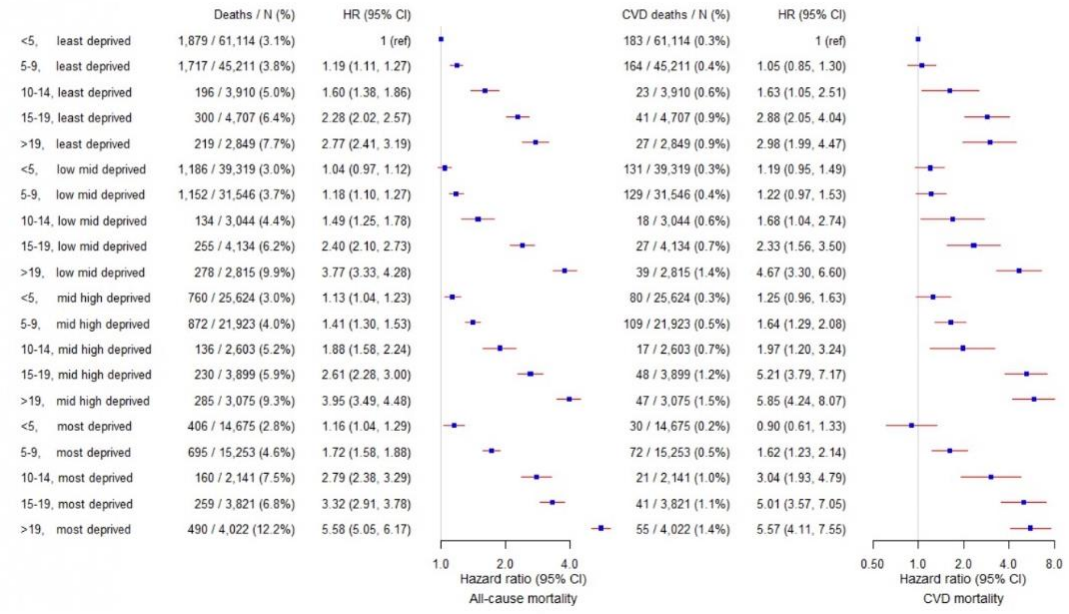


Figure S16 (Deprivation stratified) Association between weighted score and outcomes stratified by deprivation (those with baseline CVD/cancer or who died within 2 years excluded)

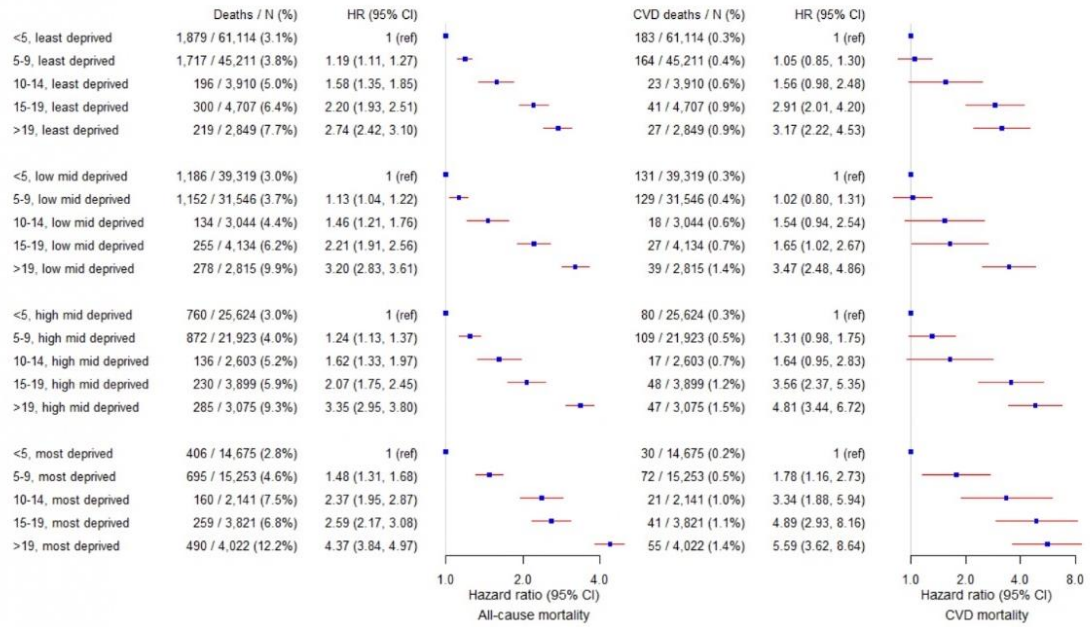
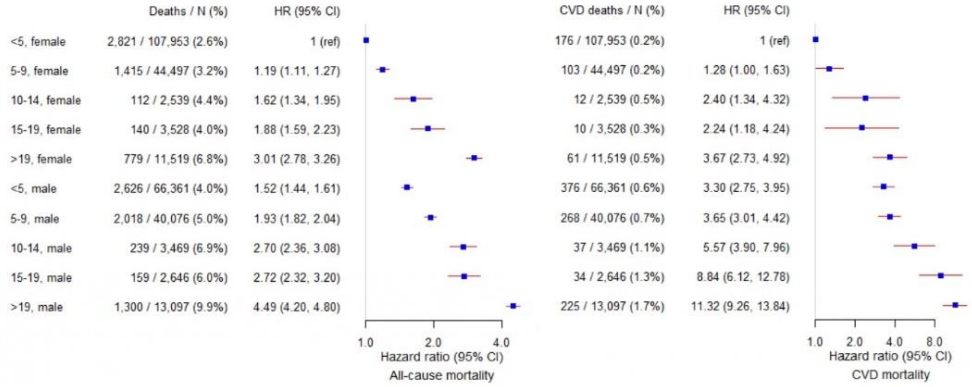


Table S10 (Interaction sex) Interaction between binary weighted lifestyle score and sex for outcomes (those with baseline CVD/cancer or who died within 2 years excluded)

	All-cause mortality			CVD mortality		
	Weighted lifestyle score		Effect of less healthy stratified by sex	Weighted lifestyle score		Effect of less healthy stratified by sex
	More healthy (≤9)	Less healthy (≥10)		More healthy (≤9)	Less healthy (≥10)	
	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]
Female	1 [Reference]	2.22 [2.08, 2.37]	2.22 [2.08, 2.37]	1 [Reference]	2.78 [2.2, 3.53]	2.78 [2.2, 3.53]
Male	1.57 [1.51, 1.64]	3.47 [3.28, 3.67]	2.2 [2.09, 2.33]	3.18 [2.75, 3.68]	8.31 [7.05, 9.79]	2.61 [2.28, 2.99]
Effect of male sex stratified by weighted score	1.57 [1.51, 1.64]	1.56 [1.45, 1.68]		3.18 [2.75, 3.68]	2.98 [2.37, 3.76]	
Multiplicative scale	0.99 [0.91, 1.08]			0.94 [0.71, 1.23]		
RERI	0.67 [0.46, 0.88]			3.34 [2.24, 4.56]		
AP	0.19 [0.14, 0.25]			0.4 [0.29, 0.49]		
SI	1.37 [1.24, 1.52]			1.84 [1.5, 2.26]		

Figure S17 (Sex combined) Combined association between weighted score, sex, and outcomes (those with baseline CVD/cancer or who died within 2 years excluded)



List of References

1. Murray CJLea. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 2020; **396**(10258): 1223-49.

2. Cockerham WC. Chapter 10: Health Lifestyle Theory. Sociological theories of health and illness. Oxford, UNITED KINGDOM: Taylor & Francis Group; 2020.

3. Ambrose JA, Barua RS. The pathophysiology of cigarette smoking and cardiovascular disease: An update. *J Am Coll Cardiol* 2004; **43**(10): 1731-7.

4. Cai C, Atanasov S. Long Sleep Duration and Stroke-Highly Linked, Poorly Understood. *Neurology International*, 2023. (accessed.

5. Mandelli L, Milaneschi Y, Hiles S, Serretti A, Penninx BW. Unhealthy lifestyle impacts on biological systems involved in stress response: hypothalamic–pituitary–adrenal axis, inflammation and autonomous nervous system. *International Clinical Psychopharmacology* 2023; **38**(3).

6. Gray ID, Kross AR, Renfrew ME, Wood P. Precision Medicine in Lifestyle Medicine: The Way of the Future? *American Journal of Lifestyle Medicine* 2019; **14**(2): 169-86.

7. Ding D, Rogers K, van der Ploeg H, Stamatakis E, Bauman AE. Traditional and Emerging Lifestyle Risk Behaviors and All-Cause Mortality in Middle-Aged and Older Adults: Evidence from a Large Population-Based Australian Cohort. *PLoS Med* 2015; **12**(12): e1001917.

8. Gascon M, Triguero-Mas M, Martínez D, et al. Residential green spaces and mortality: A systematic review. *Environment International* 2016; **86**: 60-7.

9. Doll R, Peto R, Boreham J, Sutherland I. Mortality from cancer in relation to smoking: 50 years observations on British doctors. *Br J Cancer* 2005; **92**(3): 426-9.

10. Krokstad S, Ding D, Grunseit AC, et al. Multiple lifestyle behaviours and mortality, findings from a large population-based Norwegian cohort study - The HUNT Study. *Bmc Public Health* 2017; **17**.

11. Michie S, West R. Sustained behavior change is key to preventing and tackling future pandemics. *Nature Medicine* 2021; **27**(5): 749-52.

12. Dai X, Gakidou E, Lopez AD. Evolution of the global smoking epidemic over the past half century: strengthening the evidence base for policy action. *Tob Control* 2022; **31**(2): 129-37.

13. Aneni EC, Crippa A, Osondu CU, et al. Estimates of Mortality Benefit From Ideal Cardiovascular Health Metrics: A Dose Response Meta-Analysis. *Journal of the American Heart Association* 2017; **6**(12): e006904.

14. Tsai M-C, Lee C-C, Liu S-C, Tseng P-J, Chien K-L. Combined healthy lifestyle factors are more beneficial in reducing cardiovascular disease in younger adults: a meta-analysis of prospective cohort studies. *Sci Rep-Uk* 2020; **10**(1): 18165.

15. Meader N, King K, Moe-Byrne T, et al. A systematic review on the clustering and co-occurrence of multiple risk behaviours. *Bmc Public Health* 2016; **16**: 657.

16. Holt-Lunstad J. Loneliness and Social Isolation as Risk Factors: The Power of Social Connection in Prevention. *Am J Lifestyle Med* 2021; **15**(5): 567-73.

17. Proctor AS, Barth A, Holt-Lunstad J. A healthy lifestyle is a social lifestyle: The vital link between social connection and health outcomes. *Lifestyle Medicine* 2023; **4**(4): e91.

18. Holt-Lunstad J, Robles TF, Sbarra DA. Advancing social connection as a public health priority in the United States. *Am Psychol* 2017; **72**(6): 517-30.

19. Murthy VH. Our Epidemic of Loneliness and Isolation. The U.S. Surgeon General's Advisory on the Healing Effects of Social Connection and Community. https://www.hhs.gov/sites/default/files/surgeon-general-social-connection-advisory.pdf; 2023.

20. Marmot MG, Shipley MJ, Hemingway H, Head J, Brunner EJ. Biological and behavioural explanations of social inequalities in coronary heart disease: the Whitehall II study. *Diabetologia* 2008; **51**(11): 1980-8.

21. Laine JE, Baltar VT, Stringini S, et al. Reducing socio-economic inequalities in all-cause mortality: a counterfactual mediation approach. *Int J Epidemiol* 2019.

22. Smith GD, Hart C, Hole D, et al. Education and occupational social class: which is the more important indicator of mortality risk? *J Epidemiol Commun H* 1998; **52**(3): 153-60.

23. Stringhini S, Carmeli C, Jokela M, et al. Socioeconomic status and the 25 × 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1·7 million men and women. *The Lancet* 2017; **389**(10075): 1229-37.

24. Foster HME, Celis-Morales CA, Nicholl BI, et al. The effect of socioeconomic deprivation on the association between an extended measurement of unhealthy lifestyle factors and health outcomes: a prospective analysis of the UK Biobank cohort. *Lancet Public Health* 2018; **3**(12): e576-e85.

25. Hagger MS, Weed M. DEBATE: Do interventions based on behavioral theory work in the real world? *Int J Behav Nutr Phy* 2019; **16**(1): 36.

26. Bull ER, Dombrowski SU, McCleary N, Johnston M. Are interventions for low-income groups effective in changing healthy eating, physical activity and smoking behaviours? A systematic review and meta-analysis. *BMJ Open* 2014; **4**(11): e006046.

27. Bull ER, McCleary N, Li X, Dombrowski SU, Dusseldorp E, Johnston M. Interventions to Promote Healthy Eating, Physical Activity and Smoking in Low-Income Groups: a Systematic Review with Meta-Analysis of Behavior Change Techniques and Delivery/Context. *International Journal of Behavioral Medicine* 2018; **25**(6): 605-16.

28. Michie S, Jochelson K, Markham WA, Bridle C. Low-income groups and behaviour change interventions: a review of intervention content, effectiveness and theoretical frameworks. *J Epidemiol Commun H* 2009; **63**(8): 610.

29. Katikireddi SV, Higgins M, Smith KE, Williams G. Health inequalities: the need to move beyond bad behaviours. *J Epidemiol Community Health* 2013; **67**(9): 715-6.

30. Baum F, Fisher M. Why behavioural health promotion endures despite its failure to reduce health inequities. *Sociology of Health & Illness* 2014; **36**(2): 213-25.

31. Foster HME, Polz, P., Mair, F.S., Gill, J.M.R., O'Donnell, C.A.,. Understanding the influence of socioeconomic status on the association between combinations of lifestyle factors and adverse health outcomes: a systematic review protocol. *BMJ Open (Accepted for publication)* 2021.

32. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *Bmj-Brit Med J* 2020; **368**.

33. Collins R. What makes UK Biobank special? *Lancet* 2012; **379**(9822): 1173-4.

34. OED OED. Oxford English Dictionary, s.v. lifestyle (n.), Etymology,". Oxford English Dictionary: Oxford University Press; 2023.

35. Alder A. Problems of neurosis. A Book of Case-Histories. London Kegan Paul, Trench, Trübner; 1929.

36. Weber M. Economy and society : an outline of interpretive sociology translated by G. Roth and C. Wittich. Berkley, Los Angeles, London: University of California Press; 1978/1922.

37. Abel T, Cockerham WC. Lifestyle or Lebensführung? Critical Remarks on the Mistranslation of Weber's "Class, Status, Party". *The Sociological Quarterly* 1993; **34**(3): 551-6.

38. de Gruchy J. The lazy language of 'lifestyles' – let's rid this from our talk about prevention. Association of Directors of Public Health. Published online; 2019.

39. O'Conner PT, and Kellerman, S. The life of 'lifestyle'. Grammarphobia; 2018.

40. Michel JB, Shen YK, Aiden AP, et al. Quantitative analysis of culture using millions of digitized books. *Science* 2011; **331**(6014): 176-82.

41. Wangwe S, Kawamura H. The 1960s: High Growth, High Hopes, and Looming Structural Imbalances. In: Ocampo JA, Chowdhury A, Alarcón D, eds. The World Economy through the Lens of the United Nations: Oxford University Press; 2018: 0.

42. Lytle MH. The Golden Age of Consumption. The All-Consuming Nation: Chasing the American Dream Since World War II: Oxford University Press; 2021: 0.

43. Veblen T. The theory of the leisure class. New York: Dover Publications; 1994.

44. Richards H, Reid M, Watt G. Victim-blaming revisited: a qualitative study of beliefs about illness causation, and responses to chest pain. *Fam Pract* 2003; **20**(6): 711-6.

45. Graham H. Health inequalities, social determinants and public health policy. *Policy & Politics* 2009; **37**(4): 463-79.

46. Lincoln P. Lifestyle: a plea to abandon the use of this word in public health. UK Public Health netwrok; 2015.

47. Rothman KJ. Modern epidemiology. 4th ed. Philadelphia: Wolters Kluwer; 2021.

48. Lacy-Nichols J, Nandi S, Mialon M, et al. Conceptualising commercial entities in public health: beyond unhealthy commodities and transnational corporations. *Lancet* 2023; **401**(10383): 1214-28.

49. Knai C, Petticrew M, Mays N, et al. Systems Thinking as a Framework for Analyzing Commercial Determinants of Health. *The Milbank Quarterly* 2018; **96**(3): 472-98.

50. Doll R. Uncovering the effects of smoking: historical perspective. *Statistical Methods in Medical Research* 1998; **7**(2): 87-117.

51. World Health Organization - Global action plan for the prevention and control of noncommunicable diseases 2013-2020., 2013.

52. Diet, nutrition, physical activity, and oesophageal cancer -The World Cancer Research Fund, Continuous Update Project, 2018.

53. World Cancer Research Fund/American Institute for Cancer Research , Diet, Nutrition, Physcial Activity and Cancer: a Global Perspective. Continuous Update Project Expert Report 2018.

54. Wu J, Zhang H, Yang L, et al. Sedentary time and the risk of metabolic syndrome: A systematic review and dose–response meta-analysis. *Obesity Reviews* 2022; **23**(12): e13510.

55. Itani O, Jike M, Watanabe N, Kaneita Y. Short sleep duration and health outcomes: a systematic review, meta-analysis, and meta-regression. *Sleep Medicine* 2017; **32**: 246-56.

56. Jike M, Itani O, Watanabe N, Buysse DJ, Kaneita Y. Long sleep duration and health outcomes: A systematic review, meta-analysis and meta-regression. *Sleep Med Rev* 2018; **39**: 25-36.

57. Scott AJ, Webb TL, Martyn-St James M, Rowse G, Weich S. Improving sleep quality leads to better mental health: A meta-analysis of randomised controlled trials. *Sleep Med Rev* 2021; **60**: 101556.

58. Sommerlad A, Kivimäki M, Larson EB, et al. Social participation and risk of developing dementia. *Nature Aging* 2023; **3**(5): 532-45.

59. Walser T, Cui X, Yanagawa J, et al. Smoking and lung cancer: the role of inflammation. *Proc Am Thorac Soc* 2008; **5**(8): 811-5.

60. Fedak KM, Bernal A, Capshaw ZA, Gross S. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. *Emerg Themes Epidemiol* 2015; **12**: 14.

61. Schuit AJ, van Loon AJ, Tijhuis M, Ocké M. Clustering of lifestyle risk factors in a general adult population. *Prev Med* 2002; **35**(3): 219-24.

62. Noble N, Paul C, Turon H, Oldmeadow C. Which modifiable health risk behaviours are related? A systematic review of the clustering of Smoking, Nutrition, Alcohol and Physical activity ('SNAP') health risk factors. *Preventive Medicine* 2015; **81**: 16-41.

63. Burton R, Sharpe C, Sheron N, et al. The prevalence and clustering of alcohol consumption, gambling, smoking, and excess weight in an English adult population. *Preventive Medicine* 2023; **175**: 107683.

64. Fine LJ, Philogene GS, Gramling R, Coups EJ, Sinha S. Prevalence of multiple chronic disease risk factors. 2001 National Health Interview Survey. *Am J Prev Med* 2004; **27**(2 Suppl): 18-24.

65. Funderburk JS, Maisto SA, Sugarman DE, Wade M. The covariation of multiple risk factors in primary care: a latent class analysis. *J Behav Med* 2008; **31**(6): 525-35.

66. Chiolero A, Wietlisbach V, Ruffieux C, Paccaud F, Cornuz J. Clustering of risk behaviors with cigarette consumption: A population-based survey. *Preventive Medicine* 2006; **42**(5): 348-53.

67. Poortinga W. The prevalence and clustering of four major lifestyle risk factors in an English adult population. *Prev Med* 2007; **44**(2): 124-8.

68. Ortiz C, Lopez-Cuadrado T, Rodriguez-Blazquez C, Pastor-Barriuso R, Galan I. Clustering of unhealthy lifestyle behaviors, self-rated health and disability. *Prev Med* 2022; **155**: 106911.

69. Vincent GE, Jay SM, Sargent C, Vandelanotte C, Ridgers ND, Ferguson SA. Improving Cardiometabolic Health with Diet, Physical Activity, and Breaking Up Sitting: What about Sleep? *Front Physiol* 2017; **8**: 865.

70. Burton R, Fryers PT, Sharpe C, et al. The independent and joint risks of alcohol consumption, smoking, and excess weight on morbidity and mortality: a systematic review and meta-analysis exploring synergistic associations. *Public Health* 2024; **226**: 39-52.

71. Zhang Y, Pan XF, Chen J, et al. Combined lifestyle factors and risk of incident type 2 diabetes and prognosis among individuals with type 2 diabetes: a systematic review and meta-analysis of prospective cohort studies. *Diabetologia* 2020; **63**(1): 21-33.

72. Zhang YB, Pan XF, Chen J, et al. Combined lifestyle factors, all-cause mortality and cardiovascular disease: a systematic review and meta-analysis of prospective cohort studies. *J Epidemiol Community Health* 2020.

73. Zhang YB, Pan XF, Chen J, et al. Combined lifestyle factors, incident cancer, and cancer mortality: a systematic review and meta-analysis of prospective cohort studies. *Br J Cancer* 2020; **122**(7): 1085-93.

74. Khaw KT, Wareham N, Bingham S, Welch A, Luben R, Day N. Combined impact of health behaviours and mortality in men and women: the EPIC-Norfolk prospective population study. *PLoS Med* 2008; **5**(1): e12.

75. Leventhal AM, Huh J, Dunton GF. Clustering of modifiable biobehavioral risk factors for chronic disease in US adults: a latent class analysis. *Perspect Public Health* 2014; **134**(6): 331-8.

76. Tegegne TK, Islam SMS, Maddison R. Effects of lifestyle risk behaviour clustering on cardiovascular disease among UK adults: latent class analysis with distal outcomes. *Sci Rep* 2022; **12**(1): 17349.

77. Oftedal S, Vandelanotte C, Duncan MJ. Patterns of Diet, Physical Activity, Sitting and Sleep Are Associated with Socio-Demographic, Behavioural, and Health-Risk Indicators in Adults. *Int J Environ Res Public Health* 2019; **16**(13).

78. Spring B, Moller AC, Coons MJ. Multiple health behaviours: overview and implications. *Journal of Public Health* 2012; **34**(suppl_1): i3-i10.

79. King K, Meader N, Wright K, et al. Characteristics of interventions targeting multiple lifestyle risk behaviours in adult populations: a systematic scoping review. *PLoS One* 2015; **10**(1): e0117015.

80. James E, Freund M, Booth A, et al. Comparative efficacy of simultaneous versus sequential multiple health behavior change interventions among adults: A systematic review of randomised trials. *Preventive Medicine* 2016; **89**: 211-23.

81. Maisano MS, Shonkoff ET, Folta SC. Multiple Health Behavior Change for Weight Loss: A Scoping Review. *Am J Health Behav* 2020; **44**(5): 559-71.

82. Meader N, King K, Wright K, et al. Multiple Risk Behavior Interventions: Meta-analyses of RCTs. *Am J Prev Med* 2017; **53**(1): e19-e30.

83. Health Survey England Additional Analyses, Ethnicity and Health, 2011-2019 Experimental statistics: NHS Digital, 2022.

84. Baumeister RF, Leary MR. The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin* 1995; **117**(3): 497-529.

85. Uchino BN. Social support and health: a review of physiological processes potentially underlying links to disease outcomes. *J Behav Med* 2006; **29**(4): 377-87.

86. Cacioppo JT, Cacioppo S, Boomsma DI. Evolutionary mechanisms for loneliness. *Cogn Emot* 2014; **28**(1): 3-21.

87. Holt-Lunstad J. Why social relationships are important for physical health: A systems approach to understanding and modifying risk and protection. *Annual Review of Psychology* 2018; **69**: 437-58.

88. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. *Perspectives on Psychological Science* 2015; **10**(2): 227-37.

89. Holt-Lunstad J, Smith TB, Layton JB. Social relationships and mortality risk: a meta-analytic review. *PLoS Med* 2010; **7**(7): e1000316.

90. Holt-Lunstad J. Social Connection as a Public Health Issue: The Evidence and a Systemic Framework for Prioritizing the "Social" in Social Determinants of Health. *Annu Rev Public Health* 2022; **43**: 193-213.

91. Barjakova M, Garnero A, d'Hombres B. Risk factors for loneliness: A literature review. *Soc Sci Med* 2023; **334**: 116163.

92. Hodgson S, Watts I, Fraser S, Roderick P, Dambha-Miller H. Loneliness, social isolation, cardiovascular disease and mortality: a synthesis of the literature and conceptual framework. *J R Soc Med* 2020; **113**(5): 185-92.

93. Huang J, Xiao L, Zhao H, Liu F, Du L. Living alone increases the risk of developing type 2 diabetes mellitus: A systematic review and meta-analysis based on longitudinal studies. *Primary Care Diabetes* 2024; **18**(1): 1-6.

94. Maes M, Qualter P, Lodder GMA, Mund M. How (Not) to Measure Loneliness: A Review of the Eight Most Commonly Used Scales. *Int J Environ Res Public Health* 2022; **19**(17): 10816.

95. Naito R, McKee M, Leong D, et al. Social isolation as a risk factor for all-cause mortality: Systematic review and meta-analysis of cohort studies. *PLoS One* 2023; **18**(1): e0280308.

96. Robles TF, Slatcher RB, Trombello JM, McGinn MM. Marital quality and health: a meta-analytic review. *Psychol Bull* 2014; **140**(1): 140-87.

97. Schutter N, Holwerda TJ, Comijs HC, Stek ML, Peen J, Dekker JJM. Loneliness, social network size and mortality in older adults: a meta-analysis. *Eur J Ageing* 2022; **19**(4): 1057-76.

98. Surkalim DL, Luo M, Eres R, et al. The prevalence of loneliness across 113 countries: systematic review and meta-analysis. *BMJ* 2022; **376**: e067068.

99. Teo RH, Cheng, W.H., Cheng, L.G., Lau, Y., Lau, S.T. Global prevalence of social isolation among community-dwelling older adults: A systematic review and meta-analysis. *Archives of Gerontology and Geriatrics* 2022; **107**: 104904.

100. Valtorta NK, Kanaan M, Gilbody S, Ronzi S, Hanratty B. Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. *Heart* 2016; **102**(13): 1009-16.

101. Wang F, Gao Y, Han Z, et al. A systematic review and meta-analysis of 90 cohort studies of social isolation, loneliness and mortality. *Nat Hum Behav* 2023; **7**(8): 1307-19.

102. Leigh-Hunt N, Bagguley D, Bash K, et al. An overview of systematic reviews on the public health consequences of social isolation and loneliness. *Public Health* 2017; **152**: 157-71.

103. Rico-Uribe LA, Caballero FF, Martin-Maria N, Cabello M, Ayuso-Mateos JL, Miret M. Association of loneliness with all-cause mortality: A meta-analysis. *PLoS One* 2018; **13**(1): e0190033.

104. Community Life Survey 2021/22. In: Department for Culture MaS, UK, editor. Online: GOV.UK; 2023.

105. The State of Loneliness 2023 – ONS data on loneliness in Britain (Campaign to End Loneliness). Online: Campaign to End Loneliness, 2023.

106. Families and households in the UK: 2023. Office for National Statistics - Statistical bulletin. Online: Office for National Statistics, 2024.

107. Older people living in care homes in 2021 and changes since 2011. Census 2021. Online: Ofice for National Statistics, 2023.

108. Shekelle PG, Miake-Lye IM, Begashaw MM, et al. Interventions to Reduce Loneliness in Community-Living Older Adults: a Systematic Review and Meta-analysis. *J Gen Intern Med* 2024; **39**(6): 1015-28.

109. Fakoya OA, McCorry NK, Donnelly M. Loneliness and social isolation interventions for older adults: a scoping review of reviews. *Bmc Public Health* 2020; **20**(1): 129.

110. Department of Health and Social Security. Inequalities in health. Report of a research working group - 'The Black Report', 1980.

111. Berkman LF, Kawachi I, Glymour MM. Social epidemiology: Oxford University Press; 2014.

112. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annu Rev Public Health* 1997; **18**: 341-78.

113. Engels F. The Condition of the Working-Class in England in 1844: With Preface Written in 1892. Cambridge: Cambridge University Press; 2010.

114. Mackenbach JP. Politics is nothing but medicine at a larger scale: reflections on public health's biggest idea. *J Epidemiol Commun H* 2009; **63**(3): 181.

115. Durkheim E. Suicide : A Study in Sociology (first published 1897). Milton, UNITED KINGDOM: Taylor & Francis Group; 2002.

116. Bourdieu P. Outline of a Theory of Practice. Cambridge: Cambridge University Press; 1977.

117. Waitzkin H. THE SOCIAL ORIGINS OF ILLNESS: A NEGLECTED HISTORY. *International Journal of Health Services* 1981; **11**(1): 77-103.

118. Cockerham WC, Rütten A, Abel T. Conceptualizing Contemporary Health Lifestyles: Moving Beyond Weber. *The Sociological Quarterly* 1997; **38**(2): 321-42.

119. On Airs, Waters, and Places [attributed to Hippocrates (c. 460 – c. 370 B.C.)]. New York: translated from Greek by W.H.S. Jones, from (1923.

120. Cockerham WC. Health lifestyles. The Wiley Blackwell Companion to Medical Sociology; 2021: 150-70.

121. Edgerton JD, Roberts LW. Cultural capital or habitus? Bourdieu and beyond in the explanation of enduring educational inequality. *Theory and Research in Education* 2014; **12**(2): 193-220.

122. Blue S, Shove, E., Carmona, C., & Kelly, M.P. . Theories of practice and public health: understanding (un)healthy practices. *Critical Public Health* 2014; **26**(1): 36-50.

123. Cohn S. From health behaviours to health practices: an introduction. *Sociol Health Illn* 2014; **36**(2): 157-62.

124. Cockerham WC. Health lifestyle theory and the convergence of agency and structure. *J Health Soc Behav* 2005; **46**(1): 51-67.

125. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G. Indicators of socioeconomic position (part 1). *J Epidemiol Community Health* 2006; **60**(1): 7-12.

126. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G. Indicators of socioeconomic position (part 2). *J Epidemiol Community Health* 2006; **60**(2): 95-101.

127. Bartley M. Health Inequality : An Introduction to Concepts, Theories and Methods. Oxford, UNITED KINGDOM: Polity Press; 2016.

128. Parsons T, Bryan T. The Social System. Florence, UNITED STATES: Taylor & Francis Group; 1991.

129. Antonovsky A. Social Class, Life Expectancy and Overall Mortality. *The Milbank Memorial Fund Quarterly* 1967; **45**(2): 31-73.

130. Oakes JM, Rossi PH. The measurement of SES in health research: current practice and steps toward a new approach. *Soc Sci Med* 2003; **56**(4): 769-84.

131. Galobardes B, Lynch J, Smith GD. Measuring socioeconomic position in health research. *Br Med Bull* 2007; **81-82**: 21-37.

132. Darin-Mattsson A, Fors S, Kareholt I. Different indicators of socioeconomic status and their relative importance as determinants of health in old age. *Int J Equity Health* 2017; **16**(1): 173.

133. Stringhini S, Carmeli C, Jokela M, et al. Socioeconomic status, non-communicable disease risk factors, and walking speed in older adults: multi-cohort population based study. *Bmj-Brit Med J* 2018; **360**.

134. Witrick B, Dotson TS, Annie F, Kimble W, Kemper E, Hendricks B. Poverty and population health – The need for A Paradigm shift to capture the working poor and better inform public health planning. *Social Science & Medicine* 2023; **336**: 116249.

135. Mel B, David B. Commentary: Appropriateness of deprivation indices must be ensured. *BMJ* 1994; **309**(6967): 1479.

136. Regidor E. Measures of health inequalities: part 1. *J Epidemiol Community Health* 2004; **58**(10): 858-61.

137. Regidor E. Measures of health inequalities: part 2. *J Epidemiol Community Health* 2004; **58**(11): 900-3.

138. Graham H. Tackling inequalities in health in England: Remedying health disadvantages, narrowing health gaps or reducing health gradients? *Journal of Social Policy* 2004; **33**(1): 115-31.

139. Howe LD, Galobardes B, Matijasevich A, et al. Measuring socio-economic position for epidemiological studies in low- and middle-income countries: a methods of measurement in epidemiology paper. *International Journal of Epidemiology* 2012; **41**(3): 871-86.

140. Catalogue of Bias Collaboration, Spencer EA, Mahtani KR, Brassey J, Heneghan C. Misclassification bias. Catalogue Of Bias. 2018. http://www.catalogueofbiases.org/biases/misclassificationbias (accessed 31/01/2024.

141. Cancer Research UK: Percentage Deprivation Gap in European Age-Standardised Mortality Rates, Statistically Significant Cancers, Females, England, 2007-2011. https://www.cancerresearchuk.org/health-professional/cancer-statistics/mortality/deprivation-gradient#heading-One (accessed 12/05/2024.

142. Marmot M. Fair Society, Health Lives: The Marmot Review. Strategic review of health inequalities in England post-2010. University College London, 2010.

143. Marmot M, Allen J, Goldblatt P, Herd E, Morrison J. Build Back Fairer: The COVID-19 Marmot Review: Institute of Health Equity, 2020.

144. Solar O, Irwin, A. Towards a Conceptual Framework for Analysis and Action on the Social Determinants of Health - World Health Organisation. Social Determinants of Health Discussion Paper 2 (Policy and Practice). . Online: World Health Organisation - Commission on Social Determinants of Health, 2007.

145. Townsend P. Poverty in the United Kingdom: a survey of household resources and standards of living: Univ of California Press; 1979.

146. Townsend P. Deprivation. *Journal of Social Policy* 1987; **16**(2): 125-46.

147. Scottish Government. Scottish Index of Multiple Deprivation 2020, 2020.

148. Morris R, Carstairs V. Which deprivation? A comparison of selected deprivation indexes. *Journal of Public Health Medicine* 1991; **13**(4): 318-26.

149. Smith GD. Health inequalities: lifecourse approaches: Policy Press; 2003.

150. Townsend P, Philimore, P., Beattie, A. Health and deprivation: inequality and the North: Croom Helm; 1988.

151. Jarman B. Identification of underprivileged areas. *Br Med J (Clin Res Ed)* 1983; **286**(6379): 1705-9.

152. Carstairs V, Morris R. Deprivation and health in Scotland. *Health bulletin* 1990; **48**(4): 162.

153. McLennan D, Noble, S., Noble, M., Plunkett, E., Wright, G., and Gutacker, N. The English Indices of Deprivation 2019 - Techincal Report. Published online: Ministy of Housing, Communiies & Local Gvernment, 2019.

154. Scottish Government. SIMD 2020 technical notes. Published online, 2020.

155. The World Bank. Principles and Practice in Measuring Global Poverty. 2016. https://www.worldbank.org/en/news/feature/2016/01/13/principles-and-practice-in-measuring-global-poverty (accessed 21/01/2024.

156. Baah SKT, Aguilar, R.A.C., Diaz-Bonilla, C., Fujs, T., Lakner, T., Nguyen, M.C., Viveros, M. March 2023 global poverty update from the World Bank: the challenge of estimating poverty in the pandemic. 2023. https://blogs.worldbank.org/opendata/march-2023-global-poverty-update-world-bank-challenge-estimating-poverty-pandemic#:~:text=The%20global%20poverty%20headcount%20ratio,from%20648%20to%20659%20million. (accessed 31/01/2024.

157. Pronk NP, Anderson LH, Crain AL, et al. Meeting recommendations for multiple healthy lifestyle factors. Prevalence, clustering, and predictors among adolescent, adult, and senior health plan members. *Am J Prev Med* 2004; **27**(2 Suppl): 25-33.

158. Coups EJ, Gaba A, Orleans CT. Physician screening for multiple behavioral health risk factors. *Am J Prev Med* 2004; **27**(2 Suppl): 34-41.

159. Shankar A, McMunn A, Steptoe A. Health-related behaviors in older adults relationships with socioeconomic status. *Am J Prev Med* 2010; **38**(1): 39-46.

160. Silva DAS, Peres KG, Boing AF, González-Chica DA, Peres MA. Clustering of risk behaviors for chronic noncommunicable diseases: A population-based study in southern Brazil. *Preventive Medicine* 2013; **56**(1): 20-4.

161. Algren MH, Bak CK, Berg-Beckhoff G, Andersen PT. Health-Risk Behaviour in Deprived Neighbourhoods Compared with Non-Deprived Neighbourhoods: A Systematic Literature Review of Quantitative Observational Studies. *PLoS One* 2015; **10**(10): e0139297.

162. Morris LJ, D'Este C, Sargent-Cox K, Anstey KJ. Concurrent lifestyle risk factors: Clusters and determinants in an Australian sample. *Prev Med* 2016; **84**: 1-5.

163. Nordahl H, Lange T, Osler M, et al. Education and Cause-specific Mortality The Mediating Role of Differential Exposure and Vulnerability to Behavioral Risk Factors. *Epidemiology* 2014; **25**(3): 389-96.

164. Bihan H, Backholer K, Peeters A, Stevenson CE, Shaw JE, Magliano DJ. Socioeconomic Position and Premature Mortality in the AusDiab Cohort of Australian Adults. *Am J Public Health* 2016; **106**(3): 470-7.

165. Petrovic D, de Mestral C, Bochud M, et al. The contribution of health behaviors to socioeconomic inequalities in health: A systematic review. *Prev Med* 2018; **113**: 15-31.

166. Khalatbari-Soltani S, Blyth FM, Naganathan V, et al. Socioeconomic status, health-related behaviours, and death among older people: the Concord health and aging in men project prospective cohort study. *BMC Geriatrics* 2020; **20**(1): 261.

167. Wang Z, Zheng Y, Ruan H, Li L, He S. Promotion of Healthy Lifestyles Alone Might Not Substantially Reduce Socioeconomic Inequity-Related Mortality Risk in Older People in China: A Prospective Cohort Study. *Journal of Epidemiology and Global Health* 2023; **13**(2): 322-32.

168. Olsen JA, Chen G, Lamu AN. The relative importance of education and health behaviour for health and wellbeing. *Bmc Public Health* 2023; **23**(1): 1981.

169. Berg L, Landberg J, Thern E. Using repeated measures to study the contribution of alcohol consumption and smoking to the social gradient in all-cause mortality: Results from the Stockholm Public Health Cohort. *Drug Alcohol Rev* 2023; **42**(7): 1850-9.

170. Ogunlayi F, Coleman PC, Fat LN, Mindell JS, Oyebode O. Trends in socioeconomic inequalities in behavioural non-communicable disease risk factors: analysis of repeated cross-sectional health surveys in England between 2003 and 2019. *Bmc Public Health* 2023; **23**(1): 1442.

171. Diderichsen F, Evans T, Whitehead M. 12The Social Basis of Disparities in Health. Challenging Inequities in Health: From Ethics to Action: Oxford University Press; 2001. p. 0.

172. Diderichsen F, Hallqvist J, Whitehead M. Differential vulnerability and susceptibility: how to make use of recent development in our understanding of mediation and interaction to tackle health inequalities. *Int J Epidemiol* 2019; **48**(1): 268-74.

173. Mäkelä P, Paljärvi T. Do consequences of a given pattern of drinking vary by socioeconomic status? A mortality and hospitalisation follow-up for alcohol-related causes of the Finnish Drinking Habits Surveys. *J Epidemiol Community Health* 2008; **62**(8): 728-33.

174. Christensen HN, Diderichsen F, Hvidtfeldt UA, et al. Joint Effect of Alcohol Consumption and Educational Level on Alcohol-related Medical Events: A Danish Register-based Cohort Study. *Epidemiology* 2017; **28**(6): 872-9.

175. Katikireddi SV, Whitley E, Lewsey J, Gray L, Leyland AH. Socioeconomic status as an effect modifier of alcohol consumption and harm: analysis of linked cohort data. *Lancet Public Health* 2017; **2**(6): E267-E76.

176. Pampel FC, Rogers RG. Socioeconomic status, smoking, and health: A test of competing theories of cumulative advantage. *J Health Soc Behav* 2004; **45**(3): 306-21.

177. Nordahl H, Osler M, Frederiksen BL, et al. Combined effects of socioeconomic position, smoking, and hypertension on risk of ischemic and hemorrhagic stroke. *Stroke* 2014; **45**(9): 2582-7.

178. Veronesi G, Tunstall-Pedoe H, Ferrario MM, et al. Combined effect of educational status and cardiovascular risk factors on the incidence of coronary heart disease and stroke in European cohorts: Implications for prevention. *European Journal of Preventive Cardiology* 2017; **24**(4): 437-45.

179. Hussein M, Diez Roux AV, Mujahid MS, et al. Unequal Exposure or Unequal Vulnerability? Contributions of Neighborhood Conditions and Cardiovascular Risk Factors to Socioeconomic Inequality in Incident Cardiovascular Disease in the Multi-Ethnic Study of Atherosclerosis. *Am J Epidemiol* 2018; **187**(7): 1424-37.

180. Lantz PM, House JS, Mero RP, Williams DR. Stress, life events, and socioeconomic disparities in health: results from the Americans' Changing Lives Study. *J Health Soc Behav* 2005; **46**(3): 274-88.

181. Nabi H, Kivimäki M, Batty GD, et al. Increased risk of coronary heart disease among individuals reporting adverse impact of stress on their health: the Whitehall II prospective cohort study. *European Heart Journal* 2013; **34**(34): 2697-705.

182. von Känel R, Mills PJ, Fainman C, Dimsdale JE. Effects of Psychological Stress and Psychiatric Disorders on Blood Coagulation and Fibrinolysis: A Biobehavioral Pathway to Coronary Artery Disease? *Psychosomatic Medicine* 2001; **63**(4).

183. Karatsoreos IN, Bhagat SM, Bowles NP, Weil ZM, Pfaff DW, McEwen BS. Endocrine and physiological changes in response to chronic corticosterone: a potential model of the metabolic syndrome in mouse. *Endocrinology* 2010; **151**(5): 2117-27.

184. Kraft P, Kraft B. Explaining socioeconomic disparities in health behaviours: A review of biopsychological pathways involving stress and inflammation. *Neuroscience & Biobehavioral Reviews* 2021; **127**: 689-708.

185. Pool U. Socioeconomic inequalities in lifestyle-related health outcomes. *Lancet Public Health* 2019; **4**(2): e85.

186. Belsky DW, Caspi A, Cohen HJ, et al. Impact of early personal-history characteristics on the Pace of Aging: implications for clinical trials of therapies to slow aging and extend healthspan. *Aging Cell* 2017; **16**(4): 644-51.

187. Wagg E, Blyth FM, Cumming RG, Khalatbari-Soltani S. Socioeconomic position and healthy ageing: A systematic review of cross-sectional and longitudinal studies. *Ageing Research Reviews* 2021; **69**: 101365.

188. Raffington L, Belsky DW. Integrating DNA Methylation Measures of Biological Aging into Social Determinants of Health Research. *Current Environmental Health Reports* 2022; **9**(2): 196-210.

189. Shields AE. Epigenetic Signals of How Social Disadvantage "Gets Under The Skin": A Challenge to The Public Health Community. *Epigenomics* 2017; **9**(3): 223-9.

190. Cable N. Life Course Approach in Social Epidemiology: An Overview, Application and Future Implications. *J Epidemiol* 2014; **24**(5): 347-52.

191. Lewer D, Meier P, Beard E, Boniface S, Kaner E. Unravelling the alcohol harm paradox: a population-based study of social gradients across very heavy drinking thresholds. *Bmc Public Health* 2016; **16**: 599.

192. Matthews KA, Kelsey SF, Meilahn EN, Kuller LH, Wing RR. Educational attainment and behavioral and biologic risk factors for coronary heart disease in middle-aged women. *Am J Epidemiol* 1989; **129**(6): 1132-44.

193. Dal Maso L, Torelli N, Biancotto E, et al. Combined effect of tobacco smoking and alcohol drinking in the risk of head and neck cancers: a re-analysis of case–control studies using bi-dimensional spline models. *European Journal of Epidemiology* 2016; **31**(4): 385-93.

194. Chalabaev A, Cheval B, Maltagliati S, Saoudi I, Sniehotta FF. Beyond Individual Cognitions: Time for Intervention Science to Focus on Health Context and Audience. *Journal of Physical Activity and Health* 2023; **20**(6): 465-70.

195. Elwell-Sutton T, Marshall, L., Bibby, J., and Volmert, A. Reframing the conversation on the social determinants of health. Published online: The Health Foundation, 2019.

196. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008; **337**: a1655.

197. Skivington K, Matthews, L., Simpson, S.A., Craig, P., Baird, J., Blazeby, J.M. et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ* 2021; **374**: n2061.

198. Coombes E, Jones AP, Hillsdon M. The relationship of physical activity and overweight to objectively measured green space accessibility and use. *Soc Sci Med* 2010; **70**(6): 816-22.

199. Bukman AJ, Teuscher D, Feskens EJM, van Baak MA, Meershoek A, Renes RJ. Perceptions on healthy eating, physical activity and lifestyle advice: opportunities for adapting lifestyle interventions to individuals with low socioeconomic status. *Bmc Public Health* 2014; **14**(1): 1036.

200. Everson-Hock ES, Johnson M, Jones R, et al. Community-based dietary and physical activity interventions in low socioeconomic groups in the UK: A mixed methods systematic review. *Preventive Medicine* 2013; **56**(5): 265-72.

201. Leese C, Al-Zubaidi H. Urban green and blue spaces for influencing physical activity in the United Kingdom: A narrative review of the policy and evidence. *Lifestyle Medicine* 2024; **5**(1): e96.

202. Campbell M, Fitzpatrick R, Haines A, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000; **321**(7262): 694-6.

203. Bonevski B, Randell M, Paul C, et al. Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups. *BMC Med Res Methodol* 2014; **14**: 42.

204. Francis-Oliviero F, Cambon L, Wittwer J, Marmot M, Alla F. Theoretical and practical challenges of proportionate universalism: a review. *Rev Panam Salud Publica* 2020; **44**: e110.

205. Coupe N, Cotterill S, Peters S. Tailoring lifestyle interventions to low socio-economic populations: a qualitative study. *Bmc Public Health* 2018; **18**(1): 967.

206. Shaw RL, Larkin M, Flowers P. Expanding the evidence within evidence-based healthcare: thinking about the context, acceptability and feasibility of interventions. *Evidence Based Medicine* 2014; **19**(6): 201.

207. Baum F. Researching public health: Behind the qualitative-quantitative methodological debate. *Social Science & Medicine* 1995; **40**: 459-68.

208. Borkan JM. Mixed methods studies: A foundation for primary care research. *Annals of Family Medicine* 2004; **2**(1): 4-6.

209. Pluye P, Hong QN. Combining the power of stories and the power of numbers: Mixed methods research and mixed studies reviews. *Annual Review of Public Health* 2014; **35**: 29-45.

210. Creswell JW, Tashakkori A. Differing perspectives on mixed methods research. *Journal of Mixed Methods Research* 2007; **1**(4): 303-8.

211. Diderichsen F, Andersen I, Manuel C, et al. Health Inequality - determinants and policies. *Scandinavian Journal of Public Health* 2012; **40**(8_suppl): 12-105.

212. Allen N, Sudlow C, Downey P, et al. UK Biobank: Current status and what it means for epidemiology. *Health Policy and Technology* 2012; **1**(3): 123-6.

213. UK Biobank: Protocol for a large-scale prospective epidemiological resource 2007. https://www.ukbiobank.ac.uk/media/gnkeyh2q/study-rationale.pdf.

214. Touchscreen questionnaire ordering, validation and dependencies - UK Biobank. Online: UK Biobank, 2012.

215. UK Biobank: Verbal Interview stage. 2012. https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=100235 (accessed 27/04/2024.

216. Hutcheon JA, Chiolero A, Hanley JA. Random measurement error and regression dilution bias. *BMJ* 2010; **340**: c2289.

217. UK Biobank: Repeat Assessment Data. Version 1.0, 2013.

218. Weiss R. Loneliness: The experience of emotional and social isolation: MIT press; 1975.

219. Russell D, Peplau LA, Cutrona CE. The revised UCLA Loneliness Scale: concurrent and discriminant validity evidence. *J Pers Soc Psychol* 1980; **39**(3): 472-80.

220. Russell D. The measurement of loneliness. *Loneliness: A sourcebook of current theory, research and therapy* 1982; **36**: 81-104.

221. Elovainio M, Hakulinen C, Pulkki-Raback L, et al. Contribution of risk factors to excess mortality in isolated and lonely individuals: an analysis of data from the UK Biobank cohort study. *Lancet Public Health* 2017; **2**(6): e260-e6.

222. Hakulinen C, Pulkki-Raback L, Virtanen M, Jokela M, Kivimaki M, Elovainio M. Social isolation and loneliness as risk factors for myocardial infarction, stroke and mortality: UK Biobank cohort study of 479 054 men and women. *Heart* 2018; **104**(18): 1536-42.

223. Elovainio M, Lahti J, Pirinen M, et al. Association of social isolation, loneliness and genetic risk with incidence of dementia: UK Biobank Cohort Study. *BMJ Open* 2022; **12**(2): e053936.

224. Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A Short Scale for Measuring Loneliness in Large Surveys: Results From Two Population-Based Studies. *Res Aging* 2004; **26**(6): 655-72.

225. Hakulinen C, Pulkki-Raback L, Jokela M, et al. Structural and functional aspects of social support as predictors of mental and physical health trajectories: Whitehall II cohort study. *J Epidemiol Community Health* 2016; **70**(7): 710-5.

226. Alcohol units - NHS Alcohol advice. 2021. https://www.nhs.uk/live-well/alcohol-advice/calculating-alcohol-units/2024).

227. What is a unit of alcohol - Chapter 12: Alcohol - Delivering better oral health: an evidence-based toolkit for prevention, 2014.

228. Twohig-Bennett C, Jones A. The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. *Environmental Research* 2018; **166**: 628-37.

229. Miguet M, Venetis S, Rukh G, Lind L, Schiöth HB. Time spent outdoors and risk of myocardial infarction and stroke in middle and old aged adults: Results from the UK Biobank prospective cohort. *Environmental Research* 2021; **199**: 111350.

230. Foster HME, Ho, F.K., Sattar, N., Welsh, P., Pell, J.P., Gill, J.M.R., Gray, S.R., & Celis-Morales, C.A. . Understanding How Much TV is Too Much: A Nonlinear Analysis of the Association Between Television Viewing Time and Adverse Health Outcomes. . *Mayo Clinic Proceedings* 2020; **95**(11): 2429-41.

231. Patterson R, McNamara E, Tainio M, et al. Sedentary behaviour and risk of all-cause, cardiovascular and cancer mortality, and incident type 2 diabetes: a systematic review and dose response meta-analysis. *Eur J Epidemiol* 2018; **33**(9): 811-29.

232. Jani BD, McQueenie R, Nicholl BI, et al. Association between patterns of alcohol consumption (beverage type, frequency and consumption with food) and risk of adverse health outcomes: a prospective cohort study. *BMC Med* 2021; **19**(1): 8.

233. Anderson BO, Berdzuli N, Ilbawi A, et al. Health and cancer risks associated with low levels of alcohol consumption. *Lancet Public Health* 2023; **8**(1): e6-e7.

234. UK Biobank: Mortality data: linkage to death registries. Version 3.0, 2023.

235. International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10). WHO; 2019.

236. Zhang Y-B, Pan X-P, Chen J, et al. Combined lifestyle factors, all-cause mortality and cardiovascular disease: a systematic review and meta-analysis of prospective cohort studies. *J Epidemiol Commun H* 2021; **75**(1): 92.

237. UK Biobank: Hospital inpatient data. Version 4.0, 2023.

238. UK Biobank: Primary Care Linked Data. Version 2.0, 2023.

239. Greenland S, Senn SJ, Rothman KJ, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol* 2016; **31**(4): 337-50.

240. Ioannidis JPA. Why Most Published Research Findings Are False. *PLOS Medicine* 2005; **2**(8): e124.

241. Senn S. Two cheers for P-values? *J Epidemiol Biostat* 2001; **6**(2): 193-204; discussion 5-10.

242. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004; **15**(5): 615-25.

243. Toledano MB, Smith RB, Brook JP, Douglass M, Elliott P. How to Establish and Follow up a Large Prospective Cohort Study in the 21st Century - Lessons from UK COSMOS. *PLOS ONE* 2015; **10**(7): e0131521.

244. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol* 2013; **42**(4): 1012-4.

245. Manolio TA, Collins R. Enhancing the feasibility of large cohort studies. *JAMA* 2010; **304**(20): 2290-1.

246. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015; **12**(3): e1001779.

247. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am J Epidemiol* 2017; **186**(9): 1026-34.

248. Munafo MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol* 2018; **47**(1): 226-35.

249. Brayne C, Moffitt TE. The limitations of large-scale volunteer databases to address inequalities and global challenges in health and aging. *Nature Aging* 2022; **2**(9): 775-83.

250. Glossary of terms - Catalogue of Bias Collaboration. https://catalogofbias.org/about/2024).

251. Lee H, Aronson, JK., Nunan, D. Catalogue of bias collaboration, Collider bias. 2019. https://catalogofbias.org/biases/collider-bias/.

252. Clark T, Foster L, Sloan L, Bryman A. Bryman's social research methods. Sixth / Tom Clark, Liam Foster, Luke Sloan, Alan Bryman ; advisor, Elena Vacchelli. ed. Oxford: Oxford University Press; 2021.

253. Norman Blaikie B. Designing social research: the logic of anticipation: Polity Press; 2019.

254. Albert H. Karl Popper, critical rationalism, and the Positivist Dispute. *Journal of Classical Sociology* 2015; **15**(2): 209-19.

255. Miller D. Critical rationalism: A restatement and defence: Open court; 2015.

256. Bryman A. The Debate about Quantitative and Qualitative Research: A Question of Method or Epistemology? *The British Journal of Sociology* 1984; **35**(1): 75-92.

257. Holloway I, Todres L. The Status of Method: Flexibility, Consistency and Coherence. *Qualitative Research* 2003; **3**(3): 345-57.

258. Activity in the NHS - The NHS in a nutshell. Online: The King's Fund, 2023.

259. Mercer SW, Fitzpatrick B, Grant L, et al. Effectiveness of Community-Links Practitioners in Areas of High Socioeconomic Deprivation. *Ann Fam Med* 2019; **17**(6): 518-25.

260. National Institute for Health and Care Research: Payment guidance for researchers and professionals - version 1.5. 2022. https://www.nihr.ac.uk/documents/payment-guidance-for-researchers-and-professionals/27392.

261. Braun V, & Clarke, V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006; **3**(2): 77-101.

262. Braun V, Clarke V. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 2019; **11**(4): 589-97.

263. Byrne D. A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality & Quantity* 2022; (56): 1391–412.

264. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesising qualitative and quantitative evidence: A review of possible methods. *Journal of Health Services Research & Policy* 2005; **10**(1): 45-53.

265. Dixon-Woods M, Cavers D, Agarwal S, et al. Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Medical Research Methodology* 2006; **6**(1): 35.

266. Depraetere J, Vandeviver C, Keygnaert I, Beken TV. The critical interpretive synthesis: an assessment of reporting practices. *International Journal of Social Research Methodology* 2021; **24**(6): 669-89.

267. Stanaway JD, Afshin A, Gakidou E, et al. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018; **392**(10159): 1923-94.

268. Hart CL, Davey Smith G, Gruer L, Watt GC. The combined effect of smoking tobacco and drinking alcohol on cause-specific mortality: a 30 year cohort study. *Bmc Public Health* 2010; **10**: 789.

269. Ford ES, Bergmann MM, Boeing H, Li C, Capewell S. Healthy lifestyle behaviors and all-cause mortality among adults in the United States. *Prev Med* 2012; **55**(1): 23-7.

270. Behrens G, Fischer B, Kohler S, Park Y, Hollenbeck AR, Leitzmann MF. Healthy lifestyle behaviors and decreased risk of mortality in a large prospective study of U.S. women and men. *Eur J Epidemiol* 2013; **28**(5): 361-72.

271. McNamara CL, Balaj M, Thomson KH, Eikemo TA, Solheim EF, Bambra C. The socioeconomic distribution of non-communicable diseases in Europe: findings from the European Social Survey (2014) special module on the social determinants of health. *Eur J Public Health* 2017; **27**: 22-6.

272. Inequalities in Health: Report of a Research Working Group ('The Black report'), Department of Health and Social Security, London (UK). 1980.

273. Birch S, Jerrett M, Eyles J. Heterogeneity in the determinants of health and illness: the example of socioeconomic status and smoking. *Soc Sci Med* 2000; **51**(2): 307-17.

274. Foster HME, Ho FK, Mair FS, et al. The association between a lifestyle score, socioeconomic status, and COVID-19 outcomes within the UK Biobank cohort. *Bmc Infect Dis* 2022; **22**(1).

275. Fiorito G, Polidoro S, Dugue PA, et al. Social adversity and epigenetic aging: a multi-cohort study on socioeconomic differences in peripheral blood DNA methylation. *Sci Rep* 2017; **7**(1): 16266.

276. Mackenbach JD, Brage S, Forouhi NG, Griffin SJ, Wareham NJ, Monsivais P. Does the importance of dietary costs for fruit and vegetable intake vary by socioeconomic position? *Brit J Nutr* 2015; **114**(9): 1464-70.

277. Burgoine T, Mackenbach JD, Lakerveld J, et al. Interplay of Socioeconomic Status and Supermarket Distance Is Associated with Excess Obesity Risk: A UK Cross-Sectional Study. *Int J Env Res Pub He* 2017; **14**(11).

278. Pepper GV, Nettle D. The behavioural constellation of deprivation: Causes and consequences. *Behav Brain Sci* 2017; **40**: e314.

279. Wardle J, Steptoe A. Socioeconomic differences in attitudes and beliefs about healthy lifestyles. *J Epidemiol Commun H* 2003; **57**(6): 440-3.

280. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; **339**: b2535.

281. Foster H. 2022_12_22_PRISMA_2020_abstract_checklist.docx. 2022. https://doi.org/10.6084/m9.figshare.21770651.v1.

282. Foster H. PRISMA 2020 Checklist. 2022. https://doi.org/10.6084/m9.figshare.21770657.v1.

283. Foster H, Polz, P., Mair, F., Gill, J., O'Donnell, C. Understanding the impact of socioeconomic status on the association between combined lifestyle factors and adverse health outcomes: a systematic review. PROSPERO 2020 CRD42020172588. PROSPERO - International prospective register of systematic reviews; 2020.

284. Foster H. Foster, Hamish (2023). 2023_10_18_SES_lifestyle_systematic_rv_SUPPORTING_INFORMATION_corrected.docx. figshare. Dataset. . 2023. https://figshare.com/articles/dataset/2023_10_18_SES_lifestyle_systematic_rv_SUPPORTING_INFORMATION_corrected_docx/24649755.

285. McKenzie J, Brennan, SE., Ryan, RE., Thomson, HJ., Johnston, RV., Thomas, J.,. Chapter 3: Defining the criteria for including studies and how they will be grouped for the synthesis. In: Cochrane Handbook for Systematic Reviews of Interventions version 6.0. , 2019.

286. Wells G, Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses: Ottawa Hospital Research Institute [Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp], 2019.

287. Higgins JPT, Morgan RL, Rooney AA, et al. A tool to assess risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E). *Environ Int* 2024; **186**: 108602.

288. Luchini C, Stubbs B, Solmi M, Veronese N. Assessing the quality of studies in meta-analyses: Advantages and limitations of the Newcastle Ottawa Scale. *World Journal of Meta-Analysis* 2017; **5**(4): 80-4.

289. Andersen SW, Zheng W, Sonderman J, et al. Combined Impact of Health Behaviors on Mortality in Low-Income Americans. *Am J Prev Med* 2016; **51**(3): 344-55.

290. Eguchi E, Iso H, Honjo K, Yatsuya H, Tamakoshi A. No modifying effect of education level on the association between lifestyle behaviors and cardiovascular mortality: the Japan Collaborative Cohort Study. *Sci Rep-Uk* 2017; **7**.

291. Andersen SW, Blot WJ, Shu XO, et al. Associations Between Neighborhood Environment, Health Behaviors, and Mortality. *Am J Prev Med* 2018; **54**(1): 87-95.

292. Choi SH, Stommel M, Ling J, Noonan D, Chung J. The Impact of Smoking and Multiple Health Behaviors on All-Cause Mortality. *Behav Med* 2022; **48**(1): 10-7.

293. Zhang YB, Chen C, Pan XF, et al. Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: two prospective cohort studies. *BMJ* 2021; **373**: n604.

294. About the National Health Interview Survey. https://www.cdc.gov/nchs/nhis/about_nhis.htm (accessed 31/03/2022.

295. About the National Health and Nutrition Examination Survey. https://www.cdc.gov/nchs/nhanes/about_nhanes.htm (accessed 31/03/2022.

296. Andersson T, Alfredsson L, Kallberg H, Zdravkovic S, Ahlbom A. Calculating measures of biological interaction. *Eur J Epidemiol* 2005; **20**(7): 575-9.

297. Ejiogu N, Norbeck JH, Mason MA, Cromwell BC, Zonderman AB, Evans MK. Recruitment and retention strategies for minority or poor clinical research participants: lessons from the Healthy Aging in Neighborhoods of Diversity across the Life Span study. *Gerontologist* 2011; **51 Suppl 1**(Suppl 1): S33-45.

298. Routen A, Bodicoat D, Willis A, Treweek S, Paget S, Khunti K. Tackling the lack of diversity in health research. *British Journal of General Practice* 2022; **72**(722): 444-7.

299. Ho FK, Gray SR, Welsh P, et al. Associations of fat and carbohydrate intake with cardiovascular disease and mortality: prospective cohort study of UK Biobank participants. *BMJ* 2020; **368**: m688.

300. Foster HME, Ho FK, Sattar N, et al. Understanding How Much TV is Too Much: A Nonlinear Analysis of the Association Between Television Viewing Time and Adverse Health Outcomes. *Mayo Clin Proc* 2020; **95**(11): 2429-41.

301. Yusuf S, Joseph P, Rangarajan S, et al. Modifiable risk factors, cardiovascular disease, and mortality in 155 722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. *Lancet* 2020; **395**(10226): 795-808.

302. Buck D, Frosini, F. Clustering of unhealthy behaviours over time. Implications for policy and practice - The Kings Fund, 2012.

303. Physical Activity Guidelines for Americans, 2nd edition. In: Washington DC U, editor. Published online: U.S. Government Department of Health and Human Services; 2018.

304. UK Chief Medical Officers' physical activity guidelines 2020, 2020.

305. The Eatwell Guide (England, Wales, & Northern Ireland's recommendations on eating healthily), 2016.

306. Evans H, Buck, D. Tackling multiple unhealthy risk factors. Emerging lessons from practice - The Kings Fund, 2018.

307. Kabisch N. The Influence of Socio-economic and Socio-demographic Factors in the Association Between Urban Green Space and Health. . In: Marselle M. SJ, Korn H., Irvine K., Bonn A., ed. Biodiversity and Health in the Face of Climate Change: Springer, Cham. ; 2019: 91-119.

308. Blaxter M. Health and lifestyles. London ; New York: Routledge; 1990.

309. Booth A. Over 85% of included studies in systematic reviews are on MEDLINE. *J Clin Epidemiol* 2016; **79**: 165-6.

310. Lefebvre C, Glanville, J., Briscoe, S., Featherstone, R., Littlewood, A., Marshall, C., Metzendorf, M-I., Noel-Storr, A., Paynter, R., Rader, T., Thomas, J., Wieland, LS. . Technical Supplement to Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston MS, Li T,

Page MJ, Welch VA (eds). Cochrane Handbook for Systematic Reviews of Interventions Version 6.3 (updated February 2022): Cochrane; 2022.

311. Tran CL, Choi KS, Kim SY, Oh JK. Individual and joint effect of socioeconomic status and lifestyle factors on cancer in Korea. *Cancer Med* 2023; **12**(16): 17389-402.

312. Zhang Y-B, Li Y, Geng T-T, et al. Overall lifestyles and socioeconomic inequity in mortality and life expectancy in China: the China health and nutrition survey. *Age and Ageing* 2022; **51**(7).

313. Diseases GBD, Injuries C. 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. *Lancet* 2020; **396**(10258): 1204-22.

314. Peters R, Booth A, Rockwood K, Peters J, D'Este C, Anstey KJ. Combining modifiable risk factors and risk of dementia: a systematic review and meta-analysis. *BMJ Open* 2019; **9**(1): e022846.

315. Wakasugi M, Kazama JJ, Yamamoto S, Kawamura K, Narita I. A combination of healthy lifestyle factors is associated with a decreased incidence of chronic kidney disease: a population-based cohort study. *Hypertens Res* 2013; **36**(4): 328-33.

316. Holt-Lunstad J. Why Social Relationships Are Important for Physical Health: A Systems Approach to Understanding and Modifying Risk and Protection. *Annu Rev Psychol* 2018; **69**: 437-58.

317. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. *Perspect Psychol Sci* 2015; **10**(2): 227-37.

318. Fortmann AL, Gallo LC. Social support and nocturnal blood pressure dipping: a systematic review. *Am J Hypertens* 2013; **26**(3): 302-10.

319. Xiong Y, Hong H, Liu C, Zhang YQ. Social isolation and the brain: effects and mechanisms. *Mol Psychiatry* 2023; **28**(1): 191-201.

320. Kobayashi LC, Steptoe A. Social Isolation, Loneliness, and Health Behaviors at Older Ages: Longitudinal Cohort Study. *Ann Behav Med* 2018; **52**(7): 582-93.

321. Lauder W, Mummery K, Jones M, Caperchione C. A comparison of health behaviours in lonely and non-lonely populations. *Psychol Health Med* 2006; **11**(2): 233-45.

322. Stokes JE. Social integration, perceived discrimination, and self-esteem in mid- and later life: intersections with age and neuroticism. *Aging Ment Health* 2019; **23**(6): 727-35.

323. Shankar A, McMunn A, Banks J, Steptoe A. Loneliness, social isolation, and behavioral and biological health indicators in older adults. *Health Psychol* 2011; **30**(4): 377-85.

324. Bevilacqua G, Jameson, K.A., Zhang, J. et al. Relationships between non-communicable disease, social isolation and frailty in community dwelling adults in later life: findings from the Hertfordshire Cohort Study. . *Aging Clinical and Experimental Research* 2022; **34**: 105-12.

325. Hajek A, Kretzler B, Konig HH. Multimorbidity, Loneliness, and Social Isolation. A Systematic Review. *Int J Environ Res Public Health* 2020; **17**(22): 8688.

326. Coyle CE, Dugan E. Social isolation, loneliness and health among older adults. *J Aging Health* 2012; **24**(8): 1346-63.

327. Newall NEG, & Menec, V. H. . Loneliness and social isolation of older adults: Why it is important to examine these social aspects together. *Journal of Social and Personal Relationships* 2019; **36**(3): 925–39.

328. Kung CSJ, Kunz JS, Shields MA. Economic Aspects of Loneliness in Australia. *Aust Econ Rev* 2021; **54**(1): 147-63.

329. Kung CSJ, Pudney SE, Shields MA. Economic gradients in loneliness, social isolation and social support: Evidence from the UK Biobank. *Soc Sci Med* 2022; **306**: 115122.

330. Mansfield L, Victor C, Meads C, et al. A Conceptual Review of Loneliness in Adults: Qualitative Evidence Synthesis. *Int J Environ Res Public Health* 2021; **18**(21): 11522.

331. Zhen Z, Feng, Q., and Gu, D. The Impacts of Unmet Needs for Long-Term Care on Mortality Among Older Adults in China. *Journal of Disability Policy Studies* 2015; **25**(4): 243-51.

332. Lee H, Singh GK. Social Isolation and All-Cause and Heart Disease Mortality Among Working-Age Adults in the United States: The 1998-2014 NHIS-NDI Record Linkage Study. *Health Equity* 2021; **5**(1): 750-61.

333. Pantell M, Rehkopf D, Jutte D, Syme SL, Balmes J, Adler N. Social isolation: a predictor of mortality comparable to traditional clinical risk factors. *Am J Public Health* 2013; **103**(11): 2056-62.

334. Steptoe A, Shankar A, Demakakos P, Wardle J. Social isolation, loneliness, and all-cause mortality in older men and women. *Proc Natl Acad Sci U S A* 2013; **110**(15): 5797-801.

335. UK Biobank: Protocol for a large-scale prospective epidemiological resource https://www.ukbiobank.ac.uk/media/gnkeyh2q/study-rationale.pdf (accessed 01/12/2022 2022).

336. Health Survey for England 2017 - Summary of key findings. 2018.

337. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ) https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/ipaq_analysis.pdf (accessed 01/12/2022.

338. UK Chief Medical Officers' physical activity guidelines. . In: Care. UDoHaS, editor. Published online; 2019.

339. Jani BD, Hanlon P, Nicholl BI, et al. Relationship between multimorbidity, demographic factors and mortality: findings from the UK Biobank cohort. *BMC Med* 2019; **17**(1): 74.

340. Victor C, Sulliva, M. P., Woodbridge, R., and Thomas, M.,. Dancing with Loneliness in Later Life: A Pilot Study Mapping Seasonal Variations. *The Open Psychology Journal* 2015; **8**(1): 97-104.

341. Townsend B, Phillimore, P., Beattie, A. . Health and deprivation: inequality and the North. 1 ed. London: Routledge; 1988.

342. Myers R. Classical and modern regression with applications 2nd ed. ed. Boston: Duxbury Press, ; 1990.

343. Fox JaM, G. Generalized Collinearity Diagnostics. *J Am Stat Assoc* 1992; **87**(417): 178-83.

344. Knol MJ, VanderWeele TJ. Recommendations for presenting analyses of effect modification and interaction. *Int J Epidemiol* 2012; **41**(2): 514-20.

345. Dreyer K, Steventon A, Fisher R, Deeny SR. The association between living alone and health care utilisation in older adults: a retrospective cohort study of electronic health records from a London general practice. *BMC Geriatr* 2018; **18**(1): 269.

346. Siette J, Cassidy M, Priebe S. Effectiveness of befriending interventions: a systematic review and meta-analysis. *BMJ Open* 2017; **7**(4): e014304.

347. Friedman HS, Tucker JS, Schwartz JE, et al. Psychosocial and behavioral predictors of longevity. The aging and death of the "termites". *Am Psychol* 1995; **50**(2): 69-78.

348. Bookwala J, Gaugler T. Relationship quality and 5-year mortality risk. *Health Psychol* 2020; **39**(8): 633-41.

349. Uhing A, Williams JS, Garacci E, Egede LE. Gender differences in the relationship between social support and strain and mortality among a national sample of adults. *J Behav Med* 2021; **44**(5): 673-81.

350. Tanskanen J, Anttila T. A Prospective Study of Social Isolation, Loneliness, and Mortality in Finland. *Am J Public Health* 2016; **106**(11): 2042-8.

351. Yu B, Steptoe A, Chen LJ, Chen YH, Lin CH, Ku PW. Social Isolation, Loneliness, and All-Cause Mortality in Patients With Cardiovascular Disease: A 10-Year Follow-up Study. *Psychosom Med* 2020; **82**(2): 208-14.

352. Beller J, Wagner A. Loneliness, social isolation, their synergistic interaction, and mortality. *Health Psychol* 2018; **37**(9): 808-13.

353. Stokes AC, Xie W, Lundberg DJ, Glei DA, Weinstein MA. Loneliness, social isolation, and all-cause mortality in the United States. *SSM Ment Health* 2021; **1**: 100014.

354. Galvez-Hernandez P, Gonzalez-de Paz L, Muntaner C. Primary care-based interventions addressing social isolation and loneliness in older people: a scoping review. *BMJ Open* 2022; **12**(2): e057729.

355. Foster H, Polz P, Gill J, Celis-Morales C, Mair F, O'Donnell C. The influence of socioeconomic status on the association between unhealthy lifestyle factors and adverse health outcomes: a systematic review [version 2; peer review: 1 approved, 1 approved with reservations]. *Wellcome Open Research* 2023; **8**(55).

356. Lloyd-Jones DM, Allen NB, Anderson CAM, et al. Life's Essential 8: Updating and Enhancing the American Heart Association's Construct of Cardiovascular Health: A Presidential Advisory From the American Heart Association. *Circulation* 2022; **146**(5): e18-e43.

357. South E, Rodgers, M., Wright, K., Whitehead, M., Sowden, A. Reducing lifestyle risk behaviours in disadvantaged groups in high-income countries: A scoping review of systematic reviews. *Preventive Medicine* 2022; **154**: 106916.

358. Peters SAE, Woodward, M. Sex and gender matter in cardiovascular disease and beyond. *Heart* 2022; **108**(13): 994.

359. Ho FK, Gray SR, Welsh P, et al. Ethnic differences in cardiovascular risk: examining differential exposure and susceptibility to risk factors. *BMC Medicine* 2022; **20**(1): 149.

360. Ibsen DB, Søgaard K, Sørensen LH, et al. Modifiable Lifestyle Recommendations and Mortality in Denmark: A Cohort Study. *Am J Prev Med* 2021; **60**(6): 792-801.

361. Foster HME, Gill JMR, Mair FS, et al. Social connection and mortality in UK Biobank: a prospective cohort analysis. *BMC Med* 2023; **21**(1): 384.

362. Filippini T, Malavolti M, Whelton PK, Vinceti M. Sodium Intake and Risk of Hypertension: A Systematic Review and Dose–Response Meta-analysis of Observational Cohort Studies. *Current Hypertension Reports* 2022; **24**(5): 133-44.

363. Hakulinen C, Pulkki-Råback L, Virtanen M, Jokela M, Kivimäki M, Elovainio M. Social isolation and loneliness as risk factors for myocardial infarction, stroke and mortality: UK Biobank cohort study of 479 054 men and women. *Heart* 2018; **104**(18): 1536-42.

364. Yousaf S, and Bonsall, A. Townsend Deprivation Scores Report - UK Townsend Deprivation Scores from 2011 census data (compares 2001 and 2011). UK Data Service, 2017.

365. Chalise P, Chicken E, McGee D. Baseline age effect on parameter estimates in Cox models. *Journal of Statistical Computation and Simulation* 2012; **82**(12): 1767-74.

366. Colpani V, Baena CP, Jaspers L, et al. Lifestyle factors, cardiovascular disease and all-cause mortality in middle-aged and elderly women: a systematic review and meta-analysis. *European Journal of Epidemiology* 2018; **33**(9): 831-45.

367. Woodward M, Brindle P, Tunstall-Pedoe H, estimation Sgor. Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). *Heart* 2007; **93**(2): 172-6.

368. Hippisley-Cox J, Coupland, C., Vinogradova, Y., Robson, J., May, M., Brindle, P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ* 2007; **335**(7611): 136.

369. Hunt K, Wyke S, Gray CM, et al. A gender-sensitised weight loss and healthy living programme for overweight and obese men delivered by Scottish Premier League football clubs (FFIT): a pragmatic randomised controlled trial. *Lancet* 2014; **383**(9924): 1211-21.

370. Gao Y, Chen Y, Hu M, et al. Lifestyle trajectories and ischaemic heart diseases: a prospective cohort study in UK Biobank. *European Journal of Preventive Cardiology* 2023; **30**(5): 393-403.

371. Brauer M, Roth GA, Aravkin AY, et al. Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *The Lancet* 2024; **403**(10440): 2162-203.

372. Salman D, Farooqi M, McGregor A, Majeed A. Time spent being sedentary: an emerging risk factor for poor health. *British Journal of General Practice* 2019; **69**(683): 278-9.

373. Shams-White MM, Pannucci TE, Lerman JL, et al. Healthy Eating Index-2020: Review and Update Process to Reflect the Dietary Guidelines for Americans, 2020-2025. *J Acad Nutr Diet* 2023; **123**(9): 1280-8.

374. Loef M, Walach H. The combined effects of healthy lifestyle behaviors on all cause mortality: a systematic review and meta-analysis. *Prev Med* 2012; **55**(3): 163-70.

375. Foster HME, Polz P, Gill JMR, Celis-Morales C, Mair FS, O'Donnell CA. The influence of socioeconomic status on the association between unhealthy lifestyle factors and adverse health outcomes: a systematic review [version 2; peer review: 2 approved, 1 approved with reservations]. *Wellcome Open Research* 2023; **8**(55).

376. Wyper GMA, Mackay DF, Fraser C, et al. Evaluating the impact of alcohol minimum unit pricing on deaths and hospitalisations in Scotland: a controlled interrupted time series study. *Lancet* 2023; **401**(10385): 1361-70.

377. Cobiac LJ, Rogers NT, Adams J, et al. Impact of the UK soft drinks industry levy on health and health inequalities in children and adolescents in

England: An interrupted time series analysis and population health modelling study. *PLoS Med* 2024; **21**(3): e1004371.

378. Carey G, Crammond B, De Leeuw E. Towards health equity: a framework for the application of proportionate universalism. *International Journal for Equity in Health* 2015; **14**: 81.

379. Blane D, Lunan, C., Bogie, J., Albanese, A., Henderson, D., and Mercer, S. Tackling the inverse care law in Scottish general practice: policies, interventions and the Scottish Deep End Project. The Health Foundation: University of Glasgow, University of Edinburgh, 2024.

380. Holman D, Lynch R, Reeves A. How do health behaviour interventions take account of social context? A literature trend and co-citation analysis. *Health* 2018; **22**(4): 389-410.

381. Williams GH. The determinants of health: structure, context and agency. *Sociology of Health & Illness* 2003; **25**: 131-54.

382. Mollborn S, Lawrence EM, Saint Onge JM. Contributions and challenges in health lifestyles research. *J Health Soc Behav* 2021; **62**(3): 388-403.

383. Jebb SA, Aveyard P. 'Willpower' is not enough: time for a new approach to public health policy to prevent obesity. *BMC Medicine* 2023; **21**(1): 89.

384. Popay J, Whitehead M, Hunter DJ. Injustice is killing people on a large scale—but what is to be done about it? *Journal of Public Health* 2010; **32**(2): 148-9.

385. Williams O, Fullagar S. Lifestyle drift and the phenomenon of 'citizen shift' in contemporary UK health policy. *Sociology of Health & Illness* 2019; **41**(1): 20-35.

386. Ulucanlar S, Lauber K, Fabbri A, et al. Corporate Political Activity: Taxonomies and Model of Corporate Influence on Public Policy. *Int J Health Policy Manag* 2023; **12**: 7292.

387. Lawrence RG. Framing Obesity: The Evolution of News Discourse on a Public Health Issue. *Harvard International Journal of Press/Politics* 2004; **9**(3): 56-75.

388. Brookes G, and Baker, P. Fear and responsibility: Discourses of obesity and risk in the UK press. *Journal of Risk Research* 2022; **25** (3): 363-78.

389. Frohlich KL, Potvin L. Transcending the known in public health practice: The inequality paradox: The population approach and vulnerable populations. *American Journal of Public Health* 2008; **98**(2): 216-21.

390. Link BG, Phelan J. Social conditions as fundamental causes of disease. *J Health Soc Behav* 1995; **Extra Issue**: 80-94.

391. Phelan JC, Link BG, Tehranifar P. Social conditions as fundamental causes of health inequalities: Theory, evidence, and policy implications. *J Health Soc Behav* 2010; **51**(Suppl 1): S28-S40.

392. Koshy P, Mackenzie M, Leslie W, Lean M, Hankey C. Eating the elephant whole or in slices: views of participants in a smoking cessation intervention trial on multiple behaviour changes as sequential or concurrent tasks. *Bmc Public Health* 2012; **12**(1): 500.

393. ONS. Standard Occupational Classification 2000: SOC 2000. 2000. https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/socarchive (accessed 06/04 2024).

394. Scottish Government. The 2018 General Medical Services Contract in Scotland. Edinburgh; 2017.

395. Foster HME, Ho FK, Mair FS, et al. The association between a lifestyle score, socioeconomic status, and COVID-19 outcomes within the UK Biobank cohort. *Bmc Infect Dis* 2022; **22**: 273.

396. Foster HME, Gill JMR, Mair FS, et al. Social connection and mortality in UK Biobank: a prospective cohort analysis. *BMC Medicine* 2023; **21**: 384.

397. Braun V, Clarke, V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology* 2020; **18**(3): 328-52.

398. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. *Am J Health Promot* 1997; **12**(1): 38-48.

399. Watts S, Lloyd-Williams F, Bromley H, Capewell S. Putting a price on healthy eating: public perceptions of the need for further food pricing policies in the UK. *Journal of Public Health* 2023; **45**(4): e722-e8.

400. Rose G. Sick individuals and sick populations. *Int J Epidemiol* 2001; **30**(3): 427-32; discussion 33-4.

401. Spratt TJR, Hajizadeh, A., Heath, L., Kebbe, M., Aveyard, P. Conceptualising lifestyle "choices:" A qualitative study of GP attitudes towards patients living with "obesity" in the UK. *SSM - Qualitative Research in Health* 2023; **4**: 100325.

402. Powell K, Thurston M, Bloyce D. Theorising lifestyle drift in health promotion: explaining community and voluntary sector engagement practices in disadvantaged areas. *Critical Public Health* 2017; **5**: 554-65.

403. Duggan CP, Kurpad A, Stanford FC, Sunguya B, Wells JC. Race, ethnicity, and racism in the nutrition literature: an update for 2020. *Am J Clin Nutr* 2020; **112**(6): 1409-14.

404. Rodrigues YE, Fanton M, Novossat RS, Canuto R. Perceived racial discrimination and eating habits: a systematic review and conceptual models. *Nutr Rev* 2022; **80**(7): 1769-86.

405. Bailey ZD, Krieger N, Agenor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. *Lancet* 2017; **389**(10077): 1453-63.

406. Marteau TM, Rutter H, Marmot M. Changing behaviour: an essential component of tackling health inequalities. *Bmj* 2021; **372**: n332.

407. Blane DN, MacDonald S, O'Donnell CA. Patient and practitioner views on cancer risk discussions in primary care: a qualitative study. *BJGP Open* 2022; **6**(1).

408. O'Donnell C, Hanlon, P., Blane, D., Macdonald, S., Williamson, A., Mair, F. We should consider prevention burden in our approach to tackling NCDs. the bmj opinion; 2020.

409. Frohlich KL, Corin E, Potvin L. A theoretical proposal for the relationship between context and disease. *Sociology of Health & Illness* 2001; **23**(6): 776-97.

410. Capewell S, Capewell A. An effectiveness hierarchy of preventive interventions: neglected paradigm or self-evident truth? *J Public Health (Oxf)* 2018; **40**(2): 350-8.

411. Rothman KJ. Analysis of Interaction. Modern epidemiology. 4th ed. Philadelphia: Wolters Kluwer; 2021.

412. VanderWeele TJ, Knol MJ. A Tutorial on Interaction. 2014; **3**(1): 33-72.

413. Rothman KJ, Greenland S, Walker AM. Concepts of interaction. *Am J Epidemiol* 1980; **112**(4): 467-70.

414. Mercer SW, Higgins M, Bikker AM, et al. General Practitioners' Empathy and Health Outcomes: A Prospective Observational Study of Consultations in Areas of High and Low Deprivation. *Ann Fam Med* 2016; **14**(2): 117-24.

415. Hojat M, Louis DZ, Markham FW, Wender R, Rabinowitz C, Gonnella JS. Physicians' Empathy and Clinical Outcomes for Diabetic Patients. *Academic Medicine* 2011; **86**(3).

416. Hansen CB, Pavlovic KMH, Sondergaard J, Thilsing T. Does GP empathy influence patient enablement and success in lifestyle change among high risk patients? *BMC Family Practice* 2020; **21**(1): 159.

417. Smith KE, Anderson R. Understanding lay perspectives on socioeconomic health inequalities in Britain: a meta-ethnography. *Sociol Health Illn* 2018; **40**(1): 146-70.

418. Kaplan GA, Lazarus NB, Cohen RD, Leu DJ. Psychosocial factors in the natural history of physical activity. *Am J Prev Med* 1991; **7**(1): 12-7.

419. Locher JL, Ritchie CS, Roth DL, Baker PS, Bodner EV, Allman RM. Social isolation, support, and capital and nutritional risk in an older sample: ethnic and gender differences. *Soc Sci Med* 2005; **60**(4): 747-61.

420. Hanna K, Cross J, Nicholls A, Gallegos D. The association between loneliness or social isolation and food and eating behaviours: A scoping review. *Appetite* 2023; **191**: 107051.

421. Umberson D, Donnelly R. Social Isolation: An Unequally Distributed Health Hazard. *Annu Rev Sociol* 2023; **49**(1): 379-99.

422. Algren MH, Ekholm O, Nielsen L, Ersboll AK, Bak CK, Andersen PT. Social isolation, loneliness, socioeconomic status, and health-risk behaviour in deprived neighbourhoods in Denmark: A cross-sectional study. *SSM Popul Health* 2020; **10**: 100546.

423. Weyers S, Dragano N, Mobus S, et al. Poor social relations and adverse health behaviour: stronger associations in low socioeconomic groups? *Int J Public Health* 2010; **55**(1): 17-23.

424. Ahn J, Falk EB, Kang Y. Relationships between physical activity and loneliness: A systematic review of intervention studies. *Current Research in Behavioral Sciences* 2024; **6**: 100141.

425. de Heer C, Bi S, Finkenauer C, Alink L, Maes M. The Association Between Child Maltreatment and Loneliness Across the Lifespan: A Systematic Review and Multilevel Meta-Analysis. *Child Maltreatment* 2022; **29**(2): 388-404.

426. Hanlon P, McCallum M, Jani BD, McQueenie R, Lee D, Mair FS. Association between childhood maltreatment and the prevalence and complexity of multimorbidity: A cross-sectional analysis of 157,357 UK Biobank participants. *J Comorb* 2020; **10**: 2235042x10944344.

427. Walsh D, McCartney G, Smith M, Armour G. Relationship between childhood socioeconomic position and adverse childhood experiences (ACEs): a systematic review. *J Epidemiol Commun H* 2019; **73**(12): 1087.

428. Jamalishahni T, Turrell G, Foster S, Davern M, Villanueva K. Neighbourhood socio-economic disadvantage and loneliness: the contribution of green space quantity and quality. *Bmc Public Health* 2023; **23**(1): 598.

429. Greiner A, Clegg Smith K, Guallar E. Something fishy? News media presentation of complex health issues related to fish consumption guidelines. *Public Health Nutr* 2010; **13**(11): 1786-94.

430. Clark D, Nagler RH, Niederdeppe J. Confusion and nutritional backlash from news media exposure to contradictory information about carbohydrates and dietary fats. *Public Health Nutr* 2019; **22**(18): 3336-48.

431. Ihekweazu C. Is Coffee the Cause or the Cure? Conflicting Nutrition Messages in Two Decades of Online New York Times' Nutrition News Coverage. *Health Communication* 2023; **38**(2): 260-74.

432. Marmot M. Social determinants of health inequalities. *Lancet* 2005; **365**(9464): 1099-104.

433. Nettle D. Why are there social gradients in preventative health behavior? A perspective from behavioral ecology. *PLoS One* 2010; **5**(10): e13371.

434. Pepper GV, Nettle D. Perceived extrinsic mortality risk and reported effort in looking after health: testing a behavioral ecological prediction. *Hum Nat* 2014; **25**(3): 378-92.

435. Brown R, Sillence E, Pepper G. Perceptions of control over different causes of death and the accuracy of risk estimations. *Z Gesundh Wiss* 2023: 1-14.

436. Mialon M, Vandevijvere S, Carriedo-Lutzenkirchen A, et al. Mechanisms for addressing and managing the influence of corporations on public health policy, research and practice: a scoping review. *BMJ Open* 2020; **10**(7): e034082.

437. Carters-White L, Hilton S, Skivington K, Chambers S. Children's, parents' and professional stakeholders' views on power concerning the regulation of online advertising of unhealthy food to young people in the UK: A qualitative study. *PLOS ONE* 2022; **17**(6): e0268701.

438. Arnett DK, Blumenthal RS, Albert MA, et al. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol* 2019; **74**(10): 1376-414.

439. Brandt EJ, Tobb K, Cambron JC, et al. Assessing and Addressing Social Determinants of Cardiovascular Health: JACC State-of-the-Art Review. *J Am Coll Cardiol* 2023; **81**(14): 1368-85.

440. Anderson P, Stockwell T, Natera G, Kaner E. Minimum unit pricing for alcohol saves lives, so why is it not implemented more widely? *BMJ* 2024; **384**: e077550.

441. Rogers NT, Cummins S, Forde H, et al. Associations between trajectories of obesity prevalence in English primary school children and the UK soft drinks industry levy: An interrupted time series analysis of surveillance data. *PLoS Med* 2023; **20**(1): e1004160.

442. Short SE, Mollborn S. Social Determinants and Health Behaviors: Conceptual Frames and Empirical Advances. *Curr Opin Psychol* 2015; **5**: 78-84.

443. Timpson K, McCartney G, Walsh D, Chabanis B. What is missing from how we measure and understand the experience of poverty and deprivation in population health analyses? *Eur J Public Health* 2023; **33**(6): 974-80.

444. Zhong J, Cai Q, Zheng W, Chen S, Wu S, Dong S. Association of socioeconomic status and life's essential 8 with cardiovascular diseases and all-cause mortality in north China: Kailuan study. *Bmc Public Health* 2024; **24**(1): 2709.

445. Galante J, Adamska L, Young A, et al. The acceptability of repeat Internet-based hybrid diet assessment of previous 24-h dietary intake: administration of the Oxford WebQ in UK Biobank. *Br J Nutr* 2016; **115**(4): 681-6.

446. Batty GD, Gale CR, Kivimaki M, Deary IJ, Bell S. Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. *BMJ* 2020; **368**: m131.

447. UK Biobank Death Summary Report (March 2024). 2024. https://biobank.ndph.ox.ac.uk/~bbdatan/DeathSummaryReport.html (accessed 20/11/2024.

448. Office for National Statistics (ONS),Deaths registered in England and Wales: 2021 (refreshed populations), statistical bulletin. 2023. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsregistrationsummarytables/2021refreshedpopulations2024).

449. Office for National Statistics (ONS), Deaths registered in England and Wales (2021 edition). In: (ONS) OfNS, editor.; 2023.

450. Keyes KM, Westreich D. UK Biobank, big data, and the consequences of non-representativeness. *The Lancet* 2019; **393**(10178): 1297.

451. Littlejohns TJ, Sudlow C, Allen NE, Collins R. UK Biobank: opportunities for cardiovascular research. *European Heart Journal* 2017; **40**(14): 1158-66.

452. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J, 3rd. Factors of risk in the development of coronary heart disease--six year follow-up experience. The Framingham Study. *Ann Intern Med* 1961; **55**: 33-50.

453. Shenyang G. 734 The Cox Proportional Hazards Model. In: Guo S, editor. Survival Analysis: Oxford University Press; 2009. p. 0.

454. Stensrud MJ, Hernán MA. Why Test for Proportional Hazards? *JAMA* 2020; **323**(14): 1401-2.

455. Sterne JA, Davey Smith G. Sifting the evidence-what's wrong with significance tests? *Bmj* 2001; **322**(7280): 226-31.

456. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012; **380**(9836): 37-43.

457. Ho IS-S, Azcoaga-Lorenzo A, Akbari A, et al. Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies. *The Lancet Public Health* 2021; **6**(8): e587-e97.

458. Allik M, Brown D, Dundas R, Leyland AH. Developing a new small-area measure of deprivation using 2001 and 2011 census data from Scotland. *Health & Place* 2016; **39**: 122-30.

459. Kip E, Parr-Brownlie LC. Healthy lifestyles and wellbeing reduce neuroinflammation and prevent neurodegenerative and psychiatric disorders. *Front Neurosci* 2023; **17**: 1092537.

460. van Assen M, Helmink JHM, Gobbens RJJ. Associations between lifestyle factors and multidimensional frailty: a cross-sectional study among community-dwelling older people. *BMC Geriatr* 2022; **22**(1): 7.

461. Feldman S, Johnston C. Understanding the Determinants of Political Ideology: Implications of Structural Complexity. *Political Psychology* 2014; **35**(3): 337-58.

462. Carmines EG, apos, Amico NJ. The New Look in Political Ideology Research. *Annual Review of Political Science* 2015; **18**(Volume 18, 2015): 205-16.

463. Wang X, Bakulski KM, Paulson HL, Albin RL, Park SK. Associations of healthy lifestyle and socioeconomic status with cognitive function in U.S. older adults. *Sci Rep-Uk* 2023; **13**(1): 7513.

464. Rod NH, Lange T, Andersen I, Marott JL, Diderichsen F. Additive Interaction in Survival Analysis: Use of the Additive Hazards Model. *Epidemiology* 2012; **23**(5).

465. Sun JW, Zhao LG, Yang Y, Ma X, Wang YY, Xiang YB. Association Between Television Viewing Time and All-Cause Mortality: A Meta-Analysis of Cohort Studies. *Am J Epidemiol* 2015; **182**(11): 908-16.

466. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant* 2018; **52**(4): 1893-907.

467. de Wet C, Bowie P, O'Donnell CA. Facilitators and barriers to safer care in Scottish general practice: a qualitative study of the implementation of the

trigger review method using normalisation process theory. *BMJ Open* 2019; **9**(9): e029914.

468. May CR, Mair F, Finch T, et al. Development of a theory of implementation and integration: Normalization Process Theory. *Implement Sci* 2009; **4**: 29.

469. Ritchie J, Spencer, L. Qualitative data analysis for applied policy research. In: Bryma A, Burgess, R.G., ed. Analyzing Qualitative Data: Routledge; 1994.

470. Ritchie J, Lewis, J. Qualitative Research Practice: SAGE Publications; 2003.

471. Braun V, Clarke, V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006; **3:2**: 77-101.