Estimating the effectiveness and cost-effectiveness of body weight interventions for the prevention of non-communicable disease in local authority areas of England

Ben Amies-Cull

Green Templeton College

A thesis submitted for the degree of DPhil in Population Health

Nuffield Department of Population Health, University of Oxford

Hilary Term 2023

Acknowledgements

This thesis is not only a body of work, but a product of a much wider DPhil process, which was as much about the people I've been lucky enough to spend it with and the learning involved as it was the work. I would like to separately acknowledge the direct support and training I have received to deliver the thesis and its scientific contributions, additional guidance I've received towards my development as an academic, and the particularly special professional and personal relationships that have been essential through the DPhil.

Firstly, to my supervisors, Peter Scarborough and Jane Wolstenholme. Pete took me on as an experiment as a pre-doctoral research fellow and has dedicated a huge amount to my development – for taking a risk on me I am ever-grateful and I would not be in this position without him. I'm enormously proud to be starting work as a postdoc with Pete – it looks like the experiment paid off. Jane, too, has invested a great deal in me, particularly in her exceptional attention to detail and teaching me to read my own work with fresh eyes (though that one is still a work in progress). Both have been diligent in developing my technical skills in modelling and health economics, as well as wider research skills, and have always been available and engaged. They allowed me time to learn, develop, teach and explore. A mention is also absolutely essential for Linda Cobiac, who contributed as a supervisor with her unique brand of brilliance until the pandemic disrupted matters and she returned to Australia.

Also making technical contributions to the work were Dr Nia Roberts of Bodleian Libraries (on the literature review), Dr Chris Jackson of the University of Cambridge (on the use of the Disbayes model), Dr Michael Clark of the Smith School (on Generalised Linear Modelling) and Dr Cherry Law of the University of Reading (on healthcare cost estimates). In addition, there were all the trainers and presenters of courses, seminars and workshops at the Department, Division and across the University that made my learning so much broader and richer. Although there are too many to name individually, Jen de Beyer of the EQUATOR Network deserves a special mention as an exceptionally talented trainer who I have learned a great deal from.

I must acknowledge my funder, the Oxford-MRC Doctoral Training Partnership, for allowing the DPhil to happen, and specifically its lead, Prof Ester Hammond, for her support. I would like to thank NDPH for sponsoring me as an academic visitor, allowing me to complete important groundwork to support my application, then for taking me on for DPhil as an inexperienced researcher. At NDPH there are many other DPhils and senior colleagues who have provided great intellectual community, but I would specifically like to thank Prof Sasha Shepperd for her ongoing positivity and wisdom, and my talented friends and colleagues at the DPhil Student Committee and Early Career Research Network.

The Centre on Population Approaches for Non-Communicable Disease Prevention, as it was, is an extremely special place and group of people, both in term of its intellectual weight but also the special research culture that has grown up there, particularly due to its senior stewardship. I would particularly like to thank the DPhils ahead of me whose friendship has carried me through the process, Dr Lauren Bandy and Dr Jess Renzella.

At Green Templeton College I would like to acknowledge the support of Prof Alan Silman (a very special mention), Prof David Hunter, Dr Alison Stenton, Prof Denise Lievesley and Dr Davide Bilardi for their hours of friendship and mentorship. Most of all at Green Templeton, my College coach Kirsten Jellard – who donated a plainly unreasonable amount of time to me with her gently Socratic method of encouragement and self-development; with thanks to her I am leaving this process quite a different person to when I started.

To my colleagues at Mount Road Surgery for their understanding and flexibility throughout. To Manchester University Centre for Primary Care for keeping me on as an honorary member and providing me with ongoing office space. To Angela Pinnock for keeping track of us, often at unreasonably late notice.

I would like to thank my parents for their continuous support and encouragement for me to explore thinking and to learn to love learning, and to Jack, for always making space for me despite more important things going on.

Tom – more than anyone. The better part of me, allowing me to see what I can aim for, what I can do, refocusing me and for always pushing me on to do better, maintaining faith in me when I doubted myself, for being the other half of us. I could not have done it without you by my side.

To Dan

For putting me on this path. I hope you'd be proud of me.

Contents

List of tables

List of figures

List of abbreviations

5YFV	The NHS Five-Year Forward View
A&E	Accident and Emergency department
ACRA	Advisory Committee on Resource Allocation
AF	Atrial Fibrillation
ALS	Active Lives Survey
APC	Acute Patient Care
BIC	Bayesian Inference Criterion
BMI	Body Mass Index
CCG	Clinical Commissioning Group
CEA	Cost-Effectiveness Analysis
CKD	Chronic Kidney Disease
COVID-19	Coronavirus 2019
CPNP	Centre on Population Approaches for Non-communicable Disease Prevention
CPRD	Clinical Practice Research Datalink
CRA	Comparative Risk Assessment
CV	Coefficient of Variation
CVD	Cardiovascular Disease
d.p.	Decimal places
DALY	Disability-Adjusted Life-Year
DPP	Diabetes Prevention Programme

GBD	Global Burden of Disease
GHDx	Global Health Data Exchange
GLM	Generalised Linear Modelling
GP	General Practice
HEE	Health Economic Evaluation
HES	Hospital Episode Statistics
HFMA	Healthcare Financial Management Association
HFSS	High Fat, Sugar and Salt
HHD	Hypertensive Heart Disease
HRG	Healthcare Resource Group
HSCA	Health and Social Care Act
HSE	Health Survey for England
ICD-10	International Classification of Diseases, Tenth Revision
ICER	Incremental Cost-Effectiveness Ratio
ICS	Integrated Care System
IHD	Ischaemic Heart Disease
IMD	Index of Multiple Deprivation
IPF	Iterative Proportional Fitting
ISPOR-SMDM	Professional Society for Health Economics and Outcomes Research and Society for Medical Decision Making
LSOA	Lower Super Output Area
MeSH	Medical Subject Headings

MFF	Market Forces Factor
NCC	National Cost Collection
NCD	Non-Communicable Disease
NCMP	National Child Measurement Programme
NDNS	National Diet and Nutrition Survey
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMB	Net Monetary Benefit
NUTS	National Units of Territorial Statistics
OLS	Ordinary Least Squares regression
ONS	Office for National Statistics
OPCS-4	Classification of Interventions and Procedures, version 4
OPD	Outpatient Department care
PA	Physical Activity
PB	Programme Budgeting
PCS	Prospective Cohort Study
PHE	Public Health England
PIF	Population Impact Fraction
PLICS	Patient-Level Indicative Costing System
PMSL	Proportional Multistate Lifetable
PRIME	Preventable Risk Integrated ModEl
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta Analyses
PSA	Probabilistic Sensitivity Analysis

PSU	Primary Sampling Unit
QALY	Quality-Adjusted Life-Year
RCI	Reference Costs Index
RCT	Randomised-Controlled Trial
RMSE	Root Mean Squared-Error
ROI	Return on Investment
RUC	Rural-Urban Classification
s.f.	significant figures
ScHARR	School for Health And Related Research
SD	Standard Deviation
SDIL	Sugary Drinks Industry Levy
SE	Standard Error
SES	Socioeconomic Status
SPHR	School for Public Health Research
T2DM	Type-2 Diabetes Mellitus
UK	United Kingdom
WHO	World Health Organisation

Abstract

Background

Non-Communicable Diseases related to Body Mass Index (BMI) account for approximately 10% of all disease burden in England. Addressing this burden and its associated health inequalities are major challenges for local and national policymakers, with increasing need to also understand how BMI interventions might affect local areas differently. The aim was set to develop a new local authority-level health model to estimate the disease burden and healthcare cost implications of BMI interventions.

Methods

Diseases included in the new model were asthma, low back pain, osteoarthritis of the hip, osteoarthritis of the knee, Ischaemic Heart Disease, stroke, hypertensive heart disease, type-2 diabetes mellitus, atrial fibrillation/ flutter, colorectal cancer, breast cancer and oesophageal cancer. Local-level data were estimated via different approaches: adult BMI distributions were produced via individual-level synthetic estimation, child BMI distributions were interpolated from measured BMI and healthcare costs estimated using new high-quality cost collection data. Disease epidemiology was estimated at the Index of Multiple Deprivation quintile level using a Bayesian modelling tool. A scenario on restricting television advertising of unhealthy foods in each of 315 local authorities in England was then used to explore the new model's capabilities.

Findings

Patterns to adult and child BMI supported conventional ideas that raised BMI is associated with age and deprivation. Healthcare costs formed very positively skewed distributions without an obvious geographic pattern. Disease epidemiology showed burden was broadly higher in more deprived quintiles.

Modelling showed that restricting the advertising of unhealthy foods is likely to benefit more deprived areas more overall, but with overlaps between quintiles of deprivation. It is estimated that 0.0222-0.0521 Quality-Adjusted Life-Years per person could be saved over the lifetime of the 2018 population of children and £11.7-£42.8 per person in healthcare costs.

Conclusions

This model offers the most local-specific health modelling for BMI interventions in England. Some areas may be harder to reach with a given intervention, creating mismatch between need and ability to intervene.

Chapter 1: Introduction

Context

Body weight and health

Excess body weight is well established to be linked to health outcomes, specifically due to body fat composition. There are various ways of defining and measuring excess weight, but one of the most common is the Body Mass Index (BMI). This quantifies the healthiness of an individual's weight, relative to their height, as kg/m^2, with a BMI of 19-25 categorised as a healthy weight, BMI 25-30 as overweight and over 30 as obese. Increasing body weight constitutes a major public health risk, with excess body weight estimated to account for 9% of all morbidity and mortality in England.(1,2)

In 2015 Public Health England (PHE) estimated that 63% of adults in England were overweight or obese – with this proportion having nearly doubled since 1993. This burden is estimated to be leading to a £6.1bn annual cost to the National Health Service (NHS) and £27bn annual costs to society and the economy (for example due to lost economic output due to deaths in the workforce).(3) Rates of overweight and obesity are also strongly associated with increasing deprivation, and some ethnic minority groups are both more affected by excess body weight as well as less favourable risk relationships between BMI and some health outcomes.(3–5)

Of particular additional concern is that rates of childhood obesity continue to rise, though with hints before the Coronavirus 2019 (COVID-19) pandemic that it may have been starting

to level out or decline. There have been dramatic increases in childhood obesity since the start of the pandemic, including worsening of pre-existing inequalities.(6,7) As excess body weight earlier in life predicts excess weight later, there is a major concern that the lifelong risk exposure will contribute towards increasing burden of disease in those generations as older people.(8) Moving beyond the United Kingdom's (UK) headline childhood obesity figures also provides cause for concern – while obesity rates for children from more affluent communities may have started to decline before the epidemic, those for children in more deprived communities continued to increase.(7)

Raised BMI is associated with increased rates of Non-Communicable Diseases (NCDs), probably largely due to proinflammatory activity of adipose (fat) tissue.(9) The NCDs are a highly heterogeneous group of disorders whose development is associated with environmental factors (such as air pollution or car-dependent infrastructure), lifestyle (eg. alcohol or unhealthy food consumption), age or sex.

Public Health responses to excess body weight

The escalating health and economic implications of excess body weight make reducing its prevalence a common aim for policymakers. There are many ways they can approach the goal of reducing excess weight(10–13) and how these approaches should be combined is a major policy conundrum.(14–16). For example, we can focus on preventing people from gaining excess weight, or help those with excess weight to lose it. We can implement approaches that work one-on-one with individuals, through to strategies that work more remotely on greater populations. Geoffrey Rose (1985)(17) described the prototypical dichotomy between 'high-risk' approaches and 'population' approaches, where the former

focuses on greatly reducing the risk of those with the highest total risk, while the latter slightly reduces risk across a whole population, regardless of risk level. The important observation was that the total accumulation of smaller risk improvements across larger numbers may be greater overall. In reality, these two approaches form a spectrum.(17)

Each specific intervention will affect different people and communities differently – and none of them are easy.(18,19) Our metabolism makes losing weight very challenging and long-term behaviour change is extremely difficult to achieve. From these starting points we also find that the effectiveness of a given approach is likely to vary widely between communities due to social, cultural, economic and many other forms of variation.(20) Layered onto these issues is a question over how much we should prioritise reducing inequalities over purely focusing on reducing total disease burden or costs.(21)

Ways for us to understand the impacts of the food system on health, anticipate how interventions may affect different people or organisations differently and how to prioritise approaches are badly needed. In a now classical piece of work, this complexity was summarised as the *Foresight obesity map*, shown in figure 1.1.(19) This work demonstrated that easy wins may be unlikely and that much better basic understanding of the system would be a long-term intellectual project, alongside our attempts to improve health in the shorter term.(19)

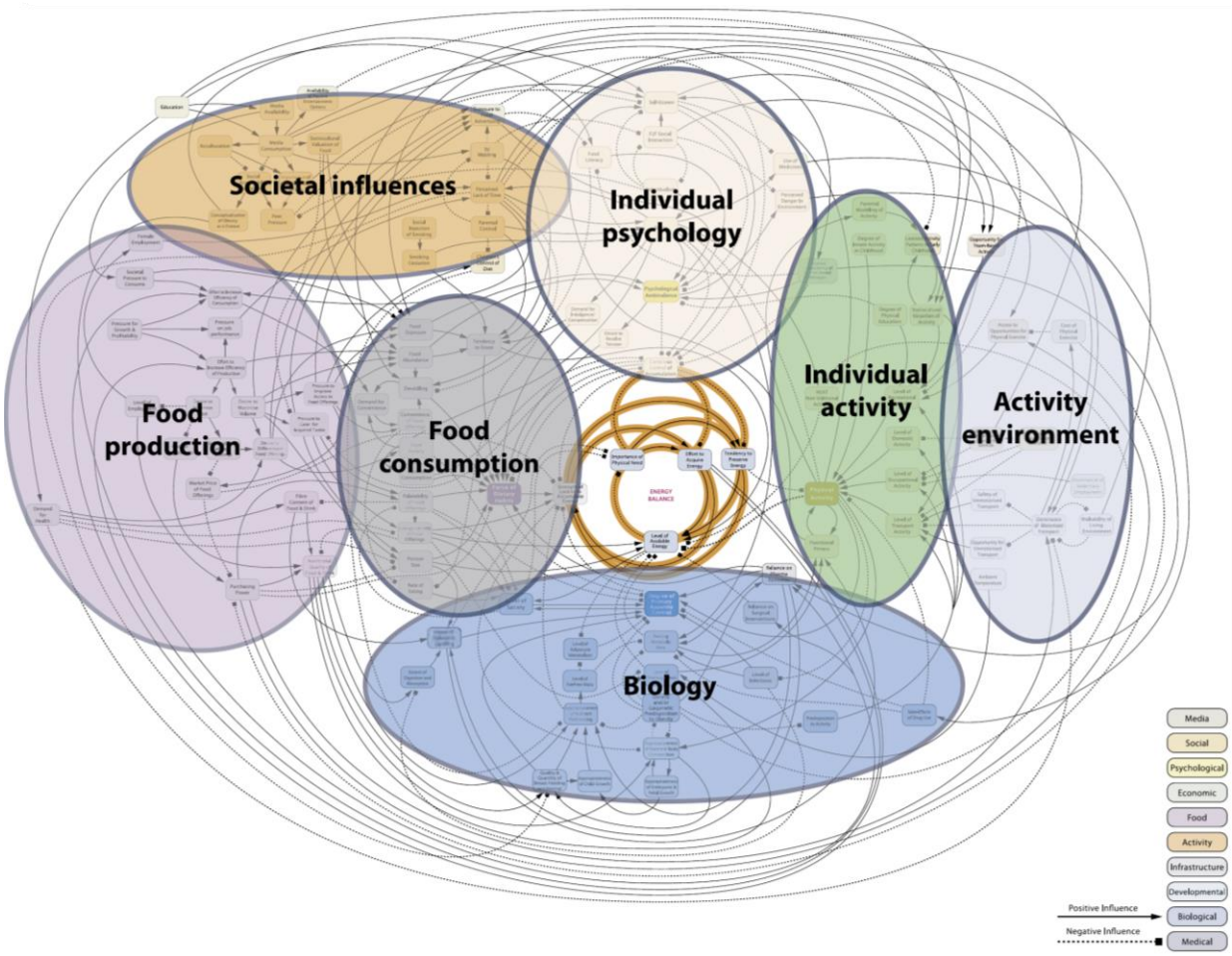


Figure 1.1: The Foresight obesity map, from Foresight Tackling Obesities: Future Choices – Project Report, 2nd Edition.(19) (This is not intended to be legible here, but to illustrate the number and diversity of factors at play in determining body weight, and the complexity of their interactions.)

Policy context

Prevention has been a persistent priority for UK governments. A great deal of attention has ostensibly been paid to preventing and treating excess body weight, including at the heart of the most influential government strategies. From 1992 to 2020, UK governments published 14 strategies for obesity,(13) most recent, the 2020 *Tackling Obesity: Empowering Adults and Children to Live Healthier Lives*.(11) Both the 2014 NHS 5-Year Forward View (5YFV)(22) and 2019 NHS Long-Term Plan(23) put prevention as top priorities. However,

fourteen years after the Wanless Report (2002) warned that avoidable illness could overwhelm the health service,(24) the 5YFV admitted the NHS had been left “on the hook” for failure to act. After years of less interventionist approaches,(13) the introduction of the 2016 UK Sugary Drinks Industry levy (SDIL)(25–27) marked new policy interest in population approaches. Former Prime Minister Boris Johnson was famously convinced to take more direct action on obesity after being critically ill with COVID-19, the severity of which he acknowledged was likely to be related to being “way overweight”.(28)

Other changes to the health landscape in the UK changed how excess body weight is being approached. The Health and Social Care Act 2012 (HSCA 2012) moved responsibility for Public Health from the NHS to local authorities. Very large cuts to Public Health budgets followed this as the decision was made not to apply the NHS ringfence to those budgets, which protected the NHS from nominal budget cuts.(29) There has also been a recent emphasis on government delegating more power to local areas, such as through the Localism Act 2011 and “devolution deals” to groups of local authorities(30) and what is referred to as the “levelling up” agenda,(31) broadly interpreted as the aim to reduce geographic inequalities at various levels.

Local authorities form the main administrative levels of government in England below Westminster. Their configuration occasionally changes, but as of 2019, there were 317 lower tier and unitary local authorities in England, with populations mostly of 50,000-500,000. They are independently elected of central government, with specific responsibilities and revenue streams (mainly made up of central government grants, council tax (residential properties) and business rates (commercial properties)).(32) Most areas have two layers of local authority. Lower tier authorities are responsible for housing,

planning, rubbish collection and recycling. Groups of these, usually at the level of the county, form upper tier authorities, which are responsible for services such as public health, education, planning, social care and trading standards. In some areas these are compressed into one layer that is responsible for all of these duties, particularly in cities.(33)

Decision-making in disease prevention

Making good decisions on what approaches to implement is critical to addressing the challenge of obesity, but to select the most appropriate interventions, policymakers need good information. Using a health economic framework, the broad dimensions they would need to consider are effectiveness, cost-effectiveness and equity, as well as unwanted effects.(21,34)

The movement towards local decision-making and "levelling up" have created two needs to understand local variation in an intervention's costs and effects. Firstly, local decision-makers need local-specific information on each intervention to allow them to select appropriate interventions for their area, considering not simply what appears cost-effective on average in England but what is most cost-effective for their area. Secondly, national decision-makers need to understand how the costs and effects of a policy implemented nationally would vary by local area. This would allow them to answer "levelling up" questions on how obesity-related geographic health inequalities could best be reduced.

Policymakers also must consider many other factors such as public acceptability, legislative context and implementation levers, all of which fall outside the remit of this research. Of these, a particularly consequential factor for this thesis was the Nuffield Bioethics Council "ladder", which arose in an influential report from 2007 on the ethics of public health

interventions. The “ladder” provides a structure of how much state involvement is required to tackle a public health problem, or (alternatively) how much individual autonomy needs to be interfered with. It encouraged policies to be selected that involve the least interference with individual autonomy.(16) This report has been influential at home in the UK but also abroad, being used as evidence in specific instances of decision-making, but also setting a more general climate around public health policy.(35) Other factors were weakly considered in the ladder (though somewhat better dealt with through the report), so when this is implemented into the real world, it is likely to bias towards interventions that may increase health inequalities as the communities with the required agency to benefit from the higher-autonomy interventions also tend to be in the most affluent communities. This may lead those with the most need to gain the least benefit, widening inequalities.(36)

Another important factor in policy choice is that there is now an expectation for policies to be underpinned by adequate evidence before being implemented, to deliver public value for money.(37,38) However, the benefits and costs of some approaches can be precisely estimated more easily than others. Prototypically, the most likely approach to produce unbiased estimates of an intervention’s costs and effects is a meta-analysis of systematically reviewed Randomised-Controlled Trials (RCTs). A problem arises that some interventions are much less amenable to examination by trials, making them potentially less appealing to policymakers. For example, RCTs of weight loss services are now routine, providing relatively comparable impacts on body weight over time, so better ones can be identified and implemented.(39,40) Conversely, it would be extremely difficult, if not impossible, to implement a trial for a tax on sugary drinks (*sugar-sweetened beverages*) that yielded an unbiased result. Even if the logistics could be arranged, the evidence on the UK SDIL indicates that most of its effects on calorie consumption came from industry reformulating

products after the policy's national announcement.(27,41) It seems unlikely this industry behaviour would have been prompted by a trial implemented for a limited time. Evaluations of population health interventions may struggle to directly measure the intervention's impact on BMI – such as the impact of a sugary drinks tax on BMI. Therefore, other ways of estimating its impacts may be necessary, such as estimating the impact of the tax on purchases and sugar consumption, with additional assumptions necessary to estimate the impact of these changes on BMI and health. Interventions that cannot be feasibly trialled need other methods to estimating body weight impact, which will vary by intervention. The preference for RCTs may also structurally lead to the selection of approaches that increase health inequalities as individual-level interventions that can be evaluated by a RCT tend also to require greater agency.(42,43) Accounting for these factors is now expedient, to reduce health inequalities and aid "levelling up".

Once an effect on body weight has been estimated, we still then do not know what the effect on disease rates will be across the lifetime. Empirically measuring the long-term health impacts of body weight change is not feasible given that people would have to consent to decades of follow up, while over time other influences over people's health would overwhelm the impact of short-term intervention and the apparent effects dissipate– due to statistical realities rather than physiological ones. The 'signal' of the intervention would be lost in the 'noise' of the real world.

Modelled evidence

Quantifying these long-term effects of interventions is valuable in that it allows the health and healthcare cost implications to be compared with other health interventions. It allows

formal health economic evaluation (HEE) to be undertaken to allow the comparison against benchmarks such as the National Institute for Health and Care Excellence (NICE) wiliness to pay threshold of £20,000-30,00 per Quality-Adjusted Life-Year (QALY). The inability to take direct measurements of public health interventions over long periods of time means these outcomes require modelling of some type. There are now a wide variety of scenario model structures for Public Health interventions, each with different strengths and limitations for a given application.(44,45)

Examples of specific NCD scenario models are provided in the literature review in Chapter 2. The PRIMEtime model is an example of a proportional multistate lifetable model (PMSL) that will be used as the model structure for this thesis. The PMSL and PRIMEtime were developed by Barendregt, Cobiac and colleagues,(46–48) and have been used to estimate the long-term impacts of interventions for NCD prevention related to obesity and other NCD risk factors.(27,39,48–53) More detail on PRIMEtime is given in Chapter 3.

Health Economic Evaluation

Modelling provides the means to perform HEE. HEE is based on a broad desire to choose a 'best' option for achieving a certain aim, given limited resources. This involves estimating efficiency, to allow the selection of the most efficient option(s) for maximising the health return from investing those limited resources. There is a variety of different methods for quantifying and comparing costs and benefits. Cost-Consequence Analysis involves listing disaggregated non-equivalised costs and benefits, leaving the user to synthesise/ interpret them in their own context. For example, a procedure might result in more nausea but less pain than an alternative. Cost-Minimisation Analysis is uncommon, but may be used where

benefits can be shown to be statistically equivalent, to identify the least costly approach. This may be useful such as when a trial finds no health difference between approaches. Cost-Effectiveness Analysis (CEA) has become widely used. This involves the calculation of aggregated costs and aggregated benefits for two interventions, then the calculation of incremental difference of one intervention relative to the other, for example the Incremental Cost-Effectiveness Ratio (ICER) or Net Monetary Benefit (NMB). Benefits can be quantified in any standard unit of health outcome, such as number of cases or units of blood pressure (millimetres of mercury, mmHg) but the QALY has become most often used, as a useful tool enabling the comparison of health benefits of widely differing interventions and across different diseases. The use of the QALY in CEA is referred to as Cost-Utility Analysis as the 'quality adjusted' component is measured in terms of utility. Cost-Benefit Analysis assumes a consistent level of resource then compares health benefit delivered by given interventions, monetising the costs and the health benefits.(34)

The QALY combines the two traditional categories of disease burden – morbidity and mortality – through a combination of quality of life and quantity of life. A QALY represents one year lived with a health *utility* (or health state preference value, where preference can be equated with value or desirability) of 1, representing perfect health. A utility of 0 represents death, with increasing levels of disability and symptom burden represented by values less than 1 (states worse than death can be represented by values less than zero). For any disease state, a utility decrement can be estimated representing the impact of that disorder on perfect health to be subtracted from 1, such that minor ailments have small utility decrements and more severe conditions may have much greater weightings.(34) One year lived in perfect health plus one year lived with a single disease that has a utility value of 0.25 (a decrement of 0.75) together are represented by 1.25 QALYs. Disability-Adjusted Life-

Years (DALYs) are calculated to represent the inverse of QALYs – one year at a utility decrement of 0.75 is represented by 0.75 DALYs. DALYs are also calculated with reference to specific disease states, rather than health/ disability states, attempting to improve cross-cultural comparability of results.(54)

The QALY is an extremely useful and widely accepted measure of health value, providing particular value in allowing different types of health outcomes to be quantified in the same metric. However, it is important to note that the QALY is also open to criticism on a few fronts. These criticisms can be broadly categorised as ethical, practical and contextual factors.(55) Ethical issues, for example, might be that it may introduce a structural bias against treatments for the elderly or terminally ill, who have a ceiling on how many QALYs it may be possible to gain from an intervention, leading to resource allocation away from those areas.(54,55) Utility values can be calculated via different methods, such as the Standard Gamble(56,57) and Time Trade-Off,(58) so there are practical concerns around methodological consistency in generating utility values, the implementation of a particular method or population sample variation between studies. There is also a failure to grasp non-health benefits such as the impact of an illness on family. Contextual factors are considerations such as whether it is reasonable to compare very different experiences of patients of very different disease types, for example, depression with a broken arm.(55) DALYs are calculated measuring utility by function rather than symptoms, and weight by age such that in older ages DALYs saved outweigh QALYs averted by the same intervention.(59)

Aims

The aim of the work described in this thesis is to build an NCD scenario model capable of estimating the health and cost effects of interventions aiming to reduce excess body weight at the lower tier local authority level in England, and to run and test this model. The purpose is to allow the examination either of how the costs and effects of a nationally-implemented intervention vary by local authority area or to estimate local-specific costs and effects of locally-implementable policy options to a given local authority area. This will allow HEE to be undertaken and differences between areas examined. The aim is primarily as an academic tool, though there are potential practical uses for national and local government.

This is done by developing the PRIMEtime PMSL structure mentioned above to capture important parts of the process of accumulating disease burden and costs that vary locally. Specifically, local authority areas' specific populations, BMI distributions, disease epidemiology and NHS costs are estimated. The scenario policy is the restriction of television advertising of food products high in fat, sugar or salt (HFSS products) to children. This model will be referred to as *PRIMEtime_local* throughout.

Chapter summary

Chapter 2: Literature review

Chapter 2 provides a narrative literature review of modelling capabilities for diet, physical activity (PA) and body weight models for subnational areas of England. This aims to establish

a baseline of what modelling capabilities exist, an assessment of how each model fits best practice criteria, considering how they may be improved with further work.

Chapter 3: Model summary

The structure of the new local authority PRIMEtime model is described. This covers the PRIMEtime model structure and how it is adapted in this thesis to allow local-specific flexibility using the estimated parameters from chapters 4-7.

Chapter 4: Estimating local disease epidemiology

As underlying disease rates influence the total quantity of disease that is mediated by BMI, knowing how these vary locally is also influential over modelled results. Most of this data is now available but specifically, case fatality rates need estimating. These represent the additional mortality risk conferred through having a disease, which varies by age, sex and geography. This work takes the existing gold standard epidemiology and applies a complex Bayesian modelling method to estimate case fatality rates, through exploiting the logical dependencies between incidence, prevalence and mortality in the published data.

Chapters 5 and 6: Estimating BMI distributions for adults and children

In two stages, these chapters describe the estimation of the mean and spread of BMI by age and sex across the population. These BMI distributions define the basic weight-related risk in the population: the higher people's BMI, the greater weight-related risk the population is exposed to. Direct estimates are not available at the local level as the sample sizes required of health surveys would be very large to be adequately accurate at the local level.

Therefore, other methods are required, involving using existing data to estimate local differences. Chapter 5 estimates local-level BMI distributions for adults and Chapter 6 produces equivalent estimates for children.

Chapter 7: Estimating local healthcare costs

NHS costs are known to vary widely between providers and across geographies. Therefore, the use of locally-specific costs could dramatically change the cost-effectiveness implications of an intervention in an area relative to applying a national average. Estimating these costs was done by extending methods for estimating national-level disease costs by taking new local cost data that are thought to be more consistent, then using these to vary national-level cost estimates.

Chapter 8: Estimating the local impacts of restricting advertising of foods high in fat, sugar and salt to children

This chapter is a scenario analysis of the potential impacts of restricting the advertising of HFSS products to children for each local authority area of England. First, a local-specific effect of the proposed intervention is estimated for each area, then PRIMEtime_local used to estimate potential impacts on costs, diseases cases and QALYs for each area. It then goes on to examine the health economic implications of the variation in modelled costs and effects.

Chapter 9: Discussion and conclusion

This summarises and discusses the work, including the implications of each step, its place in the wider literature, and its strengths and limitations. It will go on to consider what future opportunities there may be for using the model and what future research would be valuable based on this work.

Appendices

The Appendices then provide additional detail to the methods, including three longer-form appendices: a quantitative description of how disease states were selected for the model, a description of estimating local populations of people aged 90 years and over, and an assessment of the PRIMEtime_local model's performance against modelling best practice recommendations.

Chapter 2: Literature review

Introduction

Before embarking on developing a model that is able to model different local areas, accounting for local population, risk and disease characteristics, a review of current local-level England health models was undertaken to describe what local non-communicable disease modelling capabilities currently exist in England and describe the quality and scope that any new model would need to build on.

The purpose of the review was to summarise the current state of modelling capabilities for the estimation of the effects of interventions aimed at reducing the risk factors related to diet, PA and adiposity. This is to ensure that modelling functions being developed in this research programme are new contributions to capabilities in England. Previous reviews of NCD modelling capabilities have focused mainly on structure and quality criteria rather than capabilities or geographies included.(60–63). Some considered only one disease(60) or one modelling structure.(63)

The aim of this literature review is to explore what NCD modelling capabilities have been developed for local areas in England, regardless of who developed them and why.

The role of a systematic review was considered, as there would be benefits in using a gold-standard method, producing reproducible results (for some arms of the review) and ensuring relevant publications were not missed. However, the main purposes of performing a systematic review were not relevant to the review question in this case. Namely, the

Cochrane Handbook (v6.3)(64) states the core aim of systematic review as to avoid the need for individual practitioners to need the time, skills and resources to review and appraise the relevant literature on each area of practice. This review clearly falls outside this frame. Mulrow(65) also listed purposes of systematic review of avoiding additional primary research by producing generalisability and consistency in findings, increased statistical power, precision of an effect on outcomes and accuracy of clinical realities (ie. unbiased results). As this is a review of approaches, not findings, these benefits of systematic review are also not relevant. It was also apparent that relevant models may not be identified in formal database searches (especially those produced by government bodies). Therefore, the conceptual validity of a "systematic review" of literature could not be met in this search topic, even if the method was applied to the peer-reviewed literature on the topic.

Therefore, the significant additional procedural burden of a systematic review over a well-conducted formal database search was felt to out-weigh the small potential additional benefits, with the time instead being better used extracting grey literature.

Methods

Theoretical model of search strategy

The search strategy was designed to reflect the nature of the literature base and where publications were likely to be found. The peer-reviewed literature contains many examples of health modelling, so a formal literature search would be necessary. However, it was clear that models would also be found in a wide variety of different formats and different

sources, many of which would lie outside the peer-reviewed literature, especially from public bodies. A wide-ranging grey literature search was therefore designed to be systematic, transparent and as exhaustive as possible, to complement the formal search of the peer-reviewed literature. This grey literature search would widen the ability to identify relevant models given the fragmentation, heterogeneity and lack of central organisation of the relevant literature.

Despite not being a formal systematic review, for the sake of clarity, transparency and comprehensiveness, the remainder of this review are reported using the structure of the PRISMA statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)(66,67), as numbered below.

Inclusion and exclusion criteria

For these purposes, models are defined as simplified representations of reality that use data to apply mathematical relationships between input and output variables to estimate the solution to a real-world problem.(45) In this case, these models are also intended to make scenario estimates of the health (and/ or cost) impacts of modifying risk factor prevalence, with time represented. Time was included to account for NCDs often developing after many years of risk factor exposure and to allow total cumulative costs and benefits to be compared.

Models that met the following criteria were shortlisted for full paper review:

1. Population of any age group, living in England
2. Those relating to adiposity, or related dietary or PA risk factors

3. Any health outcome, healthcare or social care cost related to the risk factors listed
4. Modelling impacts on populations rather than individuals
5. Modelling geographic subgroups of England

A criterion for models to represent an element of time was moved from the initial screening stage to the full paper review stage, as it was usually not possible to tell from abstract review if time was included or not. An initial inclusion criterion of a minimum age bracket of 18-65 years was dropped as this was often not adhered to in models for valid reasons.

Additional restrictions were applied of having a full text available, being in English and being published in the window from 1 January 2000 to 24 March 2019, when the search started. The start date for this window was chosen on the basis that a previous systematic review from 2006 found that no NCD prevention models using complex structures had been published in the UK by 2000.(60)

Development of search terms

Known examples of relevant models (from peer-reviewed papers,(68–70) theses(71,72) and public organisations(73,74)) were used to identify and generate key terms. These were then expanded using MeSH (Medical Subject Headings) terms as prompts. A research librarian at the Oxford Bodleian Libraries (Nia Roberts) supported this by reviewing the full lists of keywords and how they were to be structured as a search strategy. Additional methodological guidance on Grey Literature searching was gained from the British Library,(75) the Grey Matters resource(76) (from the Canadian Agency for Drugs and Technology in Health) and published papers on Grey Literature search strategies.(77–79)

Terms are listed in groupings in Table 2.1. Terms were grouped into ‘non-communicable disease’ terms, ‘risk factor’ terms, ‘modelling’ terms and ‘spatial’ terms.

Information sources

To deal with the fragmentation and heterogeneity of the literature, a search strategy incorporating a wide variety of sources was executed, involving multiple overlapping approaches.

Five separate strands of review were conducted.

These were:

1. One formal database search (24 March 2019).
2. One thesis catalogue search (1-2 April 2019).
3. Google web search (4-5 and 15 April 2019).
4. Google scholar search (16-17 April 2019).
5. Hand search of selected public body websites (29 April 2019).

Search approaches

Database search

Medline was chosen for the peer-reviewed literature. A hierarchical approach was used for the benefits of being transparent, editable and at the lowest risk of introducing error. The hierarchy was structured as a series of “OR” searches combined into a single “AND” search.

Taking the search terms set out in Table 2.1, terms within a group were combined in an "OR" search for each group. The NCD and risk factor terms were supplemented with a MeSH term search for the same concept. These keyword searches and the MeSH terms search were combined into a single "OR" search. Finally, the combined risk factor and NCD term searches were combined into a single "AND" search with modelling and spatial terms, producing a search structure as:

[(NCD keyword) OR (risk factor keyword) OR (NCD and risk factor MeSH terms)] AND [modelling keyword] AND [spatial keyword]

Results were then filtered by English language results with available full texts published from the year 2000.

Table 2.1: Search terms forming the basis of search approaches. The * indicates a truncated term.

Non-communicable disease	**Risk factor**	**Modelling**	**Spatial**
non-communicable disease*	diet*	modelled	small area*
NCD*	physical activity	modelling	NUTS3
chronic disease*	exercise*	simulation	small geograph*
chronic illness*	sedentar*	microsimulation	city
chronic condition*	obesity	population	cities
long-term condition*	overweight	model*	county
long-term illness*	adiposity	health model*	counties
multimorbid*	fatness	statistical model*	clinical commissioning group*
multiple morbidity	body-mass index	economic model*	local NHS
diabetes	body mass index	health economic	local authorit*
T2DM	BMI	policy model	council*
cardiovascular disease	calori*	cost utility	CCG*
coronary disease	saturated fat	life year*	synthetic estimate*
coronary atherosclerosis	unsaturated fat	cost benefit	model-based estimate*
angina	polyunsaturated fat	cost effective*	
myocardial ischaemia	MUFA	Markov	
ischaemic heart disease	PUFA	lifetable*	
myocardial ischaemia	monounsaturated fatty acid*	life table*	
stroke	sugar	decision tree	
cerebrovascular disease	sucrose	state transition	
thromboembol*	carbohydrate	policy model*	
embol*	salt	deterministic	
cancer	sodium	probabilistic	
malignan*	red meat	stochastic	
neoplas*	processed meat	system dynamic	
	red or processed meat	discrete event	
	processed or red meat	agent based	
	fruit*		
	vegetable*		
	dairy		
	vegetarian		
	vegan		
	mediteranean		
	flexitarian		
	pescatarian		

Thesis catalogue search

The Ethos catalogue of the British Library(75) was searched using 30 unique search terms (in Table 2.2). The search tool cannot structure complex hierarchical “AND”/ “OR” searches, so individual complete search phrases were created. The aim was set for 25 unique phrases to return results, but of the 25 initial phrases, but five returned no titles, so five more were added, all of which returned titles. The search phrases are set out in Table 2.2, with the numbers of titles returned by the search tool, the number of abstracts screened and the number of full papers reviewed. Search terms were created as coherent English phrases by selecting from across the terms from Table 2.1 to produce a broad selection of unique terms. Terms were used from the non-communicable disease, risk factor group and modelling group, but not the spatial group to keep results broad given the lower absolute numbers of entries in the catalogue relative to the other search approaches.

Google and Google Scholar searches

As with the Ethos thesis tool, neither Google nor Google Scholar can structure complex searches, so a similar approach of using coherent search phrases was used. Ten phrases were created and the same phrases were used for both searches (in Table 2.3). These phrases included spatial terms as the volume of items available through both of these search tools is extremely large, so these were kept in order to narrow the scope of the search and optimise the sophisticated search algorithms that Google uses to prioritise online material. These searches proceeded by screening every result from the ‘organic search’ results produced for each phrase (that is, excluding paid adverts) and continuing through every page of search findings up to a page with no new relevant finding. All Google

and Google Scholar searches were executed in the *Incognito* setting, as Google's algorithm learns responsively from the content that users interact with, which could increase duplicated pages being returned in searches, or otherwise influence results. The phrases used in Google and Google Scholar searches are listed with the number of findings on each page in Table 2.3.

Organisational website hand search

The websites of ten domestic and international organisations were included to be hand-searched for relevant content that had otherwise not been identified in the approaches listed above. UK organisations were: the Department of Health and Social Care, NHS England, NICE, PHE, NHS Innovation, NHS Digital, NHS Evidence and the Local Government Association, and international organisations were the World Health Organisation and the Organisation for Economic Co-operation and Development.

Table 2.2: Search phrases used in Ethos thesis catalogue search with numbers of abstracts and full papers screened.

	Search phrase	**Search results**	**Abstracts screened**	**Full papers**
1	non-communicable disease modelling	24	4	2
2	non-communicable disease simulation	1	1	1
3	non-communicable disease forecasting	1	0	0
4	non-communicable disease system dynamic	4	1	0
5	chronic disease prevention modelling	131	5	1
6	chronic disease simulation	48	0	0
7	chronic disease forecasting	4	0	0
8	chronic disease system dynamic	32	0	0
9	chronic disease life table	2	1	0

10	chronic disease microsimulation	1	1	1
11	chronic disease cost effectiveness	83	3	0
12	chronic illness prevention modelling	24	1	0
13	chronic illness simulation	6	0	0
14	chronic illness modelling	228	2	0
15	cardiovascular disease prevention modelling	79	5	2
16	coronary disease modelling	44	6	0
17	cancer prevention modelling	214	5	1
18	cancer prevention simulation	20	0	0
19	diabetes prevention modelling	103	5	0
20	diabetes simulation	77	6	2
21	diabetes forecasting	4	2	1
22	obesity prevention modelling	65	5	0
23	obesity cost effectiveness	51	3	1
24	diet prevention cost effectiveness	11	1	0
25	physical activity prevention cost effectiveness	29	2	0
	non-communicable disease cost-effectiveness	0		
	chronic illness life table	0		
	chronic illness forecasting	0		
	chronic illness microsimulation	0		
	cancer forecasting	0		
Total:		1286	59	12

Table 2.3: Search phrases used in Google and Google Scholar searches, with numbers of papers on each page of the searches

		Google page							Google Scholar page						
Search phrase:		p1	p2	p3	p4	p5	p6	p7	p1	p2	p3	p4	p5	p6	p7
1	local chronic disease modelling	1	1	0					5	1	2	2	2	0	
2	city chronic disease microsimulation	2	4	1	4	0			5	4	2	4	1	0	
3	local non-communicable disease cost effectiveness	2	1	0					1	0					

4	city non-communicable disease forecasting		0								1	0					
5	local cardiovascular disease prevention modelling		3	3	0						3	5	3	1	0		
6	local cancer prevention modelling		0								0						
7	local diabetes prevention modelling		5	2	1	3	1	1	0		4	2	2	3	1	2	0
8	city obesity prevention modelling		0								3	0					
9	local diet prevention cost effectiveness		1	3	0						2	4	2	1	0		
10	city physical activity prevention cost effectiveness		1	0							1	0					

Study selection

Database search

Paper and abstracts were used to perform an initial screen for relevant papers. The search tool had already screened these for language, publication date and having a full text available, so remaining inclusion criteria set out above were used to screen papers for full paper review.

Thesis catalogue search

Each search term returned a list of theses, laid out by title and author. Abstracts were used to screen papers for the inclusion criteria set out above as well as for publication date and full text availability (all theses were in English as the British Library's Ethos resource includes only theses from UK universities).

Google search

Search results were displayed in pages of ten links. Every link on each page was opened, then all ten were screened by the criteria set out above. Non-English results were not opened. Some pages did not have an abstract or other concise summary, so in these cases full papers were screened for relevance or any clear indication of not meeting inclusion criteria.

Google Scholar search

Titles and abstracts were used to perform an initial screen for relevant papers. Results not in English and those without access to full papers were excluded manually at this stage. To deal with volume and the large degree of overlap between Google and Google Scholar results, an initial screen was followed by de-duplication and then papers were confirmed for screening criteria.

Organisational website hand search

Each organisation's home page was opened in turn. From the home page, potentially relevant sections were opened in a new window and explored in progressively smaller subsections. Where abstracts were available, these were used to screen papers, but where pages and papers did not have abstracts, full papers underwent an initial screen for clearly not meeting inclusion criteria and otherwise compiled for full paper review. These organisations often publish a series of linked papers describing a model (eg. an overview paper, a user guide, a technical document, appendices). In these cases, these added only one to the counts provided below, and the technical document initially used for the screening before using others if necessary..

Data collection process

After initial screening and de-duplication, 39 full papers were each then reviewed for all inclusion criteria. Appendix 1a.1, outlines the ways each of these 39 papers meets or does not meet final inclusion criteria. The title, first author and year are given for each paper, followed by the inclusion criteria.

Models that met the following criteria were included:

- Population: Any age group, living in England
- Intervention: Those relating to adiposity, and related dietary and PA risk factors
- Outcomes: Any health outcome, or healthcare or social care costs
- Time represented

Additional information was sought for included papers

- Modelling structure
- Risk factor inputs
- Physiological intermediate risk factors
- Disease burden output measures
- Costs included
- Accounting for future disease burden and costs
- If other societal benefits were included
- What user-defined options can be set
- How the model is accessed (described in Appendix 1a.2)
- How results are presented (described in Appendix 1a.2)
- Full descriptions of compliance with best practice criteria

Assessment of modelling best practices for included models

The ISPOR-SMDM (Professional Society for Health Economics and Outcomes Research and Society for Medical Decision Making) Modelling Good Research Practices Task Force recommendations for modelling best practice were used as the basis to assess the quality of each included model.(80) This includes an assessment of model validation. The relevance of each criterion to population-level NCD modelling criteria was considered and pertinent items were aggregated and summarised into those listed in table 2.4.

Table 2.4: Summarised ISPOR-SMDM recommendations for the assessment of modelling best practices

Domain	Criterion
Conceptualising the model To include spectrum of disease, perspective of model, target population (baseline), alternative interventions, what health and cost outcomes, what other outcomes	Stakeholder engagement undertaken. Statement of the decision problem provided, including perspective, objective and scope of modelling. Explanation for the choice of modelling structure. States represented can capture impact of intervention. The right balance of simplicity and complexity is made.
State transition models Including states reflecting disease process, important outcomes included, relevant population definition, cycle length, time horizon (assume lifetime unless justified otherwise)	Cohort simulation is used where the number of states is manageable. Model structure consistent with decision problem. Half-cycle correction applied. Clearly communicated (in non-technical language, tables and figures). Transition probabilities calculated transparently.
Parameter estimation and uncertainty	Parameters estimated on the basis of unbiased evidence, with formal synthesis. Clear distinction between uncertainty analysis and sensitivity analyses.

	Appropriate deterministic and probabilistic uncertainty analysis included, with appropriate distributions, communicated in standard statistical methods.
Transparency and validation Type of model and intended applications, funding, structure/ inputs/ outputs, other components and their relationships to model structure, data sources, validation methods and limitations	Technical and non-technical documentation available. Process for assessing face validity. Structural verification. Comparison with similar modelling analyses performed. External validation performed, including data sources, setup, simulated and observed results.

Results

Study selection

The numbers of papers assessed at each stage of each search approach is summarised in Figure 2.1. Details of how many titles and abstracts were screened for each thesis search phrase are given in Table 2.2. The numbers of papers screened for each page of each search phrase for the Google and Google Scholar search approaches are given in Table 2.3 (with ten links screened per page).

After de-duplicating and reading the 39 full papers, four papers met all inclusion criteria.

Included papers:

A table of all 39 shortlisted papers is included in Appendix 1a.1. This sets out which inclusion criteria were and were not met for each paper.

Four papers met all inclusion criteria. These were:

- School for Public Health Research (SPHR) Diabetes Prevention Programme (DPP) Return on Investment Tool v1.0(73)
- PHE Cardiovascular Disease Return on Investment Tool(74)
- NICE Estimating Return on Investment for interventions and strategies to increase physical activity(81)
- PHE Weight management economic assessment tool, version 2(82)

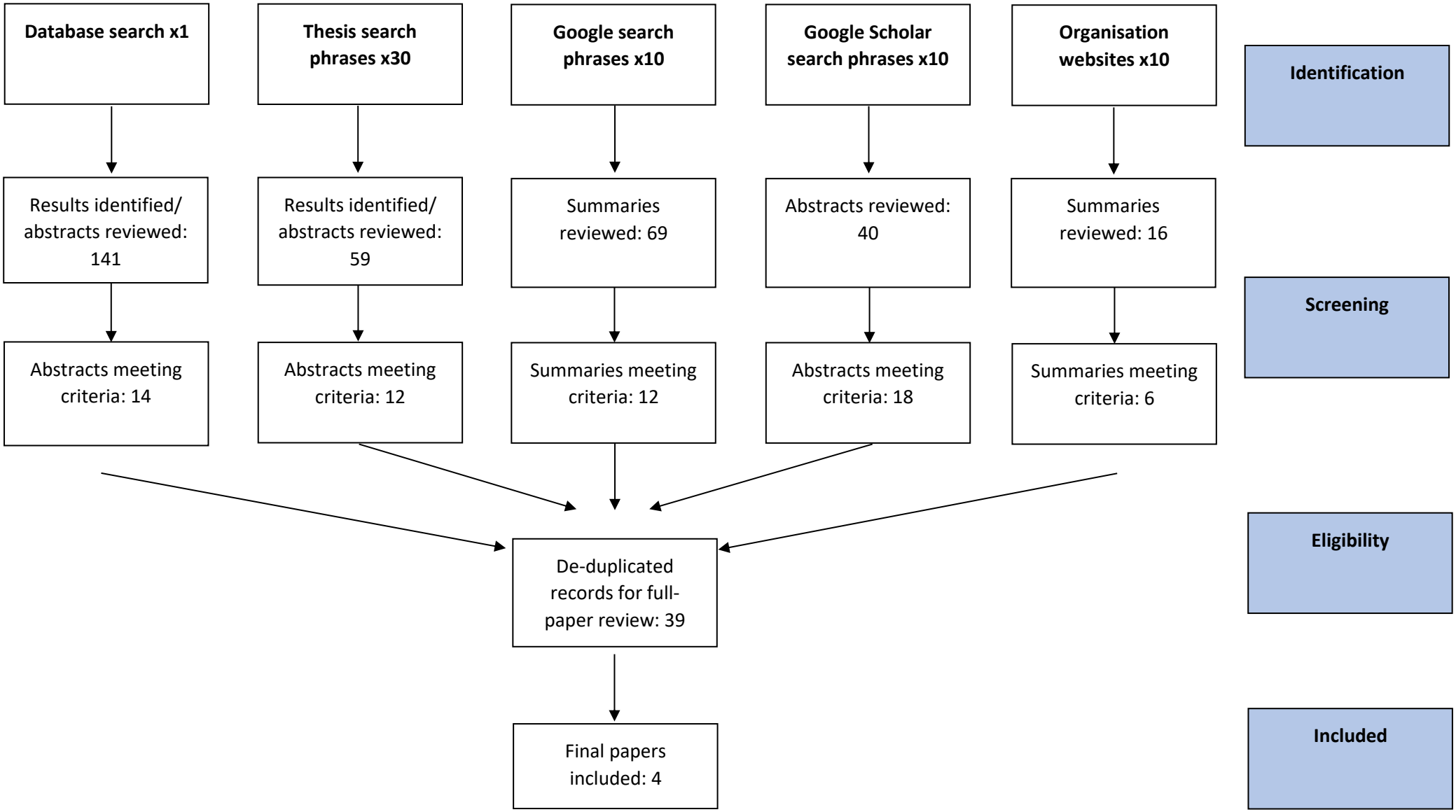


Figure 2.1: Flow diagram of papers at each stage of the review process, with PRISMA steps labelled to the right of the diagram

Study Characteristics

SPHR Diabetes Prevention Programme Return on Investment Tool

Description:

The DPP Return on Investment (RoI) Tool was funded by the National Institute for Health Research SPHR to a team at the School of Health and Related Research (ScHARR), University of Sheffield, with the explicit purpose to support the use of the NHS DPP. The tool is structured as a microsimulation, whose population characteristics are built from the Health Survey for England (HSE) 2011. It functions by creating new 'cases' of disease based on stochastic risk generated using UK clinical risk equations at the individual level (UK Prospective Diabetes Study (UKPDS) for Type-2 Diabetes Mellitus (T2DM), QRISK2 for first Ischaemic Heart Disease events and a health technology assessment on statin use for subsequent IHD events, as QRISK2 is designed for primary risk prediction only). Osteoarthritis is included by modelling BMI and T2DM as modifiable risks for the disease. Depression is modelled as a binary lifelong diagnosis, with increased relative risks related to T2DM and stroke. The inputs for these diseases are four major modifiable physiological risk factors of BMI, HbA1c, systolic BP and LDL cholesterol, plus age, sex, ethnicity, deprivation, waist/hip ratio, HDL cholesterol, smoking history and diagnoses of rheumatoid arthritis and AF. The four major physiological risk factors update on each annual cycle to reflect their progression with ageing, determined in the model by regression models built from trajectories in the Whitehall II study. Individual baseline breast and colorectal cancer risk trajectories were calculated from EPIC Norfolk cohort study and their risks related to BMI from a large meta-analysis. Healthcare costs are carefully attached to each disease state in a

granular fashion dependent on disease subtype (stable angina, unstable angina year 1, unstable angina year 2, stroke year 1, etc) using NHS Reference Costs, social care unit costs, the BNF and health economic reports from the peer-reviewed literature and the NHS. The model also estimates costs of disease to the economy from an employer perspective, by estimating the difference in the number of days absent from work (with a fixed mean daily income 'lost' from the economy) and the cost of recruiting a replacement after a death (a fixed cost per death).

The model is designed to compare the implementation of the DPP intervention in a local area to baseline. Users can vary which one of six pre-programmed high-risk groups are targeted (South Asian ethnicity, most deprived quintile of postcodes, HbA1c>42, BMI>35kg/m2, age 40-65 years, or estimated probability of developing diabetes in ten years >10%). The modelling allows users to compare cost implications (and perform CEA) of the DPP intervention at different levels of spend and targeting different at-risk groups.

Approach to local area modelling:

Local authority or Clinical Commissioning Group (CCG, the former local NHS commissioning bodies, that had approximately the same borders as lower tier local authorities) areas can be modelled, by population weights on the nationally-representative population. Weights were generated using iterative proportional fitting (IPF) using Census 2011 data from ONS and English Index of Multiple Deprivation 2015 (IMD). Demographics used in the IPF process were age (16 age categories for the local authorities, 4 for CCGs), sex, ethnicity (3 categories: white, Asian, other), deprivation (5 quintiles of IMD), diabetes prevalence. User-defined data can be used to account for local intervention effect, duration and costs.

PHE Cardiovascular Disease Prevention Return on Investment Tool

Description:

The PHE Cardiovascular Disease (CVD) tool was built as an incremental development of the above SPHR DPP tool by the same research group. The model is structured around six high-risk conditions: hypertension, AF, hypercholesterolaemia/ high cardiovascular risk, diabetes (type-1 and type-2), non-diabetic hyperglycaemia and Chronic Kidney Disease (CKD). The Diabetes module is taken in most part from the SPHR DPP tool. A similar synthetic population to the SPHR DPP tool is constructed from HSE 2014 and risk trajectories from the Whitehall II study. Risks are updated each cycle (age, BMI, BP, HbA1c, LDL and total cholesterol) and new cases stochastically arise based on clinical risk equations (QRISK2 for IHD, QStroke for stroke/ Transient Ischaemic Attack, UKPDS for T2DM and the Framingham AF Risk Equation, for AF). HSE 2014 was used for the baseline population with missing data imputed using clear procedures. Local population weightings were generated from ONS Census 2011 data and IMD 2015 data. Risk trajectories for the physiological risk factors were taken from Whitehall II study. Costs are taken from NHS Reference Costs, Personal Social Services Research Unit primary care reference costs, the BNF, and peer-reviewed articles for IHD and stroke. No social care perspective is included. Users can simulate either a user-defined intervention or one of 15 pre-set interventions identified through systematic review (NHS health checks, annual review, detection of familial hypercholesterolaemia and cascade testing, lipid modification, antihypertensive therapy, anticoagulant therapy, blood glucose lowering therapy, the DPP, structured diabetes education, weight management programmes, smoking cessation, nutritional advice for CKD, insulin pumps, pharmacist new medicine review, or blood pressure self-monitoring).

Approach to local area modelling:

Local authorities in England can be modelled, which alike with the SPHR model is achieved by applying population weights on a nationally-representative population. Weights were generated using IPF with data from the Census 2011 and English IMD 2015. Demographics used were age (16 age categories for the local authorities), sex, ethnicity (3 categories: white, Asian, other), deprivation (5 quintiles of IMD), diabetes prevalence. An association between deprivation and ethnicity was also added to account for the socioeconomic distributions of different ethnicities. User-defined data can be used to account for local intervention effect, duration, costs and baseline implementation.

NICE Physical Activity Return on Investment tool

Description:

Two Markov cohort models are used, one for children and one for adults. The adult model has a population based on ONS data, with Active People Survey data for PA. The population is divided into three PA levels ('inactive', 'low-active', 'high-active'), each with different Relative Risks for each disease state (CVD, stroke and T2DM), which are applied to age-specific baseline absolute risks. CVD/non-CVD mortality risks are applied (invariant by age or sex). No physiological intermediates are accounted for. For the adult model, users can model packages of 11 predefined interventions with emphasis varying between these. Each intervention impacts PA through pre-set percentages of participants who move from inactive to low-active or high-active after taking part in the intervention. Estimates of the percentages of the chosen local population in each PA category are saved in the tool. The tool can run between 2 and 48 1-year cycles (maximum age range 33-81 years). The number

of years in each condition are accumulated and utility weights are applied to calculate QALYs. After 10 years Relative Risks revert to those for the inactive state.

Technical information for the children's model is scarce. The children's model has two integrated interventions. It uses ONS data as its baseline but differs from the adult model in using HSE data for PA (categorised as 'inactive' or 'meets recommendations') and runs a single one-year cycle. Two pre-set interventions can be modelled, with fixed percentages of participants moving from inactive to the meets recommendations category.

Approach to local area modelling:

Local population characteristics were accounted for in terms of age and sex composition by allowing the model's input population to vary according to ONS local population estimates. Local PA measures were taken from Sport England's Local Sport Profile tool, which bases its estimates on the Active People Survey 2012. Baseline implementation of modelled interventions can be accounted for (user defined).

PHE Weight Management Economic Assessment Tool

Description:

The PHE Weight Management Tool (v2) functions around a cohort simulation with two levels of body weight: normal and high. It works by only simulating individuals included in interventions, with actual numbers being defined by the user on the basis of commissioning options. Users define the numbers of participants, their sex composition, mean age, mean BMI, mean BMI impact of the intervention, recruitment time, duration of intervention, dropout rate and time to return to baseline BMI, along with intervention costs. BMI is

simulated to decrease (for the scenario) then slowly increase with each cycle of the model and estimate disease incidence (diabetes, IHD, stroke, colorectal cancer, breast cancer) based on BMI. It runs separately for men and women, with a cycle length of 1 year and time horizon 25 years (with no explicit discussion over time horizon). Changes to disease incidence and mortality rates for the included diseases are the health outcome, with no further modelling of utility. Healthcare costs are calculated using a peer-reviewed method based on programme budgeting cost per case and social care costs are calculated by estimating the probability someone will need care directly from BMI, age and sex by the average cost of care per person nationally, without accounting disease-specific variations in social care costs. The probability of being in employment is calculated from age, sex and BMI based on logistic regression models built from the HSE 2011-13. The cost of unemployment is a fixed national average representing the benefit to the economy and the treasury, estimated by New Economy Manchester and the Department for Work and Pensions.(82)

Approach to local area modelling:

A local population is represented via user-entered data on population characteristics. Specific numbers of individuals can be entered, either numbers recruited to a healthcare intervention or a whole local population subject to change (eg. the effect of a sugary drinks tax over England or over a single local area). It is the user's responsibility how closely the population characteristics, intervention effect and intervention costs reflect those in a given area. SES is not represented in modelling but could be incorporated to intervention effect if users had access to relevant data that may vary by SES (dropout rate, effect size, effect duration). Costs to healthcare, social care and the economy use national averages.

Assessment of best practice criteria

SPHR Diabetes Prevention Programme Return on Investment Tool

Best practice criteria were well satisfied. Relevant risk and disease states reflect the disease processes it models, with a logical justification for its structure. It has been validated against longitudinal studies for IHD incidence and mortality, T2DM incidence and physiological risk factor trajectories. Probabilistic Sensitivity Analysis (PSA) is performed and presented clearly for exemplar scenarios, but is not available in the online tool due to the computational demand. Structural uncertainty explored via comparison of IHD outcomes with the UKPDS model. Methodological uncertainty is not included in sensitivity analyses. The online tool does not allow for any user-defined intervention to be modelled, rather a comparison between a given spend on DPP implementation versus no implementation. This inflexibility arises from the tool's purpose to support users in a specific commissioning setting with funds already earmarked by government; comprehensive economic analyses of broader competing options would be appropriate at the national government level. The only other best practice criterion not met is the absence of a half-cycle correction.

PHE Cardiovascular Disease Prevention Return on Investment Tool

In terms of satisfying the best practice criteria, there is no parameter uncertainty incorporated and there are no sensitivity analyses; no PSA is presented in the documents or included in the tool. It only includes a 20-year time horizon rather than lifetime, without an explanation, and there is no half-cycle correction. However, all feasible options are

modelled by limiting scope to individual-level options and identifying them formally by systematic review. A variety of other interventions including population-level approaches such as a sugary drinks tax are also modelled in the technical documents for illustrative purposes. Verification was performed for proportions of the population being tested/ treated remaining constant and that individual trajectories were behaving as intended. Comparison of MI/ stroke incidence against Hospital Episode Statistics (HES) identified that projected rates were far too high so the algorithms were adjusted to reflect current disease rates. Outputs were also compared with the previous version of the model (the SPHR DPP model). Population variation is accounted for in the individual-level simulation and the average of 500 model runs are used to account for stochastic variation.

NICE Physical Activity Return on Investment tool

A broad range of ISPOR-SMDM criteria were not met, ranging from the model lacking explicit stakeholder engagement, to missing an explicit justification of model choice, having no half cycle correction, no assessment of face validity, structural verification, model comparisons or external validation, lacking parameter uncertainty and using the concepts of "uncertainty" and "sensitivity" interchangeably (for example "the uncertainty posed by discount rates"). Otherwise, the model was transparently parameterised, including transition probabilities and utilities from good-quality sources.

PHE Weight Management Economic Assessment Tool

Best practice criteria are missing in terms of lacking a clear description of how the model is structured, weak user and technical documentation, only representing mean sample BMI, and not reporting any sensitivity analyses. Some deterministic sensitivity analyses are suggested, but no PSA is possible owing to the model not incorporating variability, parameter uncertainty or stochastic uncertainty. There is also no stakeholder engagement, a half-cycle correction or assessment of face validity.

Discussion

Summary of evidence

This review aimed to identify models that are capable of modelling change to dietary, PA and adiposity risk factors and NCD burden in local areas of England. It identified four models capable of modelling local areas of England. These were produced or commissioned by government bodies for modelling around specific interventions.

The PHE CVD tool and SPHR DPP tool were based on microsimulations and the NICE Physical Activity RoI tool and PHE Weight Management tool were based on Markov cohort models. The microsimulation structures allowed individual-level age-related change to physiological risk factors to be incorporated and for disease cases to arise from individual-level clinical risk prediction equations (such as QRISK2).(83) These two models and the Weight Management tool did not include behavioural risk factors, whereas the NICE PA tool includes PA. None included dietary risk factors. The simpler structures of the Markov models

involved the trade-off of not representing the full range of intermediate physiological risk factors: the PHE Weight Management tool had an input for BMI effect, but the NICE PA tool linked PA directly to disease rates without linking via BMI/ overweight/ obesity, HbA1c, blood pressure or cholesterol.

The tools all include the common diseases of T2DM, IHD and stroke. The CVD tool adds major bleeds and congestive heart failure to these, while the Weight Management tool adds breast cancer and colorectal cancer. The DPP tool breaks down these diseases and costs into subgroups (eg. stable and unstable angina, year 1 costs and subsequent costs, etc) and includes other diabetic complications. The DPP tool also only models the prevention of diabetes and its consequences in those with impaired glucose tolerance, not accounting for the impacts of changes to T2DM risk factors on other diseases such as IHD/ stroke in non-diabetic individuals, so cannot model population-level interventions.

All the tools include healthcare costs and the DPP tool and Weight Management tool also estimate impact on social care costs. The DPP tool, NICE PA tool and Weight Management tool account for costs to the economy, in different ways. All four models presented their results clearly.

Regards best practice criteria, the DPP and CVD tools both met most criteria, whereas the NICE PA tool and the PHE Weight Management tool did not. Both the DPP and CVD tools were successful due to their detailed structures populated with good-quality data, including physiological intermediates and careful cost attributions. Both are limited by excluding some obesity-related diseases such as cancers.

The two simpler models did not explicitly discuss the choice of model structure, though the Markov cohort model is the default recommended by the ISPOR-SMDM Task Force.(80) The

states represented are limited and neither includes heterogeneity of risk factor exposure in their populations other than by age and sex.

Completing the full complement of uncertainty analyses, validation and deterministic sensitivity analyses is a substantial task, so it is not surprising that none of the included models presented all of these. Given the CVD tool largely met other best practice criteria, that it does not include parameter uncertainty and PSA constitutes its major limitation. This model even managed the difficult criterion of modelling all feasible options, by defining the scope of feasibility via systematic review.

The NICE PA tool offers no modelling of uncertainty, no assessment of face validity, no structural verification and no discussion of external validation. The weight management tool had structural verification and external validation processes. The PA tool and Weight Management tool also have weak technical documents.

Approaches to modelling local areas varied. The SPHR DPP tool and PHE CVD prevention tool both used population weights based on age, sex, ethnicity and deprivation, on a nationally-representative microsimulation, to represent subnational heterogeneity. The NICE PA tool used locally-specific populations by age and sex, and local-specific PA data. The PHE Weight Management tool used the approach of requiring the user to enter appropriate local data.

Strengths and limitations

The review took a thorough, systematic and transparent approach to rationalising the grey literature, including using multiple overlapping approaches to maximise the chances of

every potential candidate model being identified. Each component of the review was designed to be as thorough and comprehensive as possible. The five arms of review remained structured along the steps of the PRISMA Systematic Review approach and the numbers of papers at each step monitored. Models were identified for inclusion using a consistent approach that accepted the variation in how models may be presented by different bodies. Each identified model was described in detail and assessed based on gold-standard best practice criteria.

It could be criticised for not taking a Systematic Review approach to the formal peer-reviewed literature, foregoing potential benefits such as the increased reliability of double assessment and allowing greater reproducibility of that branch of the search strategy. The other arms of the review would not be reproducible by any design, with the possible exception of the thesis search. The Google and Google Scholar searches are not reproducible as the algorithms determining their results both change over time, while hand searching websites involves an element of arbitrary choice in approach.

As this is not a formal systematic review, no review registration was made and no formal protocol was published in advance, though where changes have been made to the approach (such as over age groups included) this has been transparently documented.

Conclusion

NCD health economic modelling tools for local areas in England are scarce and variable in quality. One of the identified models has been retired due to limitations (the PHE Weight

Management tool) and one other also has significant limitations (the NICE PA tool). The remaining two are both of good quality. What broadly divides the more and less successful models is that the more successful have more complex structures including for more disease states and intermediate risk factors, and that they paid greater attention to uncertainty, population variation, validation processes and structural verification. Though complexity alone does not indicate a superior model, in these cases it has conferred benefits. Of the two better-quality models, the CVD model structure is more representative of disease burden and healthcare costs, but lacks social care costs and exploration of uncertainty including PSA.

There is a gap in local-area NCD modelling for the development of a new tool. Areas to build on current modelling capabilities include: physiological risk factors; population risk heterogeneity; healthcare, social care and a wider social and economic perspectives; the use of locally-collected data, and comprehensive validation and sensitivity analyses.

Chapter 3: PRIMEtime_local – model description

Introduction

This Chapter provides a description of PRIMEtime_local, including a technical description of the model's structure and function. The estimation process for most of its local-level parameters are described in Chapters 4-7, though those with less detailed methods are described in this chapter.

As discussed in Chapter 1, local-level modelling would be valuable and this requires local-level parameters to be estimated and structured into a model. This new local-level model will be built on the PRIMEtime model structure, which uses a PMSL approach, developed by Cobiac, Barendregt and colleagues.(46–48) PRIMEtime was originally based on PRIME (the Preventable Risk Integrated ModEl),(84,85) a comparative risk assessment (CRA) model, now available open source. The innovation of PRIMEtime was to move the static CRA into a dynamic model by embedding the principles of PRIME into a lifetable, allowing the estimation of cumulative effects over time. The structure allows multiple disease endpoints to be modelled on the basis of one or more risk factors, giving it particular strength for modelling health and cost implications of non-communicable diseases (NCDs). PRIMEtime has been used for many papers analysing NCD risk factor interventions, including for sugar, salt, obesity, tobacco, PA, dietary consumption of fruit and vegetables, salt, and red and processed meat, across many countries, including Australia, Finland, France, Italy, New Zealand, Sweden, the UK and Vietnam.(47,49–53,86–88) In PRIMEtime local, local-level

parameters will populate the PMSL structure in turn, depending on the local area to be modelled. When the simulation has finished for one area, the model then populates with data for the next required area and begins a new simulation.

Conceiving of the model

PRIMEtime_local was conceived of through the logic outlined in Chapter 1. This was that much public health decision-making had been devolved down to local authority level and there was increasing focus on understanding the impact of public health interventions on geographical inequalities. The PRIMEtime_local model can model point estimates for all 315 lower tier and unitary local authorities in England (excluding the Isles of Scilly and City of London due to size(2)) or PSA for up to ten areas at a time.

The model structure was designed around both of these needs with stakeholder engagement to help shape its development. PRIMEtime_local had two streams of stakeholder engagement undertaken, one with patients/ the public (October 2018 and November 2019), and one with policymaker stakeholders (November 2018 and November 2019). Five members of the public volunteered to contribute through a local General Practice (GP) and six policymaker representatives volunteered through the local authority public health public engagement officer and through Manchester CCG. The policymakers were the public engagement officer, two public health consultants, a PA public health practitioner, a public representative and the chair of the CCG. The purpose of this was to gain an understanding of how stakeholders viewed and understood modelled evidence in the field of public health, what outcome metrics were most valuable and what interventions

were of interest to be modelled. The patients/ public and policymakers felt the aims of the model were worthwhile and it was a useful piece of research. Policymakers felt this work and its principles could be used widely to support public health decision-making and that cost was the most important outcome measure to model. The public were asked about what kinds of interventions were of interest and they strongly supported the modelling of population-level changes to the food system. Examples such as product reformulation, sugary drinks taxes and changes to trade regulation were discussed. Previous work(72) has also explored stakeholder attitudes to comparable scenario modelling using the PRIMEtime model, collecting attitudes from representatives of national and local government, charities, healthcare professionals and academics. This focused mainly on analytic choices such as time horizon and format of modelled outcomes, with input on model relationships and face validity. Of 12 expert responses, 9 agreed there was broad face validity and the remaining three fed back with specific issues that they felt would improve the face validity – that more interventions should be modelled and that economic productivity and social impacts should be included.

Model structure – the Proportional Multistate Lifetable Model

PRIMEtime is a PMSL(46,48,89) and the specific version of PRIMEtime used as the basis for this thesis has been used elsewhere.(47) PRIMEtime is structured into three components: 1, the risk factor module; 2, the disease models; and 3, the lifetable.

1. Risk factor module

The risk factor module is where the user sets the baseline and scenario levels of risk factor exposure, by age and sex. The key principle of PRIME and PRIMEtime is the Population Impact Fraction (PIF) for each risk factor-disease relationship, which represents the effect of changing risk factor exposure on disease incidence. A PIF is calculated for each year of age for each sex, as:

$$PIF = \frac{\int p\ RR(x)\ dx - \int p'\ RR(x)\ dx}{\int p(x)\ RR(x)\ dx}$$

where x is the level of risk factor exposure (here, BMI units as kg/m^2), RR is the relative risk between that risk factor level and a given disease, p is baseline prevalence of that risk factor level and p' is the scenario prevalence for that risk factor level.

The specifics of how the PIF is calculated and operates in PRIMEtime have been detailed elsewhere,(46,85,90) but are summarised here. The PIF links behavioural or physiological risk factors (specifically here, BMI) to disease outcomes (for example ischaemic heart disease (IHD), strokes, cancers, T2DM) using their Relative Risks. Risk is quantified in terms of how many people are exposed to a risk factor, and how much of the risk factor they are exposed to. A risk factor such as smoking has a minimum risk of zero (people who have never been exposed to tobacco smoke) but there is no zero BMI, so risk factors such as this are framed in terms of their theoretical minimum risk – the value (or value range) associated with the lowest risk of disease. These relative risks between a risk factor and a disease outcome may also vary across the range of amount of risk factor exposure. PIFs for

more than one risk factor are combined multiplicatively (here, for example, PIFs for BMI-IHD and T2DM-IHD) as:

$$PIF_{combined} = 1 - \prod_{i=1}^{n}(1 - PIF_i)$$

Where i is a risk factor, for 1 to n risk factors. The probabilities of developing each disease are treated as independent from one another in the lifetable. To account for the violation of this assumption in the case of T2DM-IHD and T2DM-stroke relationships while avoiding double counting, an adjustment factor is applied, estimated for previous work on PRIMEtime.(86,91)

The user defines the population size, incidence, prevalence and case fatality rates from modelled diseases and the baseline distribution of risk factor prevalence (all by age and sex). Risk factor distributions are defined as a mean and SD, capturing spread as well as the average risk exposure. The scenario risk factor exposure is then set, providing a difference in exposure to be followed through to differences in disease mortality.

Change to BMI is calculated via change to weight (using the standard equation BMI = weight/ height^2). Change to weight is calculated from change to calorie consumption (constituting the specified intervention) using equations by Hall *et al*. These are calculated for adults (over 18 years) as 1kg weight change per 100kJ energy intake change per day, with 50% of the weight change happening within one year and 95% in three years.(92) For children (up to and including 18 years), this is calculated as 68-2.5*age for males and 62-2.5*age for females.(93)

Technical choices for calculating the PIF

Relative Risk may also change across the intensity of risk factor exposure, for example BMI has an increasing marginal risk with IHD: a one-unit increase in BMI from 30 to 31 has a greater Relative Risk for IHD than from 25 to 26. In the real world, both risk factor levels and relationship between RR and level are often continuous, but it is computationally easier to calculate them split into groups aggregated as ordinal levels. Previous work(90) has explored the impact of using the computationally easier approach and found that framing RR as categorical leaves large non-linear artefacts on PIF compared to the mathematically ideal approach of treating both as continuous, whereas treating risk factor levels as categorical left only very small differences from the ideal. Here, therefore, risk factor exposure is treated as categorical, but split into 14 groups to minimise risk of bias (BMI categories of <15, then 2.5 units wide up to a BMI of 40, then 40-50, 50-60 and >60).

2. Disease models

The disease models implement the age- and sex-specific PIF into changes in disease-specific disease burden. These simulate a thousand individuals from age zero to 100 years. They can exist in one of three states only: 'No disease', 'disease', and 'dead from disease'. State transition parameters are incidence rate, moving people from 'no disease' to 'disease' and case fatality rate, moving people from 'disease' to 'dead from disease'. At each age, these state transition probabilities calculate the proportions of people reallocated from healthy to diseased and diseased to dead across the life course.

The change to disease prevalence that results is then used to calculate the change to utility weight, applied additively to baseline average utility score and multiplied by the total

population to calculate change to QALYs. Healthcare costs are counted as either prevalent or incident costs, for diseases with costs that accumulate over time or those that are associated with a single treatment event, respectively, as discussed in more detail in Chapter 7. Incident costs are the multiple of disease unit cost with change to number of incident cases, while prevalent costs are the multiple of the unit cost with the number of life years lived with the disease. These healthcare costs can be modelled using local-specific or national average unit costs (from Cobiac *et al*(47)).

A static lifetable would represent a current population, with morbidity and mortality captured for a point in time. To allow the estimation of effects across the whole lifetime for the whole population, the model runs cohorts in five-year age bands. For instance, if the age range 40-59 years is of interest, the model will run the 40-44 year olds first by populating the disease models and lifetable with only that population. The sheets then calculate the impact of time on their progression to disease and death by applying incidence and case fatality as described above. By age 100 the remaining population is very small and a mortality rate of 1 is applied, ending the cohort. The 45-49 year olds then have the same process applied, then 50-54 year olds, and so on. This allows the effects on multiple cohorts to be captured in the model that take place simultaneously in simulated time. After each cohort is run, that cohort's modelled impacts on numbers of disease cases, QALYs and costs are stored in an output array and the next age group is run. Once all age groups of interest have been run, they are added together into final modelled output figures.

As discussed more fully in Chapter 4, there are eight chronic diseases that lead to mortality in the model (IHD, stroke, hypertensive heart disease (HHD), T2DM, AF/ flutter, colorectal cancer, breast cancer, oesophageal cancer and asthma) and three that exert morbidity

alone (osteoarthritis of the knee and hip, and low back pain). Asthma is alone in being framed as 'acute', with most of its morbidity relating to exacerbations (asthma attacks) rather than ongoing symptoms, and these acute attacks also have mortality applied. The age that BMI is associated with these disease outcomes arises from the papers describing the evidence, listed in Appendix 1b, specifically that IHD, stroke, T2DM and HHD are increased from age 35, osteoarthritis of the knee from age 50 and the remainder lifelong.

A lag is applied between the change to BMI and the change to disease risk. According to the WHO, stroke risk is fully altered by BMI at five years, but it probably takes ten years for IHD.(94) It is likely cancer risk takes longer, based on follow up from bariatric surgery and other weight loss interventions.(95) Therefore a lag of 5 years is applied to stroke, 10 years to IHD and 20 years to cancers. No lag is applied to risk for the remaining diseases.

3. Lifetable

The lifetable compiles the outputs of the disease models. The 'true' population size is used, rather than an arbitrary 1000 individuals at the start of the cohort. There is no disease-specific morbidity or mortality in the lifetable – it is restricted to states of 'alive' or 'dead' (from any cause). All-cause mortality is applied at each year of age, so the population shrinks down the table in line with that in the general population. Average utility is also applied, and quality-adjusted person years lived then calculated. This is done twice in parallel, using first baseline then scenario inputs. Baseline all-cause mortality rates are taken from the Global Burden of Disease study (GBD)(96,97) then the scenario all-cause mortality is calculated by additively combining this with the changes to disease-specific mortality from

the disease sheets. The same is the case for total utility rate being calculated via baseline utility plus changes to disease-specific utility rates. Utilities are used to calculate QALYs lived for baseline and scenario, then finally the numbers of cases, costs and QALYs are netted out from the scenario (ie. the difference in these outputs between the baseline and scenario model runs are calculated) to calculate the net effects of the intervention on these summary outcomes.

Input parameters

National-level parameters

The utility score inputs were those estimated by Sullivan *et al*(98) for the UK, following the method originally used in the USA.(99) These are derived from EQ-5D-3L responses (a generic instrument for patient reported outcome measurement that can assess patient's health-related quality of life(100) from the US annual Medical Expenditure Panel Survey 2000-03, using the UK-based preference set that used a 'Time-Trade Off' method to value the responses.

Risk factor-disease relationships are quantified using Relative Risks. These are derived from Meta Analyses (MA) of Prospective Cohort Studies (PCS) examining the relationships between BMI and given disease outcomes. The specific studies used are listed in Appendix table 1b.1.

Local level parameters

The process for estimating locally-varying parameters is described below in Chapters 4-7. The populations are taken from Office for National Statistics (ONS) and single year estimates for ages over 90 years olds are smoothed out from the local estimates of populations 90 years and over. As described in Chapter 4, epidemiology is derived from GBD data. All-cause mortality is taken from the GBD local authority estimates and aggregated as population-weighted averages at the local authority Index of Multiple Deprivation 2019 (IMD) quintile level. Disease incidence and prevalence are also taken from the GBD aggregated to the IMD quintile level. These are then used as inputs to the Disbayes model(101,102) that estimates consistent sets of incidence, prevalence and case fatality, which are used as PRIMEtime inputs. The effects of the scenario intervention can also be set as national or local estimates.

Populations are taken from the ONS. Data on populations by year and sex are available at the local authority level (the mid-year population estimates) up to age 89 years. Populations aged 90 years and above are not published by the ONS at the local authority level due to the unreliability of small numbers, but they provide an aggregated value for those 90 and over. Reliable year-specific populations are available at the national level, based on the national register of births and deaths via the Kannisto-Thatcher method,(103) so local populations of over-89s were estimated by apportioning national year-specific populations by the proportion of the national population of over-89s in each area. For example, if an area has 1% of the over-89s, then 1% of the 90 year-olds would be apportioned to that area, 1% of the 91 year-olds and so on. Details of this process are provided in Appendix 2b.

Analytic options

A variety of analytic options are available to the user, summarised in table 3.1.

Table 3.1: Analytic options available for PRIMEtime_local.

Analytic option	Definitions	Choices
Discount rates	The annually-applied weighting given to outcomes closer to the present.	Any can be defined. Pre-set options of NICE public health rates (1.5% for health outcomes and 3.5% for financial outcomes) or UK Treasury Green Book rates for health of 1.5% for the first 30 years, 1.29% for 31-75 years and 1.07 thereafter, or for costs 3.5%, 3% and 2.5%, respectively.
Open or closed cohorts	Open cohort introduces new members of the population at the bottom of the age range. A closed cohort does not introduce new members, instead modelling only a fixed population.	Open cohort. Closed cohort.
Time horizon	The length of time that an intervention and its effects are simulated over.	1-100 years for a closed cohort 1-135 years for an open cohort
Trends	Assumptions to how baseline input values are likely to change in future.	Incidence and case fatality/ mortality rates
PSA	Whether or not parameter uncertainty should be captured in modelled	Full complement of 315 local authority area point estimates (no uncertainty).

	estimates via Monte Carlo analysis. Uncertainty distributions are normal for Relative Risks and for utility weights for disease, age, sex and adult BMI have normal uncertainty distributions, while utility weights for child BMI have a beta distribution. Relative Risks and costs have lognormal uncertainty distributions. Intervention effects can have any uncertainty distribution appropriate to the intervention.	Approximate run time 3 minutes. 1-20,000 iterations of uncertainty estimates for 1-10 areas. Approximate run time 12 minutes per 1,000 iterations.

Conclusion

PRIMEtime_local takes a PMSL structure and can be used to estimate the effects of BMI interventions on local areas. The model overcomes challenges with the PMSL structure of lacking subgroup granularity. Local-level inputs are population, BMI distribution, healthcare costs, disease epidemiology (to the IMD quintile level) and intervention effects and the intervention effect (if available).

Chapter 4: Estimating local disease epidemiology

Introduction

The PRIMEtime model structure uses data describing the epidemiology of each disease being modelled (usually incidence, prevalence and case fatality rate). Therefore, it is useful to have local-level disease epidemiology to capture the contribution of its real-world variation to local differences in the impacts of interventions. The selection of disease states is explained below, then for each modelled disease these epidemiological parameters need to be estimated to populate a disease model (a disease-specific lifetable) as explained above in Chapter 3. These sets of epidemiological data have a few requirements. The disease model for each chronic fatal disease (namely IHD, T2DM, stroke, AF, oesophageal cancer, breast cancer and colorectal cancer) allows people to occupy one of three mutually exclusive and collectively exhaustive disease states: 'disease free', 'disease', or 'death from disease'. There are a series of diseases that do not exert mortality (low back pain and osteoarthritis of the hip and knee) so only the former two disease states are allowed in these cases. Asthma is modelled as an acute condition, with mortality dealt with differently, explained below. Local-specific all-cause mortality is also required in the PRIMEtime structure.(46,47)

Each fatal chronic disease needs a baseline prevalence, incidence and case fatality rate. These are parameterised as probabilities, with incidence representing the transition probability of someone in the 'no disease' group moving into the 'disease' group between

years t and t+1, prevalence representing the probability of a random member of the living population being in the 'disease' state in a given year t, and case fatality represents the additional probability of death between years t and t+1 for those with a disease compared to those without it. This is schematically represented below in figure 4.1, where *i* is incidence rate, *r* is remission rate, *f* is case fatality and *a* is a given age. In Disbayes, as in PRIMEtime, remission *r(a)* is always assumed to be zero and case fatality *f(a)* for the non-fatal diseases assumed to be zero.

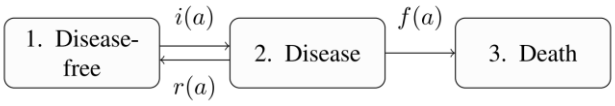


Figure 4.1: Schematic diagram representing the three-disease state model and transition parameters. Reproduced from Jackson *et al*, 2021 (102) under Creative Commons CC BY License.

The 'disease-free' state represents the pool of people at risk of disease, the 'disease' state represents population prevalence (of having ever been a case) and the 'death' state represents *disease-specific* deaths. The impacts of multiple diseases are combined later in the PRIMEtime structure.(47,88) The assumption that remission equals zero is based on the framing of disease prevalence as *of having ever had* the condition, ie. it includes current cases *and* cases in remission. As case fatality rates are based on deaths reported (across the whole population), these are adjusted downwards to accommodate the larger denominator populations.

Case fatality rates are needed rather than mortality rates, as case fatality represents the state transition probability of moving from the 'disease' state to the 'death' state, which

needs to remain consistent if prevalence rate changed, which would not be the case if population-level mortality rates were used. This chapter describes the process of using an established approach (Disbayes)(101,102,104,105) that is required to estimate internally-consistent sets of case fatality, incidence and prevalence rates for diseases to be used in PRIMEtime_local. Here, incidence and case fatality rates are from the GBD (per 100,000 people per year) with their numerator and denominator populations.

The global burden of disease approach

The GBD studies provide a rich source of epidemiological data for most countries, designed to give directly comparable sets of data across areas. GBD 2019 analyses 369 causes of disease or injury, 286 causes of death and 87 risk factors across 204 countries including the UK,(96) with subnational analyses for England and each of England's 150 upper tier local authorities (excluding the City of London and Isles of Scilly).(2) An important feature of the GBD data is they are *internally-consistent*, which is a feature also required in PRIMEtime modelling. This means that the logical relationships described above in figure 4.1 are respected in the estimated incidence, prevalence and case fatality rates used in the model. For example, across time, the sum of all new cases minus all deaths equals the prevalence.

The local authority GBD data for England provide incidence, prevalence and disease-specific mortality in terms of rates, numbers or percentages, but does not provide case fatality rates. The data are synthesised from multiple data sources including a wide range of academic literature, surveys (including the HSE and the WHO's UK Health Survey) and NHS administrative records (such as HES and Hospital Patient Discharge Data).(2)

The Disbayes model

Disbayes is a model that can be used for estimating case fatality rates by exploiting the logical relationships described in figure 4.1 from incidence, prevalence and mortality rates. This was developed by Jackson *et al*,(101,102) at the Centre for Diet and Activity Research, University of Cambridge, building on the broad approaches of DISMOD, DISMOD-II and DISMOD-MR models,(89,105,106) which have been used by the GBD and other epidemiological studies.(48,96,107) These were built primarily to estimate consistent sets of epidemiological parameters from primary data where these raw data for each parameter were estimated from different studies, so were not necessarily consistent due to bias or sampling error. Raw data cannot be used in a lifetable if they have not been estimated together as they can produce illogical results, for example if incidence was too high or case fatality too low, then prevalence rate could escalate far beyond that seen in the real world. The DISMOD packages were built to overcome this issue by estimating consistent sets of data while keeping as close as possible to the raw data. The DISMOD-II package uses an assumption that age-specific case-fatality rates are stable over time, so case fatality and incidence rates are purely a function of age. More recently, these packages developed capabilities to estimate the values of missing parameters by exploiting their logical interdependencies outlined above in figure 4.1.(102,104)

In terms of the methods, DISMOD-II broadly uses a maximum likelihood approach via minimising squared error and DisMod-MR takes a Bayesian approach (put forward by Flaxman 2015)(102,104)). The aim of developing Disbayes was to closely reflect the capabilities of DISMOD-II, using a Bayesian approach put forward by Flaxman (2015), while relaxing some limiting assumptions.(101,104,105) Specifically, Disbayes avoided

assumptions that: 1. Error variances were constant with age, 2. Time trends to incidence or case fatality rates applied equally to all ages, and 3. There were no non-linear associations between age and the parameters being dealt with (implemented in Disbayes with spline functions). Relaxing the second and third of these assumptions also allowed the modelling of age and calendar year to be combined, allowing variation through time to be captured. The compromise taken is that the Bayesian approach is much more computationally expensive.(101,102)

The process of operating the Disbayes model requires four steps:

1. Record the theoretical disease model (what is the natural history of the disease to be captured)
2. Record the statistical model (how is the theoretical model to be estimated from existing data)
3. Define priors for unknown parameters (explained further below)
4. The Disbayes model then calculates the posteriors

The disease model allows people to occupy one of three states via four transition parameters, as outlined in figure 4.1. Disease-specific mortality is assumed to be independent from death by other causes, so all-cause mortality is ignored (because of the assumption of independence, individuals are equally likely to leave the model from background mortality if they are in the 'no disease' or 'disease' state – therefore background mortality does not affect transition in the disease-specific models). Remission, too, here, is framed as zero, meaning the disease process can be entirely defined by incidence and case fatality, such that:

$n_a = N_a – P_a – M_a$

$P_{a+1} = n_a * I_a$

$M_{a+1} = P_a * F_a$

where n is the population at risk (disease free), N is the total population (across all three states), P is the number of prevalent cases, M is the number of deaths from the given disease, I is the incidence probability and F is the case fatality probability, each at a given age, a. These absolute numbers can be converted into rates and probabilities of someone from the baseline population entering the disease state. The missing parameter (usually incidence or case fatality rate) can then be calculated, using a defined prior distribution if available. The Bayesian process then provides a posterior distribution, given the prior, data inputs and assumptions described above. This is implemented through the stan package,(101,102) in these analyses via R 3.1.2 in RStudio.

The process starts with any temporal disaggregation required, estimating smoothed year-specific rates or numbers from 5-year age bands. Disease state numerators and their uncertainty (ie. absolute numbers in each of the three disease states allowed) can be estimated from a rate point estimate (proportion in the state), standard error of the mean (SE) (of that point estimate) and denominator (population or sample size). This process assumes a prior beta distribution Beta(0,0) and posterior calculated as Beta(r,n-r) where r is the point estimate given above and n is the interval of the SE. The prior can be set as vague (entirely unknown), a constant with age, a smoothly increasing linear function of age or a smoothed non-linear function of age (linked via spline functions – the preferred option). An age can be provided, below which the output estimate (case fatality or incidence) can be estimated (a requirement where the prior distribution is vague).

Using a Markov Chain Monte Carlo approach, the Disbayes package iteratively draws values, calculates possible case fatality probabilities and tests their combined fit, with the best fitting median value and its uncertainty is estimated. Outputs are then given in terms of rates, rather than probabilities.

Selection of disease states

Diseases were chosen for inclusion in the model on the basis of causing the greatest BMI-attributable disease burden. To identify these, the GBD 2019 was used to rank diseases ("causes") from the Category B of the GBD Cause taxonomy, making up NCDs. An arbitrary cut-off of accounting for 90% of BMI-attributable disease burden was chosen, making up the top 11 diseases, other than CKD and dementias.

CKD was excluded as only a small minority (<1%) of all CKD is end-stage CKD (that treated with dialysis or kidney transplant), with the remaining CKD exerting the majority of its morbidity and mortality via acting as a risk factor for IHD and stroke. Therefore, it is not included to avoid double counting.(108) It was also chosen not to model the impact of BMI interventions on dementias. The links between BMI and dementia are less well established than for other diseases and there is also a large overlap between dementias and stroke (termed vascular dementia) that could lead to double counting. Dementias are also likely to affect social care more than healthcare costs and it may exert additional healthcare costs and disease burden via other diseases than due directly to its symptoms. For example, people may present late (or not at all) with other conditions (for example, stroke or IHD) that lead to increased disease burden. Due to the poor understanding of how dementias link

to BMI, other diseases and healthcare costs, it was chosen not to include these as a disease state.(86)

Breast cancer is one of the five diseases linked directly with PA in the GBD so is useful to include in addition to the above, so that all these related diseases are also included. Breast cancer is also in the top ten causes of BMI-related death.

Therefore, the final list of included diseases is: asthma, low back pain, osteoarthritis of the hip, osteoarthritis of the knee, IHD, stroke, HHD, T2DM, AF and flutter, colorectal cancer, breast cancer and oesophageal cancer. Atrial fibrillation and flutter are referred to throughout as AF, and the term IHD is used for to refer to the spectrum of coronary arterial diseases that includes myocardial infarction and angina, though in places the term CVD is used instead when referencing other researchers' work using the term.

The Index of Multiple Deprivation

The English Index of Multiple Deprivation 2019 (IMD) is used in the development of the epidemiology in this chapter. The IMD is a score-based system estimated every 3-5 years, to describe the level of deprivation in local geographies of England. It is calculated from indicators across seven domains: income (such as the proportion on income support), employment (eg. proportion of adults on unemployment support), health and disability (eg. years of life lost to premature death), education (eg. proportion of over-16 year-olds not continuing education), barriers to housing and services (eg. proportion of overcrowded houses), crime (eg. frequency of violent crimes) and living environment (eg. frequency of pedestrian/ cyclist traffic accidents). These are then weighted and aggregated for each neighbourhood (Lower-Layer Super Output Area, LSOA, of mostly 1,000-3,000 residents).

These can then be aggregated to higher levels of geography, such as upper or lower tier local authority.

Methods

Chronic fatal diseases

Disbayes implementation

The Disbayes model was used to estimate consistent sets of incidence, prevalence and case fatality rates for IHD, stroke, T2DM, oesophageal cancer, colorectal cancer and AF, and Breast cancer for females only.

Due to limitations around runtime of the Disbayes package, it was not feasible to run the estimation process for all 150 upper tier local authorities separately, so instead areas were aggregated to the English Index of Multiple Deprivation (IMD) 2019 quintile level and diseases modelled for these five levels instead. (Test runs on the full set of diseases for females only in one local authority and one year implied that running the process for 150 local areas, two sexes, for 2019 and 2009 (to estimate trend) separately would have taken approximately five months.)

Prevalence, incidence and disease-specific mortality rates were taken from GBD 2019 and population-weighted average rates estimated at the IMD quintile level by age and sex, using populations from the ONS(109) (aggregated to IMD quintile). Disbayes was run in turn for

each quintile, for each chronic fatal disease. The priors were set to a non-independent smoothly non-linear function of age and with a constant case fatality value below 35 years of age. Trends were estimated as a tenth of the log of the proportional difference between 2019 and 2009 (ie. trend = $(\log(rate_{2019}/rate_{2009}))/10$.

Other parameters

Some diseases do not need the same treatment of estimating case fatality rates through Disbayes. The non-fatal diseases – low back pain, osteoarthritis of the knee and osteoarthritis of the hip – do not have mortality or case fatality rates. They have their baseline incidence rate and prevalence rate data input into PRIMEtime directly from the results of the GBD 2019 study.

As asthma has a high remission rate (that is also difficult to measure), keeping the assumption in place that remission equals zero would result in case fatality rates estimated by the above methods to be very suppressed. It is therefore modelled in terms of its acute attacks, with no prevalence rate. This means that incidence and mortality rates (ie. deaths per total population, not deaths per prevalent cases) can be applied directly to the whole population (by age and sex). (Change to incidence and mortality rates are modelled as a direct multiple of *1- PIF*, explained in Chapter 3.)

All-cause mortality rates were also taken from GBD 2019 (by age and sex) for maximum consistency with the disease-specific data. Hypertensive heart disease has no local-level estimates of incidence so England-level estimates from previous work by Cobiac *et al*(47) were used. As no new synthesis has been undertaken for the epidemiology of musculoskeletal diseases, asthma or hypertensive heart disease, these data are not further

presented or discussed in the remainder of the chapter, though some graphing is provided in Appendix 1c.1.

Examination of data

- Examination by deprivation: To assess how Disbayes output results varied between IMD quintiles, data were averaged across the age range using a standard population of England, by sex, for each disease and by incidence, prevalence and case fatality rates. These were then compared by calculating percentage differences from IMD quintile 5 (most deprived). To assess the impact of age structure on disease epidemiology, this was also repeated using local-specific populations (shown in Appendix 1c.2).
- Examining the relationship between the sets of estimates: The overall relationship between the GBD input data and Disbayes output estimates was plotted to identify patterns of conformity and deviation between the two. Due to the 'black box' nature of Disbayes, it is particularly important to consider if its outputs are reasonable approximations of the inputs. The relationship was quantitatively examined by regressing Disbayes estimates with their paired GBD estimate (by age, sex, IMD quintile, disease and measure (incidence or prevalence)). Outputs and normalised root-mean-square error (RMSE) were examined (with *normalised RMSE = RMSE/ mean dependent variable value* where the GBD estimate is the dependent variable). As case fatality has no matched input value, it was not examined in this way.
- Examining outliers: To identify where relationships between the two sets of estimates were weaker, residuals of Disbayes estimates were examined. Outliers

were defined as outside the 2.576 Standard Deviations (SD) from the mean (approximating to the 99% central spread of the residuals). The data underlying these outliers were then examined to identify drivers of their wider differences.

Results

Examination by deprivation

Full results are graphed in Appendix 1c.3 from age of 50 years (below which values are generally very small). These broadly show that the most deprived quintile of areas had the highest rates of most diseases for each of the three metrics, with many cases where quintiles cross through the age range or where they are clearly out of order – for example with the second least deprived quintile having lower rates than the least deprived. T2DM incidence and prevalence for both females and males provide a clear example of this. Lifetime prevalence for breast cancer and colorectal cancer in females is higher in *less* deprived groups.

There are also examples of trends for one or two quintiles deviating from the trend of the rest, for example, patterns for T2DM for males offers a few questions. As may be expected, case fatality rates for all quintiles increases with age, and faster for more deprived quintiles, with prevalence gently falling throughout these increases to incidence. The question is why despite similar case fatality and prevalence rates, quintile 5 solves to a very high incidence in the oldest ages while quintile 4 solves to a very low value. The underlying GBD data are not enlightening to this, as they are low values at these ages for all quintiles, as shown in Appendix 1c.4. Ultimately, the consistency requirement is likely to be driving these

differences, while the actual impact on results is likely to be extremely small due to the very low populations of people in their 90s affected by these issues.

Figure 4.2a/b shows the percentage difference in incidence, prevalence and case fatality rates (y-axis) across quintiles of deprivation (x-axis). For example, the population-weighted average case fatality rate of oesophageal cancer in females was approximately 15% higher in IMD quintile 4 than 5, and males' incidence approximately 40% lower in IMD quintile 2 than 5. There is broad between-disease variation in patterns across the quintiles and by sex. For example, colorectal cancer prevalence is highest in the middle quintiles of deprivation for males but decrease with deprivation for females, while AF rates are broadly flat by deprivation in males and increase for females. Case fatality rates showed somewhat less consistent patterns between diseases than incidence or prevalence, especially for males. For both males and females, the least deprived quintile (1) had higher case fatality rates for some of the diseases than the second least deprived.

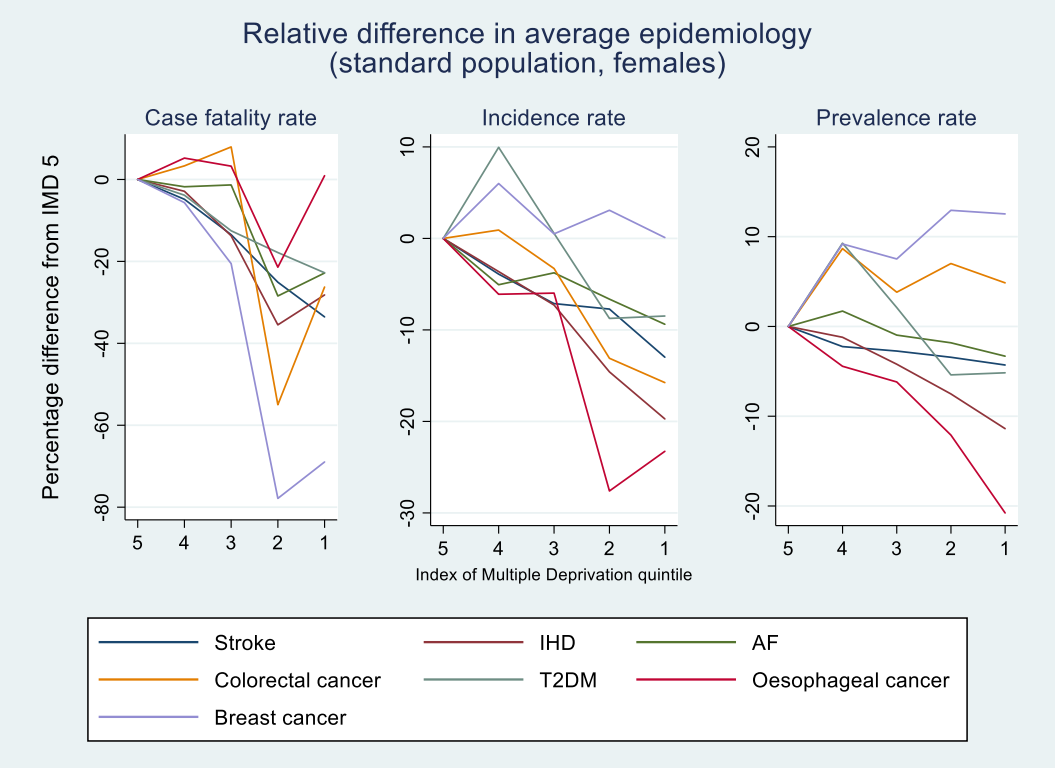


Figure 4.2a: Proportional differences in Disbayes outputs across IMD quintiles (females).

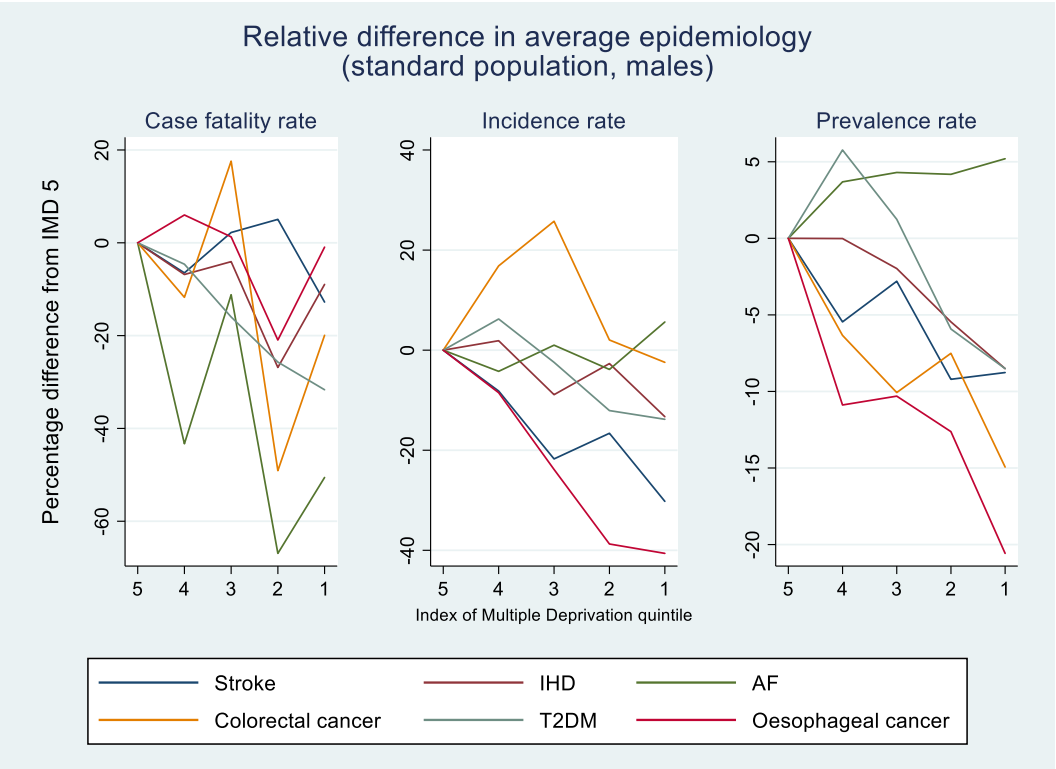


Figure 4.2b: Proportional differences in Disbayes outputs across IMD quintiles (males).

Relationship between sets of estimates

Regressing the Disbayes output estimates against the GBD input estimates across all diseases, all ages, for incidence and prevalence for females and males, we see a close relationship, shown in table 4.1. The coefficient for a one-unit increase in GBD estimate is an increase of 0.983 in the Disbayes estimate value (SE 0.00059). RMSE was 0.00467 and mean dependent variable value of 0.0283, giving a normalised RMSE of 0.165 (16.5%).

Plotting these data in figure 4.3 is done by regressing the pairs of each estimate by age, sex, IMD quintile, disease and epidemiological parameter between. For example, the incidence of stroke for females aged 58 in IMD quintile 2 from the GBD is plotted against that from the Disbayes estimates. The vast majority closely conform between the datasets, with a small number diverging from trend around GBD estimate values of approximately 0.20-0.25, progressively deviating from predicted values. On closer examination, these are incidence rates for IHD for males, and to a lesser extent females, across all IMD quintiles, which are further explored below.

The greatest difference for case fatalities was for breast cancer, nearly 80% lower in IMD quintile 2 than 5. High quintile-to-quintile volatility was observed for many diseases, while one would generally assume 'natural' data to form a grossly bell-shaped distribution, with the greatest differences between quintiles 2 and 3, and 3 and 4.

Table 4.1: Outputs from regression model of Disbayes outputs against GBD inputs.

Output	**Value**
Coefficient	0.983
(SE)	(0.00059)
Constant	0.000945

(SE)	(4.11 x10^{-15})
Observations	15,000
Root Mean Squared Error	0.00466
Adjusted R^2	0.994

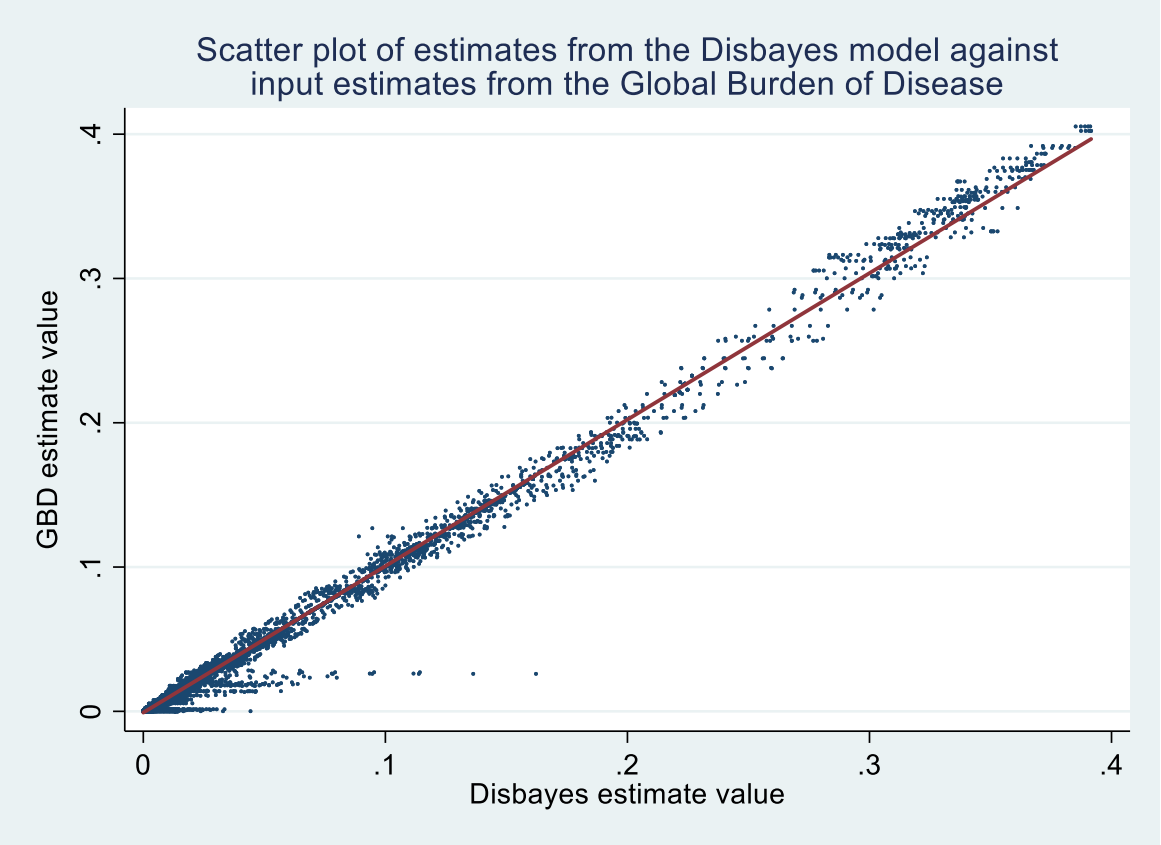


Figure 4.3: Scatter plot of Disbayes outputs against matched GBD values.

Examination of outliers

On regressing Disbayes estimates on GBD estimates at the age-sex-IMD quintile level, the SD of Disbayes residuals was 0.00470 (16.2% of the mean prediction). The 50th centile of absolute residuals represented 3.34% of the mean, 90th centile 15.4% of the mean, 95th centile 29.1% of the mean and 99th centile, 70.0%. Observations with crude residuals lying outside 2.576 SD of the mean made up 2.76% of all dependent variable values. Of residuals

with GBD estimate greater than the Disbayes estimate, observations were mostly for the prevalence of T2DM, for both sexes, with smaller numbers of observations of IHD prevalence for males. For residuals with Disbayes estimate greater than GBD estimate, the most represented metrics were for IHD prevalence and incidence in males, IHD incidence in females, and T2DM prevalence in females and males, alongside smaller numbers of stroke, oesophageal cancer and AF observations. These observations can be seen branching from the regression line at the lower end of the distribution in figure 4.3. To examine where these differences arose, the GBD and Disbayes estimates of incidence and prevalence rates for males and females for IHD and T2DM are graphed in Appendix 1c.4. This shows that the fitting of Disbayes estimates to the GBD data is resulting in under- and over- projecting of the trend compared with the GBD data, likely as a result of the need to create consistency between the three parameters.

Discussion

Summary of findings

Overview

This work estimated consistent sets of incidence, prevalence and case fatality across five quintiles of deprivation for seven non-communicable diseases: IHD, T2DM, stroke, AF, oesophageal cancer and colorectal cancer, plus breast cancer for females only. This used the Disbayes model. Previous work(47) was used for hypertensive heart disease rates at the

England level, as local incidence data were not available. Rates for asthma and musculoskeletal diseases (low back pain and osteoarthritis (OA) of the hip and knee) were taken directly from the GBD study(2) at the upper tier local authority level and aggregated to IMD quintile level.

Strengths and limitations

Strengths

This work uses local-level data aggregated to IMD quintile level to estimate the differences in disease epidemiology by deprivation. There is little similar work and nothing directly comparable. This has allowed disease epidemiology to vary by local area while working within the confines of what is feasible given the method's runtime. The Disbayes model has the strength of using the same three disease states as PRIMEtime's PMSL, allowing its outputs to be easily used. It is a well-established approach using a fully Bayesian method.

The input data used are considered gold-standard, coming from a well-funded ongoing project from the World Health Organisation (WHO) that has a very good reputation. The team of epidemiologists that worked on the local data in England come from among England's best known university epidemiology departments and government agencies.(2,96,97) The consistency and international acceptance of the GBD method are important strengths here, and despite some limitations mentioned below, the overall balance of trust in these estimates is high.

This work has also highlighted how important it is to understand local context – exactly how a disease's burden varies in ways other than simply headline deprivation should be better understood, and then what those driving forces might be once the patterning is well

understood. For example, why is T2DM incidence highest in the fourth quintile and not first or third is an interesting example – are people in this quintile more likely to have sedentary jobs than quintile 5, or higher incomes resulting in eating more takeaway food, or with less health education and time to prepare healthy food than quintile 3, or cultural or infrastructural factors related to the types of areas that by chance are in quintile 4, or combinations of all these explanations and many more.

Limitations

The Disbayes model can be deemed a black box, making it difficult to know how the precise sets of estimates were arrived at, only leaving comparison with input data to judge appropriateness of the results (as has been done here). This means that in places there are solutions provided that are dissimilar from the input values that must then be considered in terms of their impact on modelled outcomes, despite not having an alternative consistent set of estimates for that population. Overall, however, the very high concordance between the datasets that was described above ($\beta = 0.983$, normalised RMSE 16.5%) is reassuring.

The GBD data themselves have been critiqued(96,110) for multiple issues such as the use of poor-quality input data in some countries, extensive missing disease data and even lacking population sizes in some places. Once again, the internal consistency requirement adds complexity by moving the GBD estimates further from their raw input data. The England local authority data have additional specific issues, such as a lack of transparency about exactly how they were synthesised from their crude data inputs and gaps in local-specific data leading to the use of modelled estimates.(111)

It is difficult to define outliers of these estimates. Many of the residuals are very close to or equal to zero, so using a standard marker of 1.5 times interquartile range results in a large proportion of the values being categorised as outliers (indeed, any measure of spread of these data will be very narrow). Instead, a stricter definition of the crude value of the residual being greater than 2.576 SD from the mean was used (prototypically equating to the 99% central spread of a normally-distributed variable). However, this fails to account for the proportional difference between input and output values. Measuring outliers by a proportional measure of difference is not helpful as many of the values of both datasets are zero (and these often do not match) and proportions cannot be calculated. Where a value in one dataset is very near zero, small absolute differences imply very large proportional differences, while homoskedasticity results in residuals of approximately the same order of magnitude along the range of the independent variable – so the selection of outliers would be driven by the smallest values, not by weak relationship between values. It was therefore accepted that it would not be helpful to calculate proportional differences and crude residual differences were used.

It is unfortunate that within-quintile heterogeneity could not be estimated, as mentioned, due to runtime of the Disbayes model. This is an area for further work, as noted.

Uncertainty around these sets of estimates was not pursued. It is not a straightforward thing to incorporate uncertainty on epidemiology into modelling process, due again to the consistency requirement. To allow one parameter to vary while allowing the other two to cope – both mathematically and conceptually – is challenging. For example, if prevalence is allowed to start at a lower value than its central estimate then any risk of it becoming negative needs to be coped with, while we might also ask if it alone should be lower and not

having a correction to lower incidence or increase case fatality rates to explain how that prevalence came to exist in that simulated population. Alternatively, if incidence is allowed to vary, then without the numbers of deaths mirroring that variation (via case fatality) there is a risk of prevalence rates becoming unrepresentative of any real-world population. If the input or outputs values of the model are not consistent with a reasonable estimate of the real world (given any necessary or helpful simplifying assumptions) then the model should not be emulating that situation and results would be rendered uninterpretable. Further work is necessary to understand what approaches are possible to incorporating uncertainty on these estimates, alongside their effects on modelled point estimates and cumulative uncertainty.

Conclusion

This work used high-quality publicly available data to re-estimate consistent sets of epidemiology using a Bayesian fitting method. It was not feasible to estimate sets of epidemiological data for each local area separately, so it was done at the level of the IMD quintile. It found that burden of diseases tended to increase with deprivation, though not always monotonically or smoothly, implying that other dynamics are also important in predicting disease epidemiology other than headline geographic deprivation. The unexpected nature of this patterning of disease burden underlines the importance of using as locally-specific data as possible.

Chapter 5: Estimating adult BMI by age, sex and local authority

Introduction

For modelling to accurately reflect local areas, local risk needs to be captured, in the form of BMI distributions. PRIMEtime has previously parameterised BMI assuming a lognormal distribution separately for age group (in 5-year bands) and sex for the whole population of England. This is an appropriate approach that will be followed here, while also estimating separate BMI distributions by local area. This Chapter describes the methods for estimating local-level adult BMI and Chapter 6 describes the methods for local Child BMI values.

Survey data exist for estimating risk factor prevalence in England and the UK from studies such as the HSE(112), National Diet and Nutrition Survey Rolling Programme (NDNS)(113) and Active Lives Survey (ALS),(114) but each of these has limitations that prevent them from giving precise estimates of risk factor prevalence by age-, sex- and local authority area-level. PHE provides estimates of the percentage of adults overweight or obese in each local authority area through the Public Health Outcomes Framework (based on ALS data), but these are not broken down by age or sex.(6) All these surveys suffer with the common problems for small area estimates(115) that they do not specify the respondent's geographic area and their sample sizes are too small to deliver precise subgroup estimates for risk factors. To overcome these problems with the survey data, methods have been developed to allow researchers to estimate the prevalence of risk factors for small

geographical areas. A wide variety of small area estimation methods have been developed but they can be grouped into three broad approaches: indirect standardisation, model-based estimation and microsimulation approaches.(116–119) Each takes a different approach to using the associations between the characteristic in question (risk factor, health behaviour) and socio-demographic variables in survey data to estimate the prevalence of the characteristic in an area based on the pattern of those same socio-demographic markers in that local population.

Summary of small area estimation methods

The indirect standardisation approach takes survey data on a variable of interest and weights the population based on the composition of local populations. For example, HSE could be used to estimate the prevalence of obesity in subgroups defined by age group (17 levels), sex (2 levels) and ethnicity (5 levels), giving a total of 170 subgroups. The percentage of people with BMI >30 in each of the 170 groups is first drawn off, then a weighted average proportion calculated based on the proportions of each of the 170 subgroups in a local area, for example from the Census.(120) These methods are intuitive and fairly easy to apply but amount to demographically-weighted means without accounting for unmeasured variation around those means between areas, for example related to more granular socio-demographics, local context and culture.(121)

Model-based estimation approaches represent an improvement on this method by using regression modelling (usually logistic regression) to estimate the probability of the risk factor in the survey dataset depending on the basis of either individual-level features, area level features, or both.(116,117) Using regression modelling allows a greater variety of

socio-demographic variables to be incorporated and for their contributions to be assessed (removing those that fail to contribute), but as the population data still is usually still in the form of population counts, it is still limited in practice by the number of cross-classified variables available in the population data. For example, only 3 cross-classified variables are published in the UK Census at the postcode area level for health-related variables (namely age group, sex and marital status).(122) Charlton(123) expanded this approach in 1998 by using the Census microdata that had been made available for the first time in the 1991 Census. The Census microdata 1991 comprised 2% of all individual lines of the Census responses, stratified at the output area level (postcode areas of around 125 households).(124,125) A multilevel logistic model was built treating the individual as one level and the individual's GP practice as another. The coefficients produced were used to estimate the probability of people consulting their GP for a serious illness by drawing practice-level variables applied to Census areas (local authority groups) using values from other datasets (eg. estimates for the proportion of the population moving house each year as a proxy for annual practice population turnover). This approach greatly expands the number of variables available for use but may still be limited by the need to exactly match the survey variables with equivalents in the Census.(117)

An addition to these methods was developed by Twigg *et al* (2000)(122) to overcome the difficulty in accounting for local-specific variation in demographic-based estimates. This approach was later used by the same team(126) to estimate the prevalence of overweight and obesity by age, sex and ethnicity in each local authority in England. This used postcode-sized sampling unit identifiers in the HSE to apply area-level variables in a multilevel logistic model, capturing more between-group local variation. They assume an equivalence between HSE Primary Sampling Units and Census output areas (available in the cross-

tabulations but not microdata). Again, these coefficients are applied to Census subgroup counts, but in combination with local area-level variables, applied to wards in the Census to estimate the prevalence of overweight and obesity in each local authority. The main drawback to this approach is the extensive data requirement.

Microsimulation approaches combine spatially aggregated survey data with geographically disaggregated microdata. Spatially representative microdata also often needs simulating as it is not available for all of the required geographic areas. The local population can be represented either by creating a line of microdata for each individual in the area from a survey, or by weighting survey data to reflect local populations with appropriate data such as Census data.(115) These methods have an advantage of capturing local variation and can be used to estimate distributions of characteristics (rather than the proportions of categorical characteristics as with the standardisation and synthetic estimation methods), but estimating credible intervals and validation are challenging.(127,128) For example, a spatially granular microsimulation of the city of Leeds, UK,(129) used a "combinatorial optimisation" method that involves combining locally-collected childhood obesity data with the more spatially granular social information of the Census 2001 to estimate features of the obesogenic environment for each LSOA.(130)

Summary of available datasets

There are three main datasets to consider for drawing BMI data from: the HSE,(112) NDNS(113) and ALS.(114) The Census of England and Wales 2011 is the only real option for sourcing good local demographic data, but it comes in two formats: the standard release(120) and Census microdata.(125)

The Census

The Census is a survey conducted every ten years in the UK (with one for Scotland and one for England and Wales), last performed in 2011. This requires every resident to submit responses to a wide variety of questions on a single day (in 2011 it was 27 March). The standard data release publishes responses as tables of counts, giving combinations of different attributes in a given geography (postcode-level, local authority, county, region). For instance, a user can access tables giving the exact counts of respondents living in Hackney who answered that they were male, aged 30-35 years and of Asian ethnicity. The number of attributes that can be cross-tabulated varies by which are selected, but is generally 3 or 4. The Census microdata 2011 are Samples of Anonymised Records representing the individual-level responses given by 5% of individuals without any level of aggregation, stratified by output area (postcode areas). It has a good level of representativeness across age, sex, marital status, ethnicity, receipt of unpaid care, health status, occupational classification, hours worked, main language spoken, qualifications obtained, and religion. This data allows subgroups to be disaggregated using many more than 3-4 variables at a time. This comes with the trade-offs that they are an incomplete sample, so may lack power to make accurate estimates of subgroup characteristics, and that the total number of available variables is lower than the standard release. The lowest level of geography reported is the grouped local authority Census area. This is usually a lower tier local authority, or if that area has a population of less than 120,000 persons then it is grouped with an adjacent area, consistent with disclosure control principles (ie. avoiding the risk of an individual's data being identifiable).(124)

Health Survey for England

The HSE is a cross-sectional annual survey of health status in England, using a clustered multistage stratified probability sampling design. The design aims for equal probability of household selection and therefore a representative sample of England. A random sample of PSUs based on postcode sectors was first drawn, then random sampling selected households within each PSU. A minimum number of PSUs was selected from each NUTS-1 region (National Units of Territorial Statistics level 1, the former Government Office Regions, of which there are nine in England). Sampling for the HSE in the most recent available wave (2017) identified 9,612 households in 534 PSUs. Participants are invited to interview by letter evenly across the year, then asked if they would be willing to have a nurse visit (in March 2018) to take certain measurements. A total of 7,997 adults and 1,985 children were interviewed in 2017 and nurse visits were conducted for 5,196 adults and 1,195 children. The survey covers questions over basic demographics (eg. on age, sex, ethnicity), social features (eg. on socioeconomic status (SES), education), health behaviours (eg. on smoking, fruit and vegetable consumption), physical health (met and unmet care needs, general health), mental health and wellbeing. The nurse visit takes measurements including blood pressure, height, weight, a urine sample and blood tests (eg. cholesterol, kidney function).(131,132) BMI in the HSE has a choice of variable, to include unreliable measurements or to use estimated weights if measurement is over 200kg, where the scales used are unreliable. There is also an option for using estimated weight >130kg.

National Diet and Nutrition Survey

The NDNS is an annual survey focusing on dietary information and nutritional status, performed on adults and children resident in private households aged 18 months and over across the UK. The first stage involves a visit from an interviewer including a 4-day food

composition diary, PA questionnaire, questions on other health-related behaviours (eg. smoking and alcohol), demographic and social variables and information from the "Main Food Provider" of the household (on shopping and cooking habits). Height, weight and a urine sample are taken. The second stage requires a nurse visit to take additional measurements (such as blood pressure and a blood sample). The survey aims to collect a representative sample of the UK population of 500 children (1.5-18 years) and 500 adults (19 years and over), plus boost samples from Wales and Northern Ireland taking samples up to 200 participants each. The sample is drawn from postcode sectors clustered into PSUs, with an equal chance of selection design at the address level. In the most recent data release (covering waves 7 and 8, collected in 2014-16), a total of 8,848 addresses were selected from 316 PSUs (28 addresses selected per PSU). To achieve the target numbers of adults and children, addresses were randomly allocated for the survey to be applied to adults and children (10 addresses per PSU) or only children (18 per PSU). In this release, 1417 adults and 1306 children submitted full interviews, with 50% adults and 25% children giving blood samples. The published BMI variables are calculated from measured heights and weights only.(133,134)

Active Lives Survey

The ALS measures PA and related features in England. There are two steps of data collection for each participant, the first of which includes 'core' questions required to produce PA-related measures, demographics, height, weight and some basic dietary information. The second phase also asks about topics such as volunteering, environments and some more detailed questions on demographics. A slightly modified 'young person questionnaire' is provided for people aged 14-15 years. Sampling takes an unstratified design with equal probability of selection at the address level across England. Survey weights are provided to

help correct for response rates across local authorities, by response month and for the proportion of young people. Four reminders are sent to complete the online questionnaire, the third of which also includes a paper questionnaire to allow people to use that option. In the latest available year (2017), a total of 845,230 addresses were invited to respond and 201,579 valid responses received, 52% online, 48% on paper. The BMI variable is derived from self-reported height and weight responses (not verified at a nurse visit) and structured as a categorical variable for people with BMI in the ranges for underweight, normal weight, overweight, obese and morbidly obese.(135,136)

Aims

This chapter will describe the process for estimating and validating local authority level BMI distributions. Comparable estimates have not been produced since 2007,(126) So are now out of date and they only describe the proportions of the gross population as either overweight or obese, rather than the mean and SD BMI for each age and sex group that is needed for the PRIMEtime model structure. Therefore, the aim is to produce and validate new estimates of the distribution of BMI for each age and sex, in each local authority area of England.

Methods

Overview

BMI is best represented in the PRIMEtime structure as a mean and SD on a lognormal distribution. To do this, an approach similar to that put forward by Charlton (1998)(123) was developed, allowing the mean and distribution of BMI to be modelled by age, sex and local area.

This approach involved building a regression model to predict BMI from social and demographic variables, using health survey data (HSE 2018). This model was then used to impute a BMI for every member of a representative local authority population survey (Census of England and Wales 2011) that contained the same social and demographic variables. This allowed the population to then be aggregated into the age-sex-area subgroups required and BMI distributions to be drawn off.

Identification of candidate independent variables

Appropriate candidates for independent variables in the HSE 2018 (the most recent available wave) were mapped against equivalent variables in the Census microdata. These were adapted where necessary, for example recoding the ethnicity variable with 11 levels in the Census to be equivalent to the five-level variable in HSE 2018. Candidate variables were demographic and social markers thought to be associated with BMI that were available in the same or a translatable form in the Census microdata 2011. The table in Appendix 1d.1

shows the availability and alignment of variables in the HSE and Census microdata that were included in the final model, with the adaptations made to make them equivalent.

Candidate variables where alignment between the HSE and the Census microdata was possible were: age, sex, ethnicity (white, Asian, black, mixed race, other), self-reported health (from 1 (good) to 5 (poor)), the number of cars in the household, SES, NUTS-1 region, tenure categories (ie. living arrangements, such as renters or mortgage-holders), and dummy variables for being unemployed, being a student, having no formal qualifications, and having a degree. IMD quintile is stated in the HSE at the individual level, but is not included in the Census microdata, so was merged at the local authority level (the lowest available level of geography). Similarly, ONS Rural-Urban Classification (RUC) could be merged into the Census data at the local authority level. RUC is a taxonomy produced by the ONS to categorise local authority areas by the pattern of their population's geographic spread, from 1 (most rural) to 6 (most urban).(137)

Subgroup sample sizes

It is possible that subgroups in surveys may lack adequate sample size to make an accurate estimate of a variable mean and it has been suggested that approximately 30 entries are needed for each subgroup to minimise this risk.(138) The HSE 2018 has 35 entries in its smallest age/sex group and the Census microdata has over 5,000. However, the numbers of cases in the Census microdata stratified by age, sex and local area get much smaller in places: of the 7,968 subgroups (cutting the data by 249 area groups, 16 age groups and two sexes), 242 had populations less than 30, covering a population of 4,927 of the approximately 2.2m records (0.23% of records). All of these groups were in the highest two age groups (85-89 years, and 90 years and over).

Model design

A Generalised Linear Model (GLM) was used to cope with the skewed BMI dependent variable. GLM regression is a generalisation of the Ordinary Least Squares models, that relax Gauss-Markov assumptions of linear dependent-independent variable relationships, homoskedasticity and normally-distributed error. A variance function (family) must be specified, representing the distribution of variance of the dependent variable along the distribution of the independent variable(s). The link function acts as the translation of the dependent variable to allow a linear interpretation of the independent-dependent variable relationships, and the variance family to the gaussian. Each variance family has a standard (canonical) link function, though given there is heterogeneity within variance families, there is choice of link functions for some families.

Cluster-robust SE were required to account for the HSE's clustered stratified sampling design, at the postcode unit level. This meant that the formal tests for GLM family and link, such as the Park test, could not be used as they rely on squared error so are not valid with clustered sampling. The variable distribution was a skewed bell-shaped distribution, potentially consistent with gaussian or gamma families. Therefore, Cullen-Frey and QQ plots were examined (shown in Appendix 1d.2), which supported the choice of a gaussian family model. A log link was chosen as the pairing for this family, as has become an established approach for log-normally distributed independent variables.(139,140)

A model was built using a stepwise approach with BMI as the dependent variable and the candidate demographic and social variables tested as independent variables. First, the relationships between BMI and age, age^2, sex, ethnicity (and their interactions) were explored. The associations between the remaining candidate variables and BMI were then

explored, accounting for age, age^2, sex and ethnicity. Interactions were explored between each variable in turn with age, age^2, sex and ethnicity, but not between other independent variables. One variable at a time was added in turn to the model and the Bayesian Information Criterion (BIC) was used to assess if the added variable improved the model, using an arbitrary cut-off of -2 to indicate model preference. If the BIC supported that the variable improved the model, then it was included and subsequent variables tested with that variable included, and otherwise the variable was not included. Any individual with a missing value for a candidate variable was dropped from the dataset to test each model against the same data. BIC was used rather than other model diagnostic tests on the basis of preferring more parsimonious models (over the Akaike Information Criterion), reducing the risk of overfitting, especially around the hazard of multiple testing in this case. Likelihood ratio tests could not be used as tests of joint significance as they are invalid where robust SE are used.(141)

Categorical variables (eg. ethnicity) were assessed by adding each level of the categorical in turn as a dummy variable and included where BIC supported that it improved the model. Ordinal variables (eg. health) were first tested together, treated as continuous before treating it as categorical to test if the levels contributed to fit beyond linearity. Age was only treated as continuous, implemented by recoding the categories to their mid-point values. The top category of “90 years and over” was recoded as the mean age of that age group in England, based on the most up-to-date mid-year estimates from the ONS (2018).(142) These estimated the mean age of women in this group to be 93.0 years and men 92.4 years.

Using BIC means that the order variables are tested in can change which ones are included. This was dealt with at the end of the process by checking if the Wald tests in the model

output did not support a significant association (at p <= 0.05) then re-testing these variables by BIC to confirm they still improved the model, aiming for a parsimonious model and reducing the risk of redundant variables, overfitting or collinearity.

Imputation of BMIs

The resulting model was then used to impute a BMI for each respondent aged 16 and over in the Census microdata. Outputs of GLM log-link models are provided as log values, interpreted multiplicatively as:

$$E(Y) = \prod_{i=0}^{n} \exp\left(\beta_i X_i\right)$$

Where E(Y) is the estimate of the independent variable (BMI), X_i are independent variable values (and constant $X_0 = 1$) and β_i are output coefficients. This is done by multiplying variable value with its coefficient, exponentiating, then multiplying these together to give an imputed BMI value.

Each coefficient's values had random variation added at the individual level to reflect population heterogeneity. Random variation was produced on a normal distribution, with SD equal to the SE of the GLM coefficient value. The figure in Appendix 1d.3 demonstrates BMI distributions with and without random variation.

Calibration

Calibration of the results was required to allow the spread of values around the means in the modelled results to reflect that in the measured BMI data. For this, the SD of log-BMI was calculated in HSE for each age-sex group. These were then smoothed using linear regression against age and sex, to produce predicted SD values with out year-on-year

volatility. The uncalibrated individual estimates of log-BMI were converted to a z-score (value minus group mean, divided by SD). This was then multiplied by the smoothed modelled SD from HSE, then the values of the original group mean added back on. This was then exponentiated from log-BMI to the BMI scale.

To do this, the distribution of estimated log-BMI was calibrated to approximate the distribution of log-BMI from the HSE. The process was to subtract the group mean (age and sex group) from each observation then divide through by the group SD, to produce a z-scored (standard normal) distribution. Then the observations were multiplied by the SD of the measured BMI distribution from HSE before the estimated group mean was finally added back on. This calibration ensured a high degree of reflection of estimated BMI values with the distributions from the HSE. Estimates of BMI below 15 or above 50 were treated as non-valid and dropped.

Estimates were parameterised using a mean and SD of BMI for each sex in each age category in each local authority area (except the City of London and Isles of Scilly, due to their small populations). Where local authorities were merged in the microdata as grouped Census areas, the same age- and sex-specific parameters were used for all local authorities included in that group.

Validation steps

Four validation steps were used, based on previously laid out approaches.(128,143)

1. Face validity is the appropriateness of the modelling process to its aims. This was assessed by examining the variable coefficients and assessing whether there was a

clear rationale for why that variable should correlate with BMI including if other studies had found an association between BMI and a comparable variable.

2. Estimates were assessed for heteroskedasticity. This is the non-constant variance of residuals along the distribution of the dependent variable that is a sign of poorly fitted models. The presence of heteroskedasticity can imply issues such as the omission of an important independent variable.(143) Heteroskedasticity was examined for by calculating the residual between each individual's measured BMI in the HSE 2018 data and the predicted BMI from the model. These residuals were then plotted along the distribution of estimated BMI, separately for males and females, to examine if variance was approximately constant. Heteroskedasticity could not be formally tested for (for example using the Breusch-Pagan/ Cook-Weisberg test or White's/ Cameron and Trivedi decomposition Information Matrix test) as these tests are not valid with robust SE.(141,144,145)

3. Internal validity is the consistency of model results with the data they were based on, in this case the HSE 2018. It was assessed from the perspective of construct validity (that the model measures what it intends to) by region (NUTS-1), sex, then age and sex.

a. There is no marker of the local authority in the HSE, so it is not possible to compare internal validity at the local level. The lowest level of geography is the NUTS-1 region (of which there are nine in England), so mean estimated and measured BMI (standardising all age-sex groups to be represented equally) were plotted against region for comparison.
b. Construct validity was assessed at the sex level by comparing age-standardised histograms of estimated BMI versus measured BMI in the HSE 2018 data (again representing each age group equally).
c. At the age-sex group level, this was examined by plotting mean estimated BMI against mean measured BMI (in the HSE 2018). This relationship was also quantified using linear regression. Here, the important feature is that the groups' means are *unbiased* overall relative to the base HSE data, ie. they do not systematically under- or over-represent measured values.

4. External (or convergent) validity is the extent to which results reflect data collected entirely independently from the data the model was based on. This was assessed by two approaches:

    a. The first step was a comparable approach to that described under 3. The age-standardised BMI distribution, for each sex, across England was compared with the equivalent distribution from the NDNS 2014-16 (using NDNS as the ALS reports BMI only, the percentage of respondents in each BMI group, not a BMI value). These histograms are again standardised to represent each age

group equally. This relationship between estimated and measured age and sex group means was then quantified by linear regression.

b. The second external validation step was to compare the percent of people in each local authority estimated to have high BMI (ie. those in the overweight, obese and very obese categories) with those self-reported in the ALS 2019-20 to have high BMI.

c. Local authorities in the top decile for overweight rates and obesity rates were compared with previous estimates of these from Moon *et al* (2007).(126)

Results

Missing data

As shown in table 5.1, there were small amounts of missing data for these variables in the HSE 2018 with the exception of the BMI variable, with 18.33% missing. The Census microdata also had very little missing data. Appendix 1d.4 shows the break-down of missing data for the BMI variable by age, sex and IMD quintile, showing little difference by sex or IMD. The percent of missing data across the age range was fairly consistent at approximately 15-20% up to age 72.5 years, after which it rose continually with age up to 46% for the over 90 years group.

Table 5.1: Percent of missing data in source datasets (to 2 decimal places, d.p.), by variable (England only, respondents aged 17 and over). *Index of Multiple Deprivation(146) was merged into Census data at the lower tier local authority level from ONS data. SES = Socioeconomic Status.

Dataset: Census microdata	**HSE 2018**

Variable	% missing	Variable	% missing
Age	0.00	Age	0.00
Sex	0.00	Sex	0.00
Ethnicity	1.28	Ethnicity	0.39
Self-employed	1.28	Self-employed	4.90
Retired	1.28	Retired	0.00
Long-term sick	1.28	Long-term sick	0.00
No qualifications	1.28	No qualifications	0.12
Degree holder	1.28	Degree holder	0.46
Tenure status	2.21	Tenure status	0.23
Region	0.00	Region	0.00
Health	1.28	Health	0.04
SES	0.00	SES	4.52
IMD*	-	IMD	0.00
		BMI	18.33

Model results

The final model produced from the HSE is summarised in Table 5.2. The results showed broadly that BMI increased with age then declined again from around the age of 65. As shown in Figure 5.1a and 5.1b, after standardising for age, females had lower modal BMI, but also a broader range of BMI across the population, with a higher proportion of people with normal or low weight than for men. Note that unless stated otherwise, *all* the histograms, kernel density plots and scatter plots in the Results and Validation sections are standardised to make each age group equally represented, so do not reflect real populations. This has been done to avoid the figures from being influenced by a preference for larger age groups over smaller ones. The mean BMI for females was 27.7kg/m^2 and males 27.9kg/m^2 (or 27.6kg/m^2 and 27.8kg/m^2 respectively, weighting all age groups

equally). A total of 3.4% of individual-level estimates were <15kg/m^2 or >50kg/m^2 and dropped.

Table 5.2 shows the coefficients for each of the variables fitted in the final model, with their 95% confidence intervals (these confidence intervals also represent the distributions used to create variation in each coefficient imputing BMIs, as described in Methods). These show BMI was most associated with age, sex, ethnicity and health status.

Table 5.2: Regression model outputs in terms of variables, their coefficients and coefficient 95% confidence intervals, given to 3 significant figures (* = interaction between stated variables). IMD = Index of Multiple Deprivation 2019, SES = socioeconomic status.

Variable	Coefficient (log scale)	95% confidence intervals Lower	Upper
Demographics			
Age (years)	0.0114	0.00932	0.0136
Age (years) squared	-9.81×10^{-05}	-0.000119	-7.70×10^{-5}
Female (dummy)	0.108	0.0288	0.188
Female*age	-0.00398	-0.00709	-0.000874
Female*age squared	3.22×10^{-5}	3.66×10^{-6}	6.08×10^{-5}
Ethnicity			
White/ mixed (base)			
Black	0.0483	0.0147	0.0819
Asian	-0.0219	-0.0435	-0.000323
Other	-0.0423	-0.0901	0.00538
Social/ economic			
Self-employed (dummy)	-0.0159	-0.0337	0.00200
Retired (dummy)	0.00450	-0.0154	0.0244
Long-term sick (dummy)	-0.0221	-0.0653	0.0210
No qualifications (dummy)	-0.00621	-0.0209	0.00844
Degree holder (dummy)	-0.0286	-0.0407	-0.0164
Health status			
Very good (base)			
Good	0.0574	0.0467	0.0682
Fair	0.108	0.09207	0.124
Bad	0.109	0.0794	0.138
Very bad	0.142	0.0838	0.200
Tenure status			
Own outright (base)			
On a mortgage	-0.00130	-0.0178	0.0152
Shared ownership	0.0275	-0.0525	0.107
Rental	0.0192	0.00304	0.0354
Living rent-free	0.0684	-0.000499	0.137
ONS Social Grade			
AB – managerial and professional (base)			
C1 – intermediate occupations	0.00200	-0.0177	0.0217
C2 – Small employers and own account workers	0.0141	-0.0113	0.0395

DE – lower supervisory, technical, and semi-routine occupations	-0.0116	-0.0280	0.00480
Other/ no response	-0.0123	-0.0366	0.0120
Area-level			
Area IMD			
Quintile 1 (base, lowest deprivation)			
Quintile 2	0.00732	-0.00890	0.0235
Quintile 3	0.0130	-0.00446	0.0305
Quintile 4	0.0176	0.000301	0.0349
Quintile 5	0.0315	0.0116	0.0514
Region			
North East (base)			
North West	-0.00122	-0.0249	0.0225
Yorks and Humber	0.0209	-0.00582	0.0477
East Midlands	0.00591	-0.0194	0.0312
West Midlands	0.0202	-0.00633	0.0467
East of England	0.0122	-0.0140	0.0384
London	-0.0289	-0.0540	-0.00373
South East	0.0162	-0.00975	0.0423
South West	-0.00181	-0.0259	0.0222
Constant	2.97	2.91	3.03

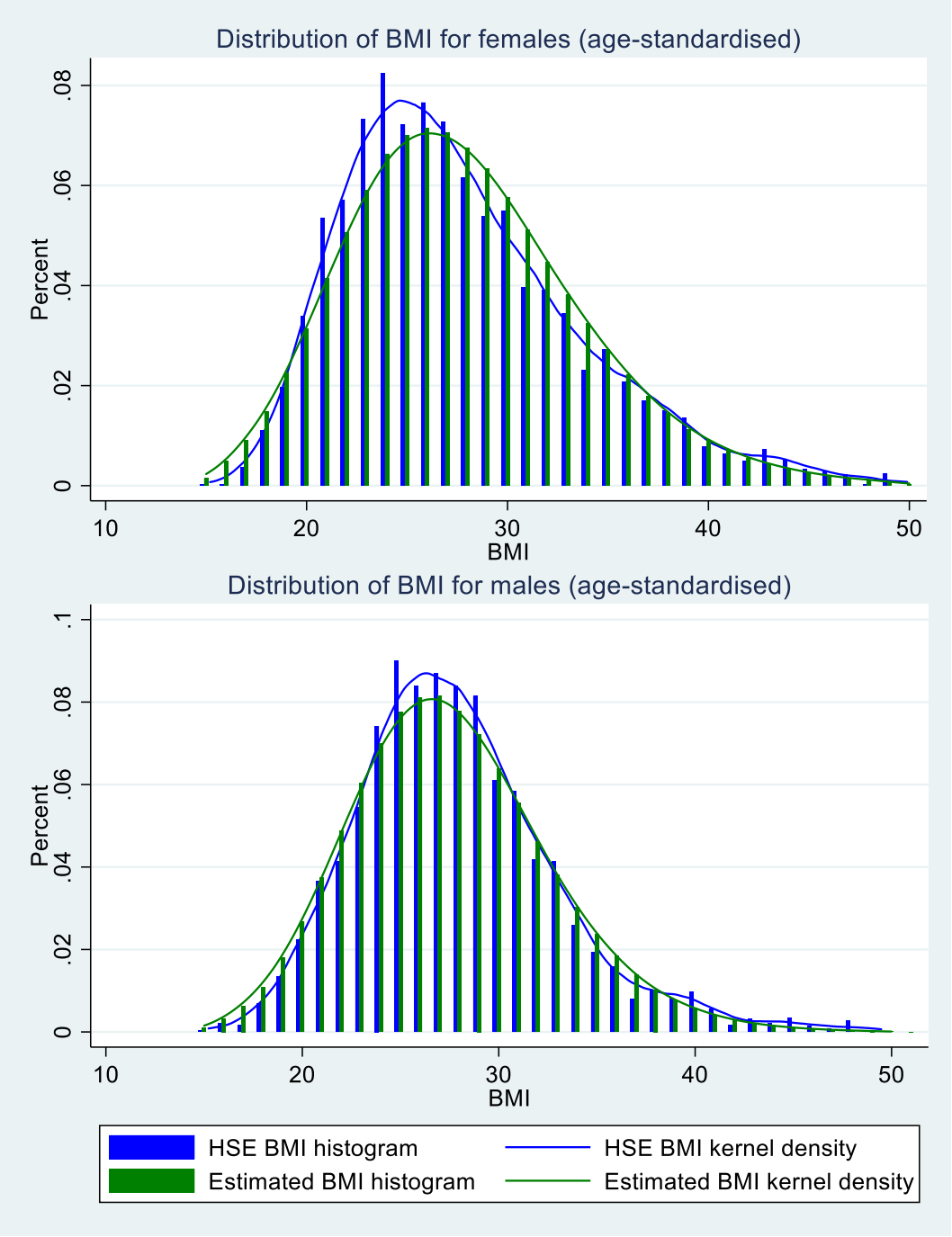


Figure 5.1a (above) and 5.1b (below): Histogram of age- and sex-standardised BMI estimates for England for females and males, respectively, including the equivalent histogram for the HSE 2018 BMI data for comparison.

Figure 5.2 shows the distributions of BMI for the whole populations of the NUTS-1 regions of England (unstandardised, representing populations of the Census year 2011). This shows small differences in mean BMI and distribution around those means. It shows that London

had the lowest overall distribution of BMI, followed by the South West. The West Midlands and Yorkshire and the Humber have the least favourable BMI distributions. Despite generally small differences between regions, these differences are cumulatively substantial; for example, the density of the population with a BMI of 30 in London is approximately half that of the highest-BMI regions. Additional local authority granularity is given in figure 5.3, mapping the differences in mean BMI between areas.

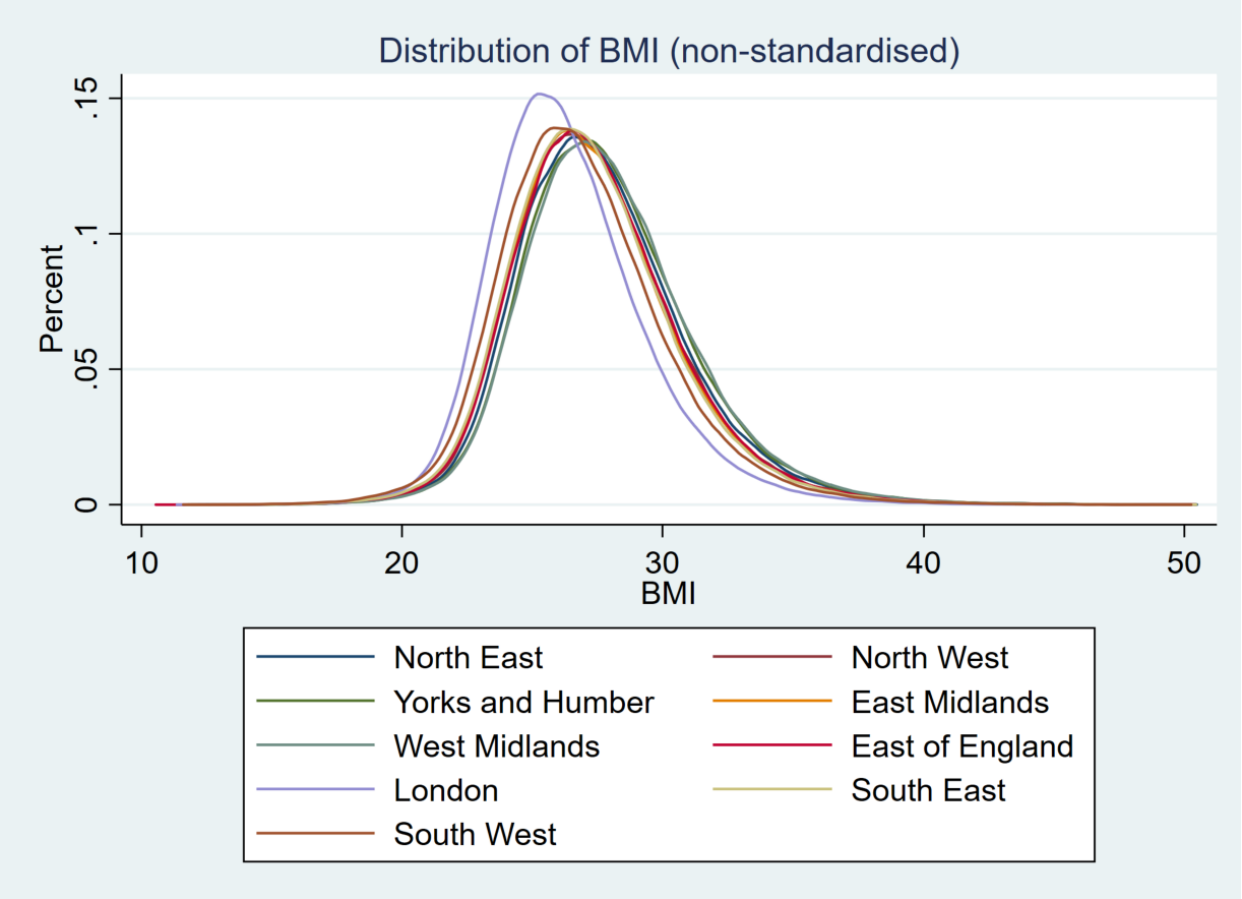


Figure 5.2: Distribution of BMI by NUTS-1 region, non-standardised, for both sexes combined.

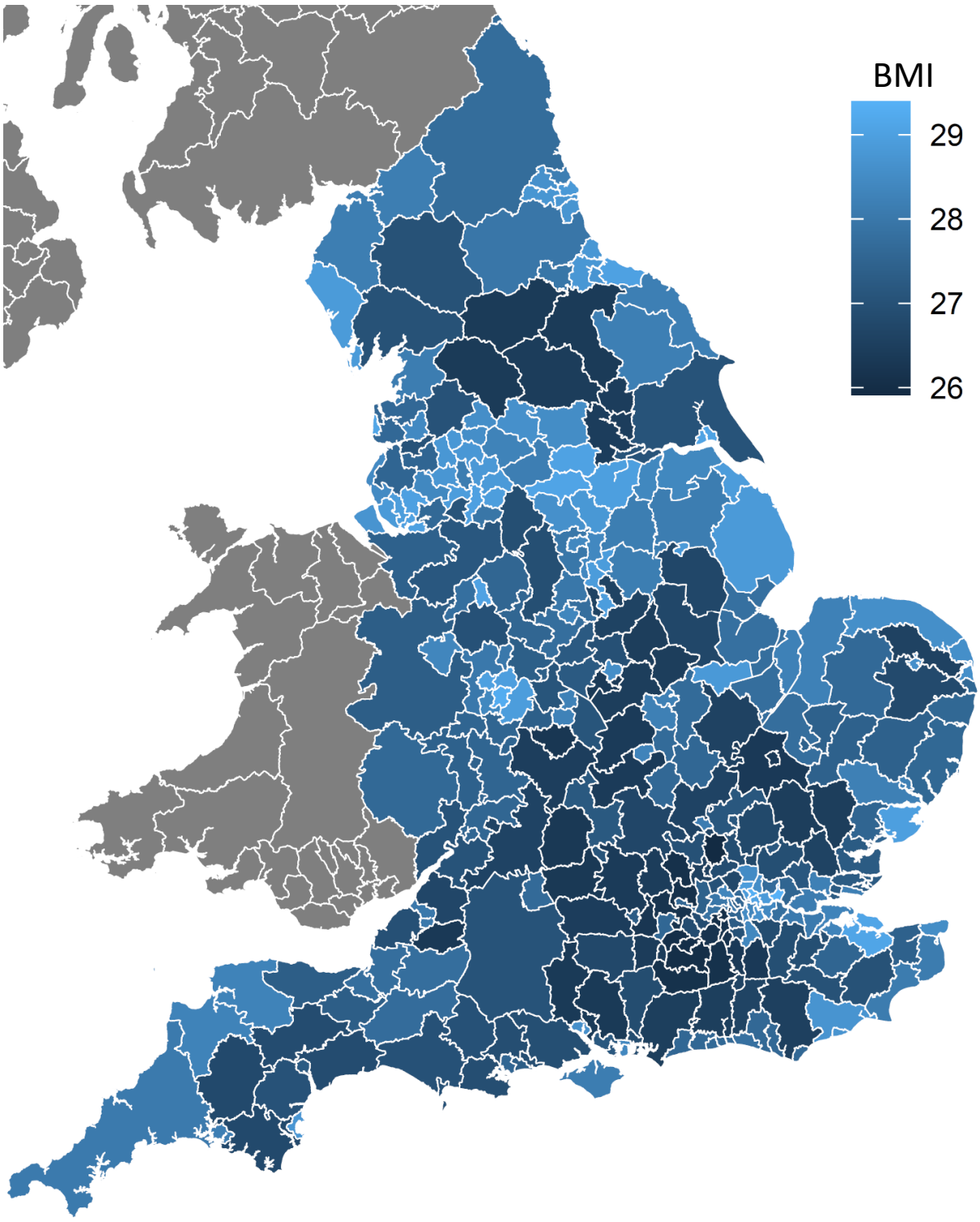


Figure 5.3: Heat map of mean BMI lower tier local authorities of England (kg/m^2) (unstandardised).

Figures 5.4 and 5.5 demonstrate the orders of magnitude of difference between BMI estimates between local authorities chosen from different quintiles of deprivation. First, Figures 5.4a and 5.4b shows the cumulative distributions of all age groups for females and males, respectively (non-standardised), with the whole distribution of BMI represented for each age group. Next, Figures 5.5a and 5.5b show distributions broken down by age group as well, for three more local areas. These show mean BMI for each age-sex group in each local authority, with the England average through the age range for reference, demonstrating the orders of magnitude of difference between local authorities across the age range. Differences between areas generally reduce over the age range and for a given area the volatility in mean BMI varies more for females than males.

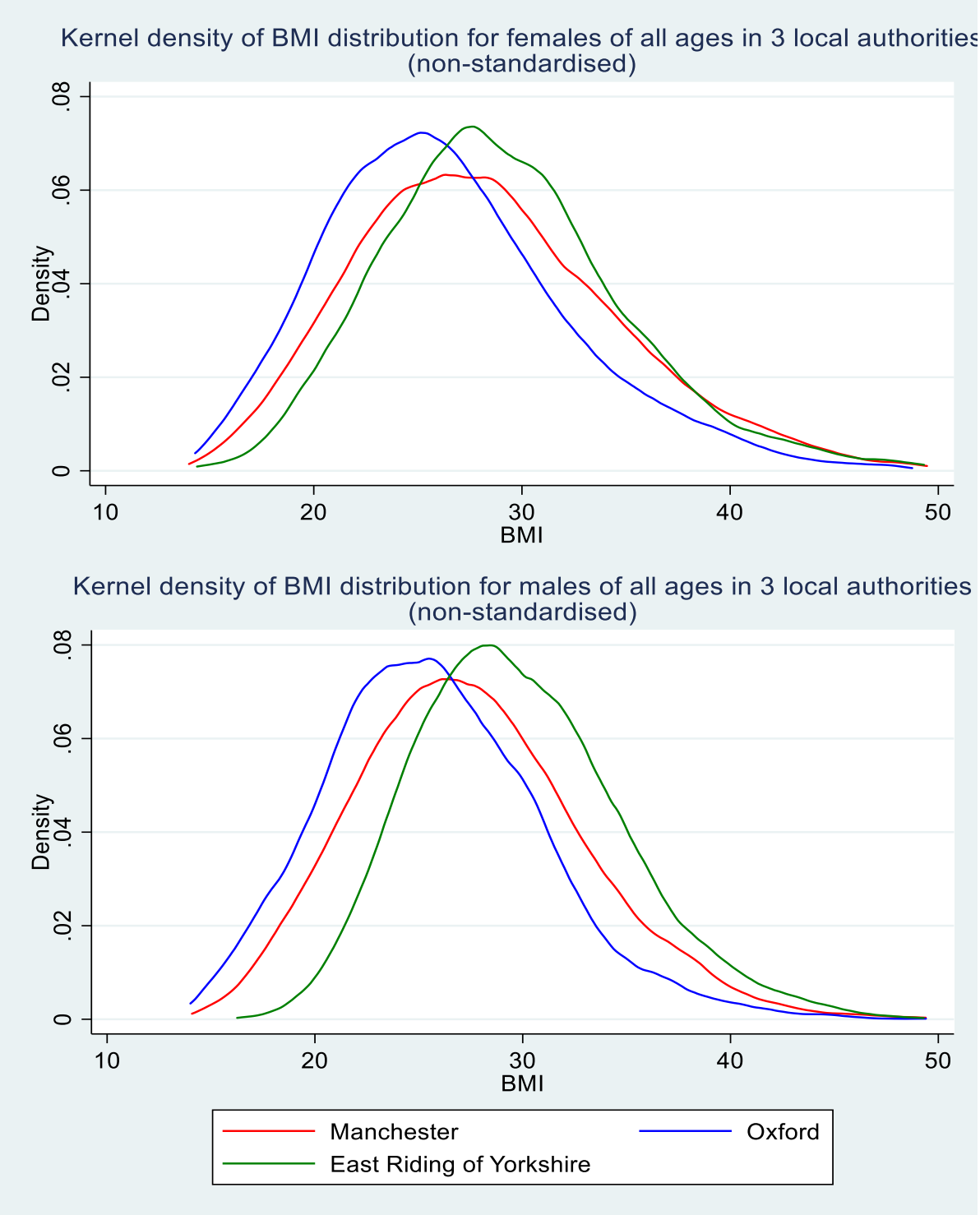


Figure 5.4a (above) and 5.4b (below): Kernel density plot of BMI of all-ages in three local authorities, for females and males, respectively. Manchester (IMD quintile 5 – highest deprivation), East Riding of Yorkshire (IMD quintile 2) and Oxford (IMD quintile 3).

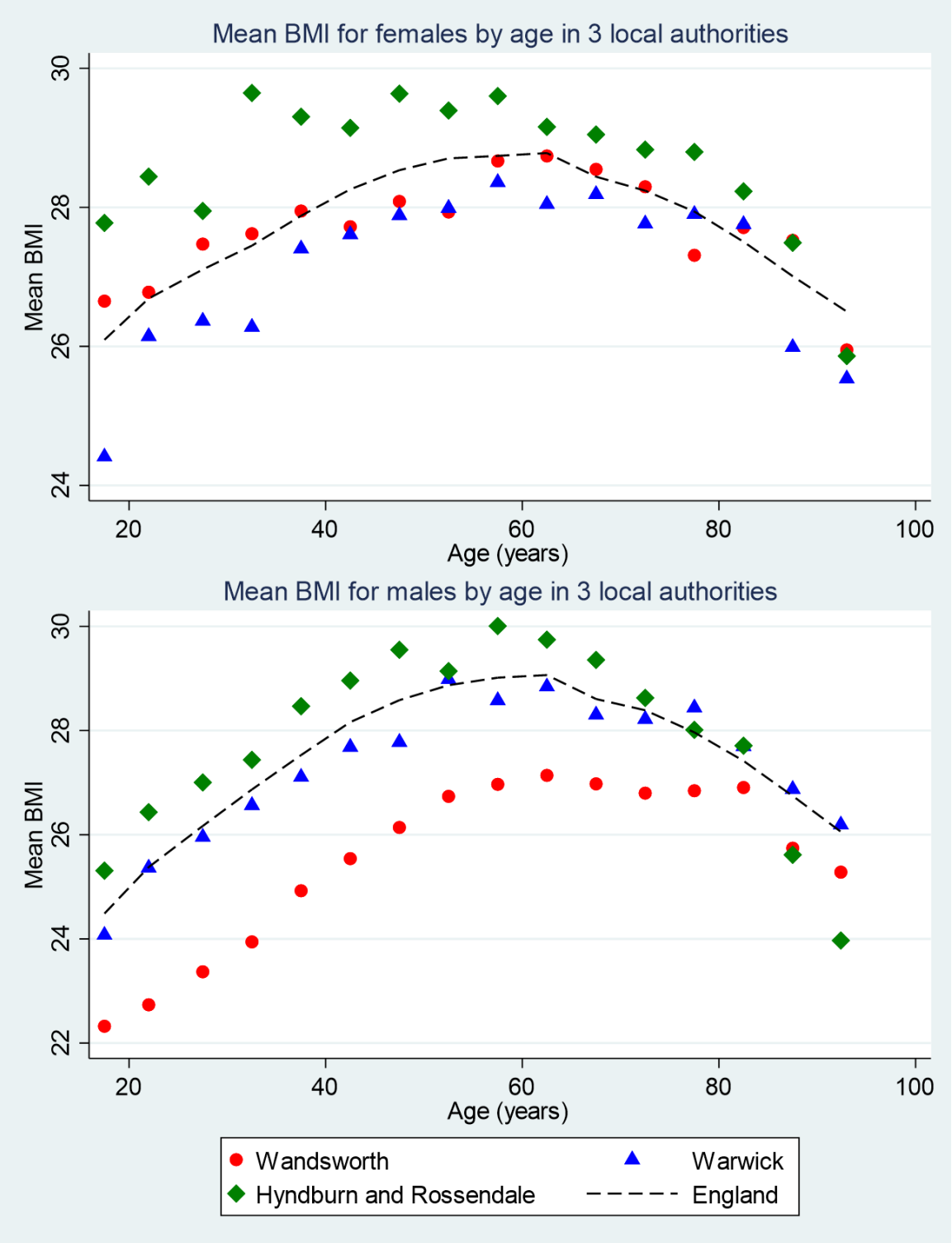


Figure 5.5a (above) and 5.5b (below): Distribution of BMI across the age range in three local authorities for females and males, respectively: Hyndburn and Rossendale (IMD quintile 5 – highest deprivation) Wandsworth (IMD quintile 3) and Warwick (IMD quintile 1)

Validation

Face validity

Face validity was considered in model design, as described above in the methods section. Variables were selected on the basis of being likely to be associated with BMI on the basis of previous data and theory, building in face validity to the model design.

Candidate variables were chosen on the basis of existing evidence and theory so all have some association with BMI. Ethnicity and sex are known to be related to BMI while age is also known to have a non-linear relationship with BMI.(126,147) Other variables tested represent different proxies for socioeconomic position. For example, tenure (after accounting for age) might represent dimensions of income and other sources of finance (for example from family) that make it more likely someone needs to access social housing or is able to save for a deposit and obtain a mortgage. The variables part-time work, self-employment, being retired, having long-term sickness, having no qualifications, and having a degree all also represent some form of individual-level dimension of socioeconomic position. Local authority-level IMD reflects area-level deprivation that may influence individual-level deprivation. Region and RUC are both known to be associated with BMI, as they may reflect geographic socioeconomic differences or local contextual variation.(126,129,147,148)

Heteroskedasticity

The plot of residuals against estimated log-BMI at the individual level is shown in figure 5.6. Examination of this plot does not identify heteroskedasticity, which is reinforced that the

error bars (in red) representing the 95% spread of data points appear consistent across five quintiles of log-BMI observations. The spread appears slightly asymmetrical around zero, though this asymmetry is found to be small, especially for females. This is demonstrated more clearly in the histogram of residuals in Appendix figure 1d.5. Formal testing was not performed as these tests are not valid with robust SE, as explained in the methods. Finally, figure 5.6 also shows trend lines for females and males in green and orange, respectively, demonstrating no overall bias in estimation across the range of log-BMI.

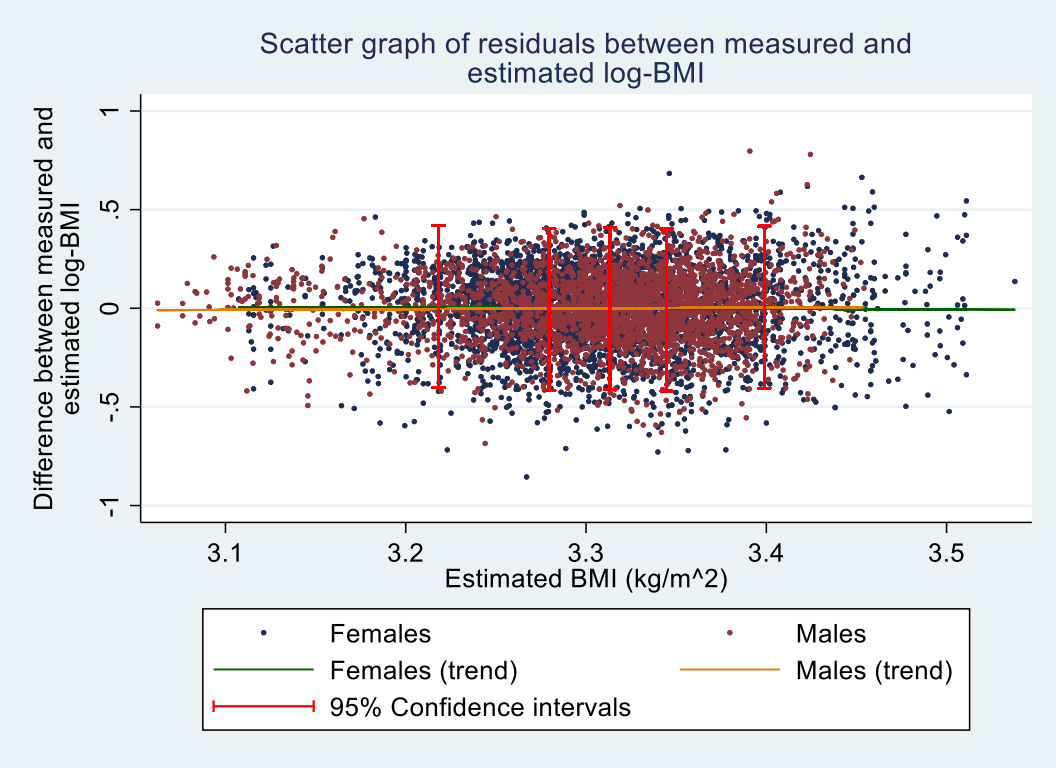


Figure 5.6: Plot of residuals between measured log-BMI in the HSE 2018 and model-based log-BMI estimates.

Internal validity

Comparing regional-level estimated and measured BMI in figure 5.7 shows similar overall patterns of BMI. There was close similarity between estimated and measured mean BMI, though it was less close for Yorkshire and Humber, which was estimated to be higher than measured, and London, which was estimated to be lower than measured.

The age-standardised England-level distributions of estimated BMI for females and males are compared with the equivalent distributions of measured BMI from the HSE 2018 (that the models were based on), shown in figures 5.1a and 5.1b at the top of the results section. By comparison, figure 5.8 shows the distributions of BMI over the age range in terms of the uncalibrated estimates (in red), calibrated estimates (in blue) and measured HSE 2018 data (in grey). The medians are given as points across the distribution, with 75th and 25th centiles of BMI distribution above and below the medians, respectively. The distribution of uncalibrated BMI is narrower than measured BMI, particularly for younger age groups, while the calibrated estimates conform well.

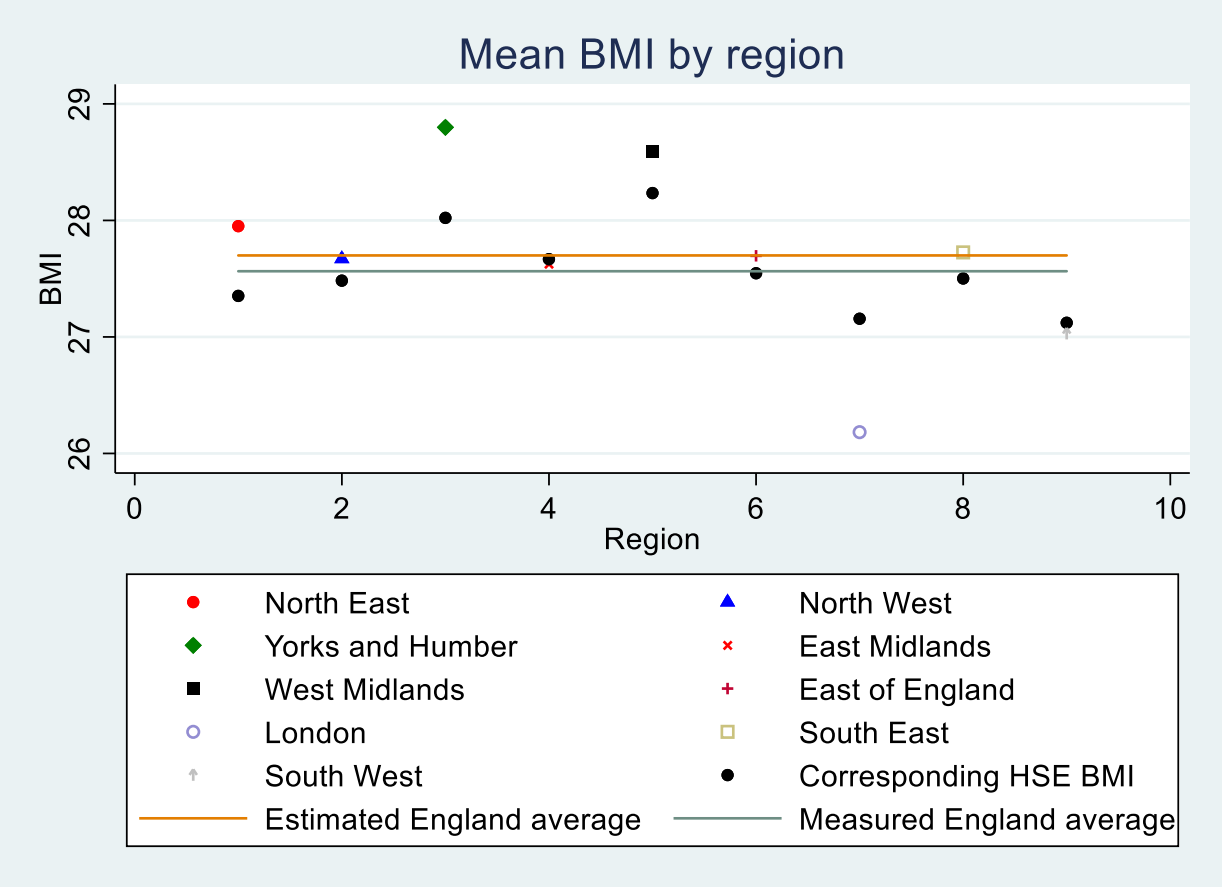


Figure 5.7: The mean estimated BMI by England NUTS-1 region and the mean measured BMI from HSE 2018 (standardising all age and sex groups equally). Regions run left to right on the X-axis, with the coloured shape marker representing mean estimated BMI (Y-axis) and black dot marker representing the mean measured BMI for that area.

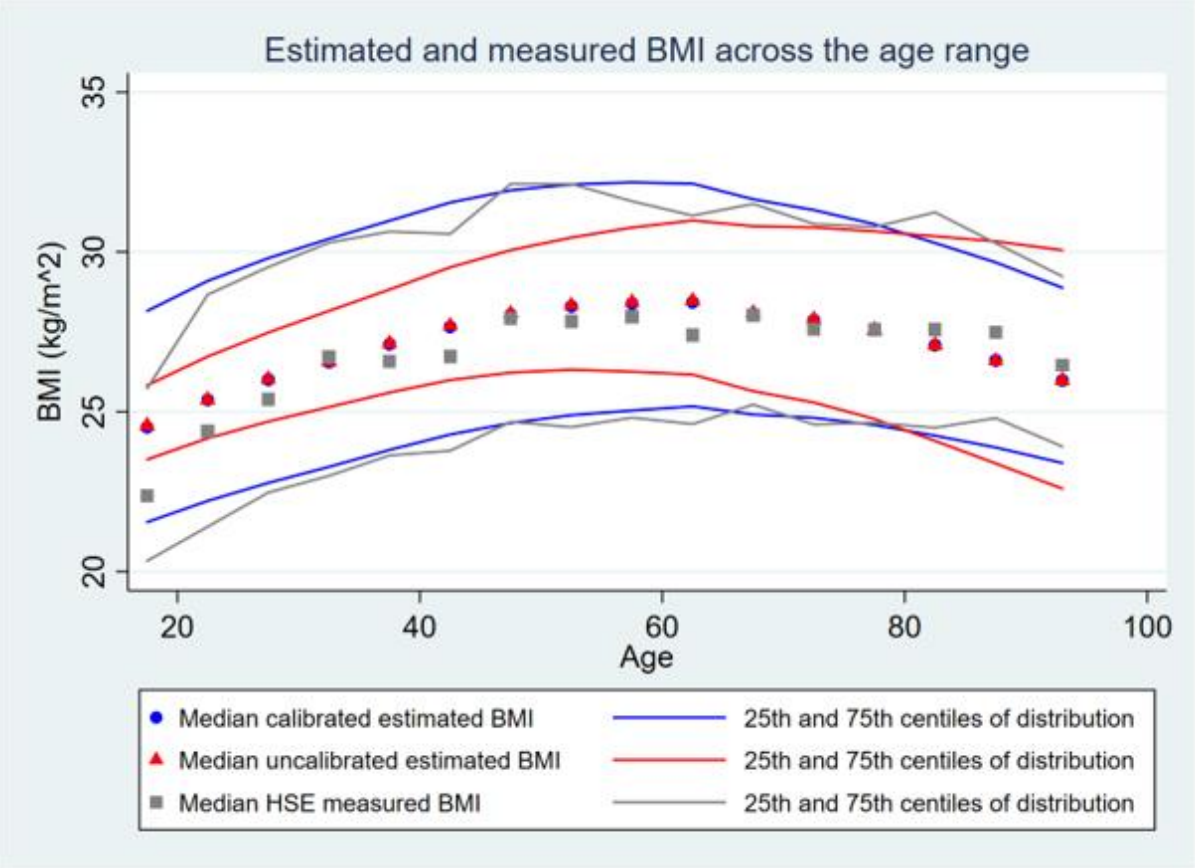


Figure 5.8: Median, 25th and 75th centiles of BMI distributions from HSE 2018, uncalibrated estimates and calibrated estimates (both sexes combined).

The relationship between mean estimated and measured age- and sex-group mean BMI is shown in figure 5.9. Each marker in the figures represents one age category, with blue for females and red for males. This shows a linear relationship between the two sets of BMIs. The regression models shown in table 5.3 quantify these differences between sexes and show that across the aggregated population there is no overall bias. Both uncalibrated and calibrated estimated BMIs have close relationships with the measured BMI data from HSE. Regression of the estimated BMIs with measured BMIs (weighting age and sex groups equally) gives a coefficient of 1.07 for the calibrated estimates or the uncalibrated estimates.

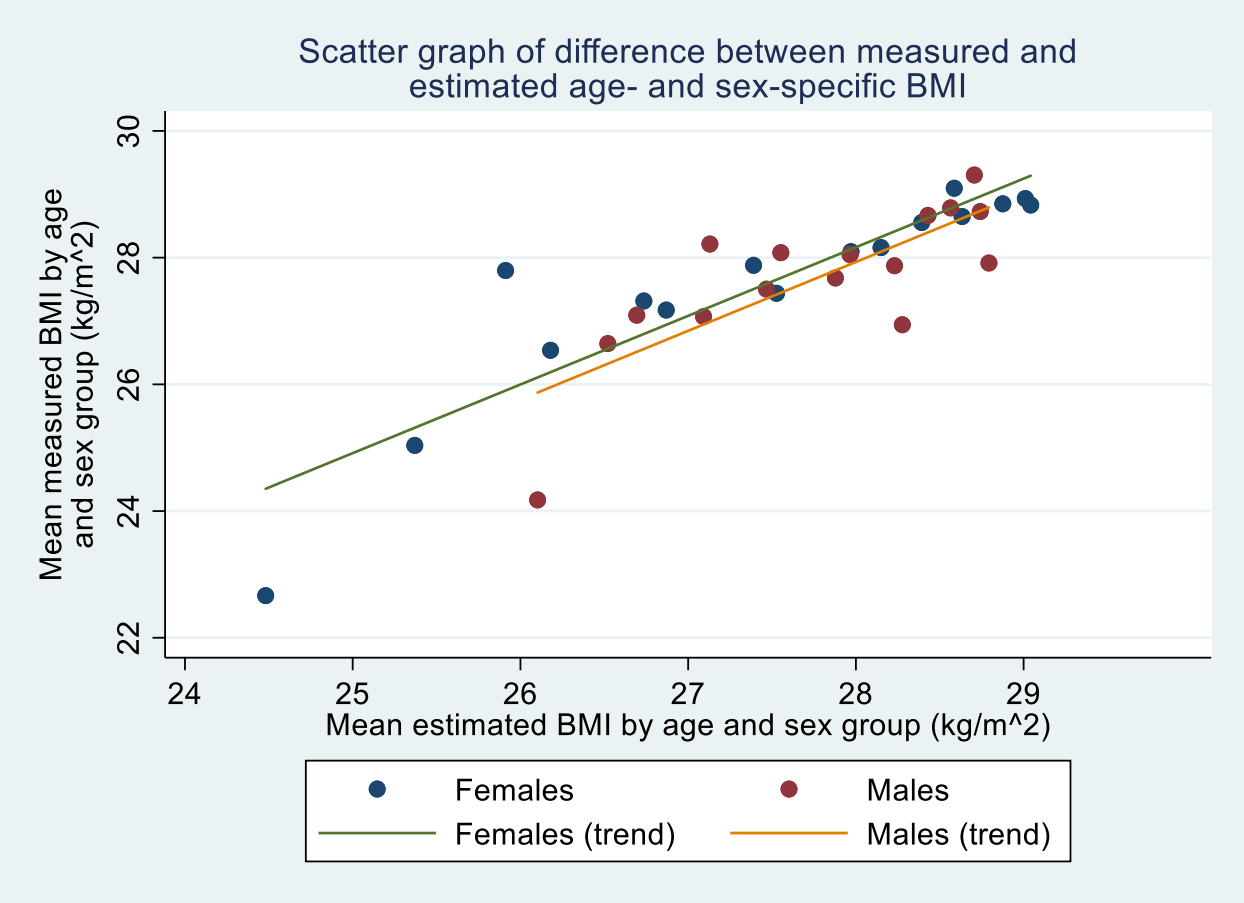


Figure 9: Plot of mean estimated (and calibrated) BMI by age and sex group against those measured in HSE 2018, with best fit lines by sex. Each marker represents one age category.

Table 5.3: Regression model outputs quantifying the relationships between age group level measured BMI with their equivalent age-group level estimated BMI, uncalibrated and calibrated, by females, males and both.

	Uncalibrated BMI estimated			**Calibrated BMI estimates**		
	Females	Males	Both	Females	Males	Both
Measured BMI	**1.11**	**1.06**	**1.07**	**1.09**	**1.06**	**1.07**
SE	0.190	0.119	0.999	0.233	0.137	0.116
Constant	**-1.49**	**-2.52**	**-1.80**	**-2.48**	**-2.17**	**-1.91**
SE	5.23	3.24	2.74	6.47	3.76	3.19

Convergent validity

Estimated BMI distributions for females and males were compared with the equivalent distributions from the independently-collected NDNS 2014-16 data (standardised to represent each age group equally). The overall distribution of BMI across England for each sex were shown to be similar. Figures 5.10a and 5.10b show similar differences between estimated BMI and measured BMI as shown in figures 5.1a and 5.1b.

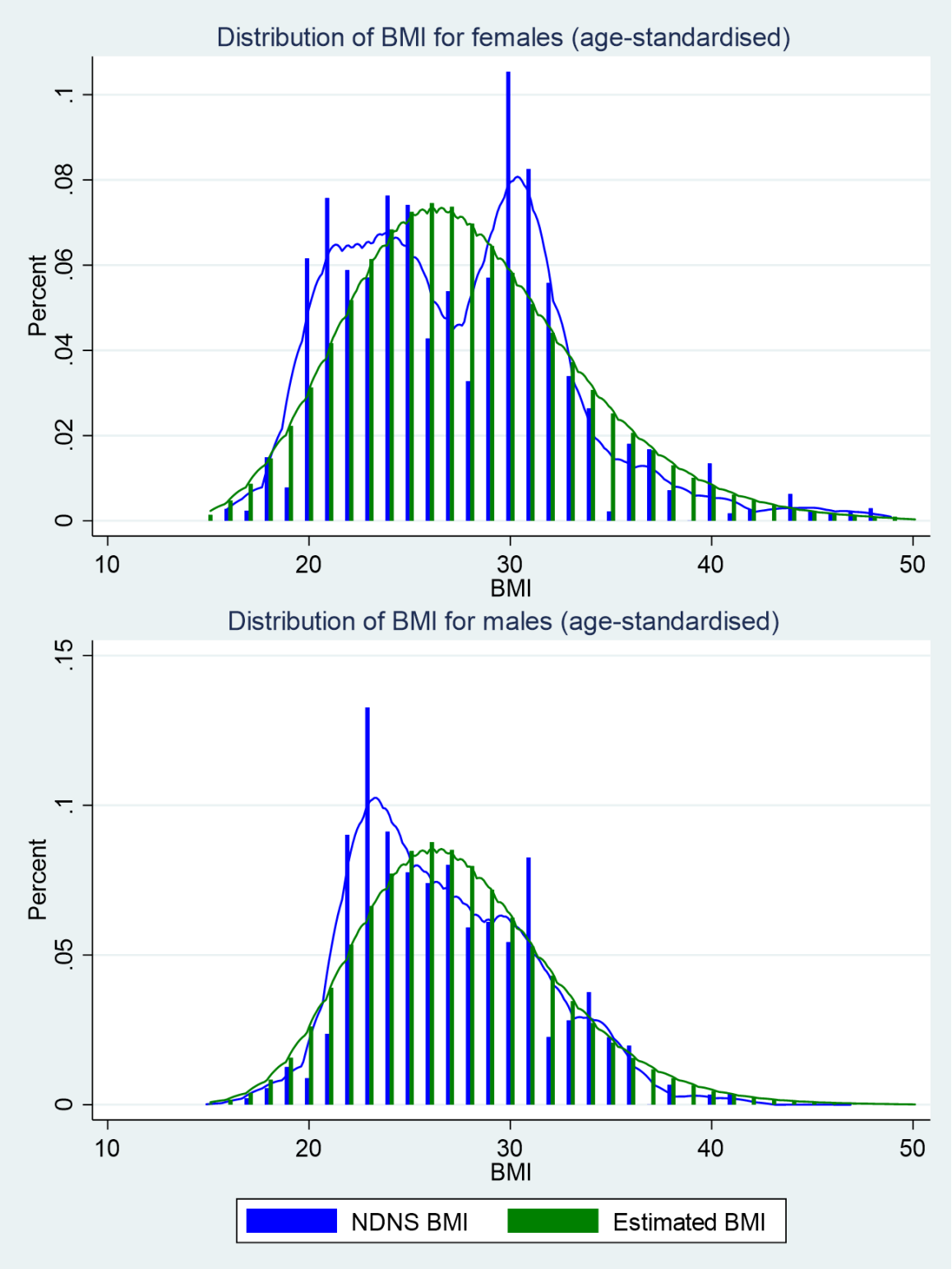


Figure 5.10a (above) and 5.10b (below): Histograms of age- and sex-standardised BMI estimates for England for females and males, respectively, including the equivalent histogram for measured BMI from NDNS 2014-16 for comparison.

There was a close relationship between the BMI profiles in the estimates to those reported at the local authority level in the ALS. This is shown in Figure 5.11, which demonstrates the percent in each local authority estimated to have a BMI >25 plotted against the percent self-reported to have BMI >25 in the ALS. The relationship is broadly linear, with higher proportions estimated to have high BMI than was reported in the ALS data. Appendix 1d.6 provides equivalent figures for each BMI category separately.

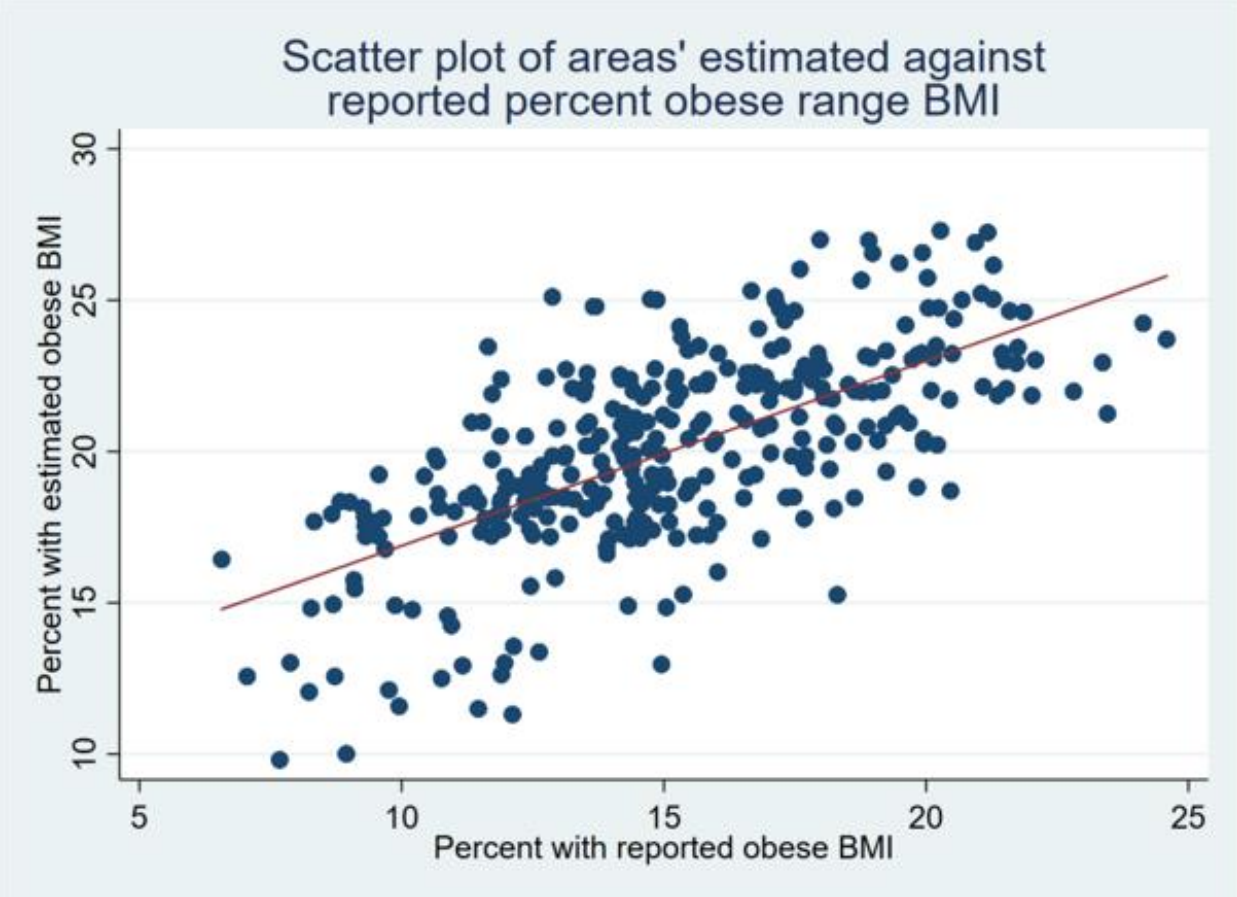


Figure 5.11: Scatter plot of local areas' percent of population with estimated BMI >25 against self-reported BMI >25 in ALS 2019-20.

Discussion

Summary of results

This method combined the individual-level associations between BMI and social and demographic features to estimate the BMI of each individual in the census microdata 2011, allowing a mean BMI and SD to be estimated for each age and sex group for each local authority. The regression coefficients broadly supported previous findings of associations between BMI and demographic/ social markers.(126) It builds on the standardisation methods, the work by Charlton (1998)(123) and previous work estimating local BMI.(126,129,149) It adds further detail by estimating BMI distributions (rather than categories) for age groups for each sex at the local authority level in England.

The distributions of group-level estimates were calibrated to replicate the national distribution of BMI in the HSE 2018. The spread of calibrated and uncalibrated estimates is shown alongside measured BMI in figure 8. That the uncalibrated estimates had a narrower spread than the measured HSE data implies that associations between BMI and the predictor variables was driven more by data points in older age groups and individual-level estimation was less accurate in younger groups. It is possible, for instance, that a variable was missing that predicted BMI in younger age groups. The regression outputs in table 5.3 show that this process made a very small impact on average age-sex group level proximity to the base HSE 2018 data.

This synthetic estimation process found that BMI varied by age, age squared, sex, age interacting with sex, age squared interacting with sex, ethnicity, being self-employed, being retired, being long-term sick, having no qualifications, having a degree, deprivation, self-reported health, tenure status, region of England, and SES.

Overall, age had a non-linear relationship with BMI, peaking in the sixties then declining again. The distribution of BMI was more skewed for women, with a higher proportion of people at lower BMIs, shown in Figure 5.1. Figures 5.2, 5.4, 5.5 and 5.7 show the orders of magnitude that BMI distributions varied between regions and between local authorities. Differences reflected conventional ideas(6,126,147) that BMIs are likely to be highest in areas with older populations and greater deprivation. The age-sex group breakdown of mean estimated BMI is summarised for three areas in Figures 5.6a/b, demonstrating generally higher average BMI in more deprived areas and the non-linear association between age and BMI.

Validation

By design, the variables used were all empirically or theoretically associated with BMI, conferring a face validity to the model building exercise. There was no evidence of heteroskedasticity at the individual level by visual inspection, though formal testing was not possible. At the regional level (NUTS-1), mean BMI in most areas aligned closely, with the exceptions of Yorkshire and the Humber and London. The measured HSE data put these further from the mean, than the measured BMIs from HSE. As shown in table 5.3, regressing estimates of age- and sex-specific BMI against those from the HSE 2018 found close linear relationships with either calibrated or uncalibrated estimates.

External validation against the NDNS 2014-16 data showed good concordance between the estimated BMI and those measured in the survey, shown in figure 5.10a/b. The bimodal distribution in females' NDNS data is likely to be related to small sample size, not true underlying distribution.

The comparison between area-level estimates of prevalence of high BMI with self-reported BMI in the ALS (figure 5.11) shows a close relationship. Estimates predicted higher proportions of the population in the high BMI categories than was reported in the ALS data. This could be for a variety of reasons. The ALS relying on self-reported BMI leaves it vulnerable to reporting bias, while any such survey may suffer from a response bias, for example of people more willing to participate being those with a better understanding of research, so may have higher level of education and lower BMI, or of people being happier to respond who are in better health, including lower BMI. Both of these phenomena could bias the ALS data to lead to the findings in figure 5.11. It seems less likely that the model-based estimates presented here have substantial bias that would lead to big overestimates of high BMI, as this would be identified on internal validation against the NUTS-1 regions and external validation against the NDNS population, neither of which suggest an overestimate of the prevalence of high BMI at the population level. What these estimates may suggest, therefore, is that the differences in BMI profiles between local authorities may be greater than previously suggested.

Strengths and limitations

This novel approach of estimating BMI at the individual level has allowed estimates to be drawn with a greater degree of detail than previously, namely estimating the whole BMI

distribution, separately for each age and sex group and each local area. GLM regression is well suited to modelling on skewed data such as BMI, picking up increasing variance along the range of the independent variables. The Census microdata's 2.1 million entries (in England) are highly granular and representative of the population. A detailed process examining for alignment between variables between datasets ensured the maximum number of variables were available in the estimation process. In combination with the HSE's detailed social and demographic information, this novel approach using individual-level and population-level modelling has provided a method of estimating the heterogeneity of BMI in each local area and with greater detail than previously.

The relative strengths of the available BMI datasets were considered and HSE chosen with benefits over the ALS of using measured rather than self-reported BMI and having continuous BMI variables rather than reporting only BMI category. It was chosen over NDNS for being much larger so more able to produce precise estimates of BMI distributions. Data quality of the HSE is generally good, being a large and representative sample. One relevant limitation to the HSE 2018 is missing data. As described in table 5.1 and in Appendix 1d.4, there is generally little missing data, except for the BMI variable, particularly in the higher age groups. It is not possible to know what impact this missing data may have on results or to account for unobserved differences, though age is included for in the model and other observed differences may be captured with other variables included. One further limitation is that the 2018 HSE data is that the new HSE 2021 data now indicate adult BMI may have risen sharply during the COVID-19 pandemic.(150) Using the lower BMI data of 2018 would confer a conservative bias on the magnitudes of modelled outcomes, though it is not possible to tell how between-area differences may be affected.

The limitations of this approach are mainly associated with the variables that can be aligned between the HSE 2018 and the Census microdata. To identify all relevant variables, a step of mapping the appropriate variables in each dataset against one another was completed (tabled in Appendix 1d.1), but limitations remain. First, the levels of geography do not align between the HSE and Census microdata, other than at the NUTS-1 level. This means that the multilevel modelling process designed by Twigg *et al* 2000(122) is not feasible as the NUTS-1 areas are far too large and heterogeneous to pick up the contextual subtleties the method intends to. The consequence here is that this method assumes that the national-level associations between the predictor variables and BMI do not systematically vary between local areas. Second, deprivation is measured at the postcode level in the HSE 2018 but it is only possible for the Census microdata to merge in deprivation at the local authority level. This means that the deprivation represented within local authorities in the Census is not heterogeneous as in the real world. This captures some of the association between deprivation and BMI but is likely to reduce variation modelled within each local area.

The additional step calibrating the estimates may have been necessary due to insufficient variables being compatible between the two datasets to capture the true breadth of BMI in younger age groups. The use of this calibration step required an additional assumption that the variables that were available were randomly distributed among a hypothetical set of variables that might predict the distribution more accurately, so the *order* that individuals' BMIs were estimated was correct. Therefore, by correcting the distribution to a wider spread, observations were made more extreme but assuming the order of BMI estimated was appropriate.

Figure 5.5a/b shows volatility in these estimates between adjacent age groups that due to the method would arise due to variation in the age-specific social and demographic variables. In the context of aggregate analysis in a longitudinal model (such as PRIMEtime) this is not important as it will be smoothed out by the cumulative passing of time for each age/ sex cohort in the model structure. However, taken as the residual variation not picked up by the modelling process, it may be an important limitation for other uses or for modelling of a specific age/ sex group at a time. Model-based small area estimates such as these can never eliminate all residual, unobserved variation, so cannot act as perfect replacement for adequately precise small-area measurements. There were also 242 age-sex-area subgroups with sample sizes <30 in the two oldest age groups, conferring increased risk of sampling error. As the overall populations of these groups are small, this is unlikely to introduce bias in population modelling, though users should be aware of this risk if they are focusing analyses on older people.

Census data is only collected every ten years, of time of writing (December 2020) nine years ago in 2011. It is difficult to know the impact of this passing of time on the estimated BMI distributions. Some areas may have aged faster or slower, but this will be accounted for elsewhere in the PRIMEtime modelling process by using the mid-year estimates as described in Chapter 3. Deprivation was merged into the Census data from the 2019 IMD(146) so change in deprivation since the last Census is accounted for. As the associations between social and demographic variables were also derived from a recent dataset (HSE 2018), the remaining factor that may have biased results due to the time since the Census is if the composition of the population within age-sex groups has changed, ie. the types of people in a given local authority has changed. This is difficult to account for, though only in a few

cases with rapid demographic change might we expect the order of local areas' estimated mean BMI or overweight/ obesity rates to change.

It is known that it is difficult to quantify uncertainty around small area estimates drawn from microdata approaches.(138) This is a key limitation to this approach, particularly as it would be helpful to incorporate uncertainty into Monte Carlo analysis in the disease modelling process.

Conclusion

This study used GLM regression in the HSE 2018 to estimate the BMI of respondents in the Census microdata 2011, providing the most recent set of disaggregated small area estimates for adult BMI at the local authority level in England since 2007. There is a suggestion that adult BMIs may be higher than acknowledged in many local areas in previous surveys and with greater difference between areas, which may arise due to response bias or sampling bias in the ALS. Models were well fitted and estimates appear unbiased against an unrelated source of measured BMI data.

These new estimates go further than previous studies by two dimensions – by estimating BMI by age-sex groups and by modelling the distribution of BMI rather than percentages in overweight and obesity groups. As such, this study has produced the most demographically granular estimates for BMI in local areas of England to date, which, when incorporated into PRIMEtime_local, will allow BMI-related risk to be modelled to a high degree of local granularity.

This Chapter has been peer reviewed and published as listed in Appendix 3.

Chapter 6: Estimating child BMI by age, sex and local authority

Introduction

To allow interventions targeted at children to be modelled at the local level it is also necessary to estimate the distributions of BMI by age and sex for each local authority. Children develop some disease outcomes related to BMI change, and/or change to children's BMI run through into adulthood, impacting adult disease burden. Children's BMI data are therefore required, but appropriate survey data are not available so they must be estimated.

More accurate local data on children's BMI exist than for adults, due to the National Child Measurement Programme (NCMP).(7) This is a survey established in 2006 to measure the height and weight of every child in reception year (aged 4-5) and year 6 (aged 10-11) at state-funded schools (assuming consent and school attendance).(7) Participation is very high, at approximately 95% of students across 99% of eligible schools. The HSE also measures a nationally-representative random sample of children as well as adults via its household survey approach (described in Chapter 5, above).

The aim was to extrapolate these data and produce estimates of the height and BMI of children by age and sex, aged 5-16 years. Mean BMI and SD BMI have to be estimated by local area separately to capture local variation in spread of BMI as well as overall tendency. Local height is also required as interventions exert an effect by changing calorie consumption, which is translated to body weight using the Hall(92,93) equations. Height

allows baseline body weight to be back-calculated in the model from the BMI parameters, which is then added to scenario weight change, to then recalculate scenario BMI.

The NCMP is much larger than the HSE and also includes a lower tier local authority identifier. The benefit of the HSE is that it measures the height and weight of children of every age (though they are reported as age groups). To estimate BMI for each year of age for each sex and local authority, a pragmatic method was developed using the strengths of the two datasets.

It is worth noting that child obesity rates rapidly increased in the COVID-19 pandemic. Rates of obesity in reception class increased by 4.5 percentage points on average – but with a dramatic increasing in inequalities, with an increase of 7 percentage points for the most deprived decile of children and just 1.8 percentage points for the least deprived. The increase in inequalities at Year 6 was marginally less severe, at 6.3% and 2.4% respectively, though partly as rates in the least deprived groups increased faster in year 6 than in Reception class.(151) For Reception class, the overall rate of increase was over 25 times higher than the average seen over the five years since the previous nadir in 2014-15 and for Year 6 it was over 11 times.(152)

Methods

Estimating mean and SD BMI by age, sex and local authority area

The local specificity of the NCMP with the full child age range in the HSE are complementary strengths, allowing local BMI distributions across the child age range to be estimated. To do

this, the most recent two waves of the NCMP were used (2018-19 and 2017-18) and mean BMI calculated by age, sex and local authority. Reception year was taken to represent age 5 and year 6 to represent age 11. National-level mean and SD of BMI by age and sex was also estimated. Using these, a z-score of log-BMI was calculated for each age/sex group as *z = (local mean – national mean)/ national SD*.

Between these two ages (5 and 11), z-scores were interpolated linearly. For example, if the z-score for girls in Barnet was 1 at age 5 and 1.6 at age 11, it would be interpolated as 1.1 at age 6, 1.2 at age 7, 1.3 at age 8, and so on.

The HSE release 2019 was used to estimate England-level BMI across the age range. The smallest children's age categories available are the seven groups 0-1 years, 2-4, 5-7, 8-10, 11-12, 13-15 and 16-19. Mean BMI was taken by sex for each age group to represent the central age (eg. 3 years for the 2-4 years group, 11.5 years for the 11-12 years group). To capture the gentle S-shaped curves of BMI change over childhood (shown in figures 6.1 and 6.2(153)) it was chosen to regress these means against age, age^2 and age^3 on a theoretical basis (Ordinary Least Squares, OLS). A smoothed distribution of BMI for each year of the age range was then estimated from the predicted values from that regression model, for females and males separately.

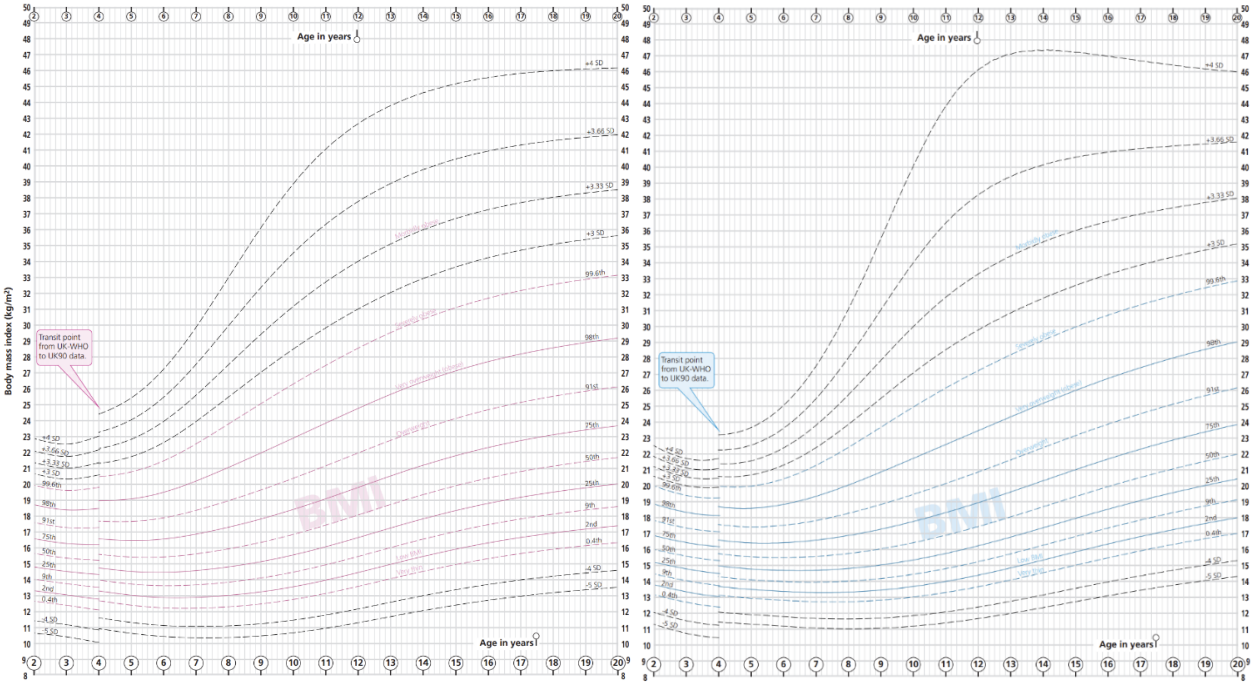


Figure 6.1: BMI centile chart for females (left) and males (right) aged 2-20 years (from the Royal College of Paediatrics and Child Health/ World Health Organisation)(153)

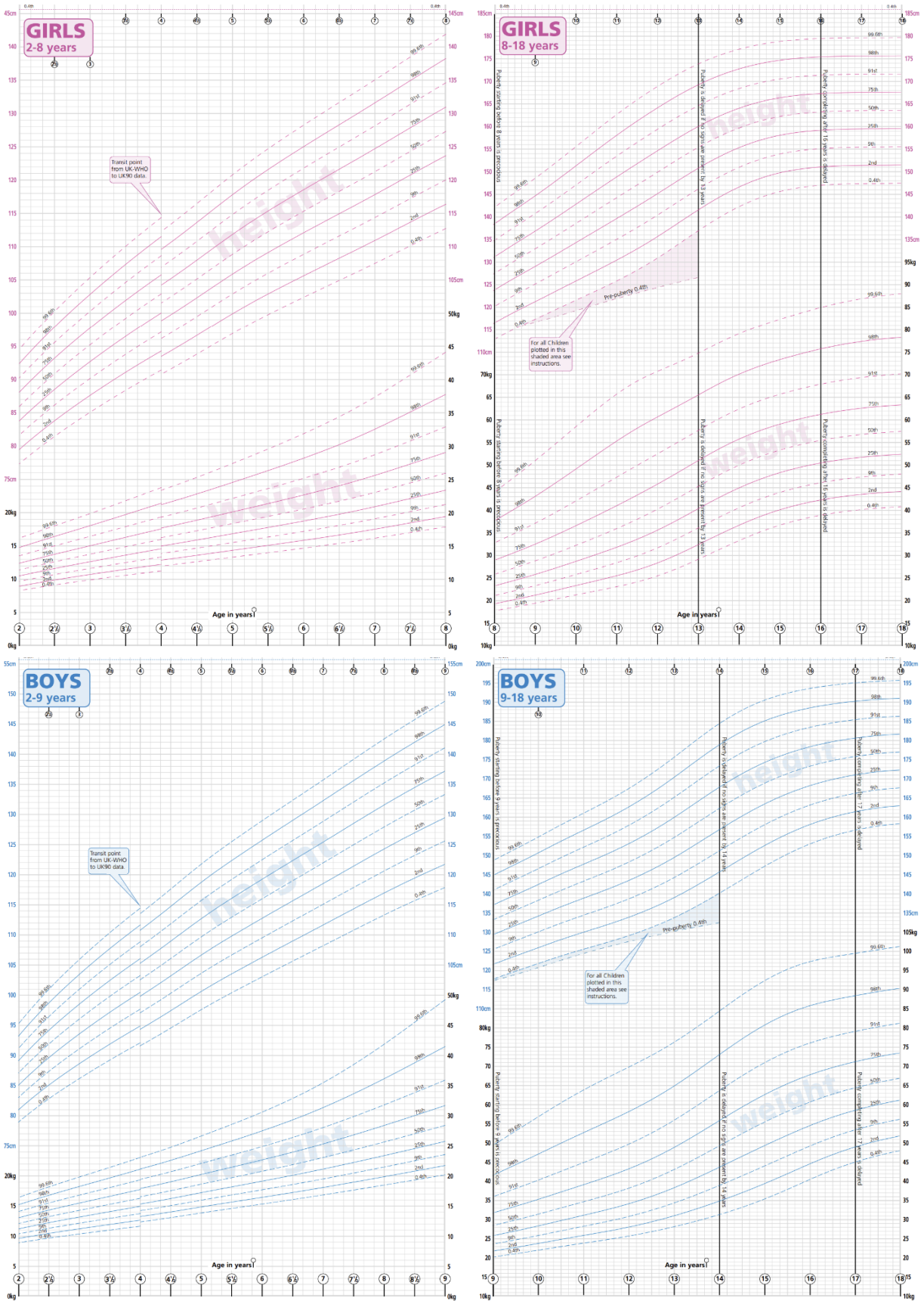


Figure 6.2: Height (and weight) centile chart for females (above) and males (below) aged 2-9 (left) and 9-18 years (right) (from the Royal College of Paediatrics and Child Health/ World Health Organisation)(153)

National-level SD of BMI was linearly smoothed in the same way as z-scores, estimating SD of BMI for ages between those measured in the NCMP dataset. For example, if SD of BMI was 0.050 for age 5 and 0.056 at age 11, they would be interpolated as 0.051 for age 6, 0.052 for age 7, 0.053 for age 8, and so on.

Local area mean BMI was then estimated by back transforming interpolated z-scores in combination with predicted mean BMI from HSE and smoothed SD BMI from NCMP as:

$$\mu_i = (z_i \bar{\sigma}) + \bar{\mu}$$

where $\mu_i$is the local estimate of mean BMI, z_i is the local z-BMI, $\bar{\sigma}$ is national average SD BMI and $\bar{\mu}$ is national average mean BMI. A similar process was used to estimate mean and SD BMI between 11 and 17. For younger children, the NCMP data could be used to produce z-scores at age 5 and 11 which dictated how the area-level mean BMI changed between those ages. For the older children, a second age point was needed to provide a z-score by local authority. This was provided by the age/ sex/ local authority level estimates of BMI for 17 year-olds from the adult local-area estimated above in Chapter 5. Again, the same process was used to estimate mean height by age, sex and lower tier local authority.

SD of BMI and mean height were estimated in an equivalent way, calculating z-scores by age, sex and local authority from NCMP 2017-19 data, then linearly interpolating between the ages of 5 and 11 years, and 11 and 17 years, taking 17 year-olds' data from the BMI estimates in Chapter 5. The estimates of BMI at ages 5, 11 and 17 years from HSE were compared with the equivalent estimates from NCMP and synthetic estimates from Chapter 5. Differences between HSE and NCMP were 1-2% at age 5 and 11 for females and males.

Differences between HSE and the synthetic estimates were slightly greater, as would be anticipated due to the methods, at 5-6%.

Managing data censorship in NCMP

An issue with data quality in the NCMP renders some areas' microdata unusable for this purpose. This issue arises as some areas have less than five observations in a given BMI category (usually underweight) so are suppressed, then to prevent the numbers in that category being back-calculated, another category also has to be suppressed. NHS Digital choose to suppress the 'healthy weight' category as this second category, as data users are generally more interested in overweight/ obese rates than the healthy population. The impact on the microdata for affected areas is demonstrated in Appendix 1e. For these areas, if one age group's data was unusable in both waves, then mean and SD of BMI was assumed to be the same as the other, usable, age group. If neither were usable then both were assumed to be the same as in the most socially similar lower tier local authority, as judged by the ONS residential-based area classifications model.(154)

Results

Summary statistics from the raw data are presented in table 6.1 to demonstrate mean area characteristics, and *variation in area means* presented as SDs (taken after dropping data affected by suppression and merging the two NCMP survey waves).This shows that males were slightly more represented in both year groups. Mean height was slightly greater for

females in both school years. Variation between areas' mean BMI, height and age was consistent between sexes. Six areas did not have a usable school year of data in either wave due to data censorship, and results for these were substituted in.(154)

Table 6.1: NCMP data 2017-18 and 2018-19 showing average and spread of anthropometrics, presented as mean of mean values, and the SD of mean values, by school year and sex (to 3 significant figures, s.f.).

	Reception		Year 6	
	Females	Males	Females	Males
N (individuals)	522,181	543,050	563,950	587,000
Proportion of population (%)	48.9	51.1	49.0	51.0
Mean BMI (kg/m^2)	16.1	16.1	19.1	18.9
Mean height (cm)	109	110	145.6	144.8
Mean age (years)	4.98	4.98	10.8	10.8
N (areas)	266	266	309	309
SD of local mean BMI (kg/m^2)	0.321	0.368	0.461	0.468
SD of local mean height (cm)	0.908	0.889	0.123	0.104
SD of local mean age (years)	0.123	0.123	0.160	0.161

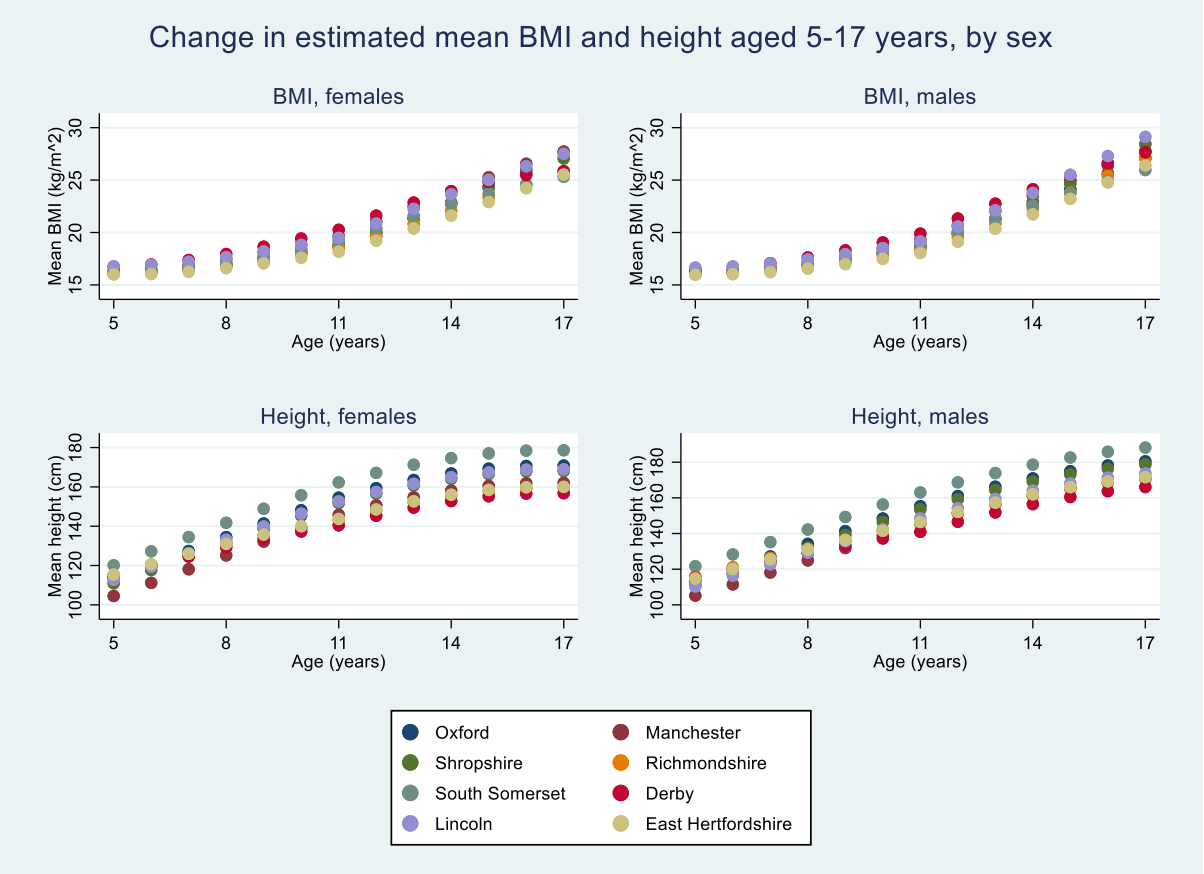


Figure 6.3: Estimated BMI (above) and height (below) for females (left) and males (right)

Figure 6.3 shows the changing of estimated BMI over the age range in the above panels, for females and males, for eight selected local authorities. These figures demonstrate gentle concave curvatures. Figures 6.3 also shows the range of heights across the age range by sex in the below panels, with slightly convex forms. Finally, examples of distributions of BMI across the 95% spread of the full population are shown in figure 6.4, showing gradually increasing spread, slightly more-so for males than females and more for younger ages. These are calculated from SD via log-BMI to appropriately capture the skew of the distributions.

For comparison, figures 6.1 and 6.2 show the 1991 centile charts (from the Royal College of Paediatrics and Child Health/ World Health Organisation)(154) for BMI and height,

respectively. These corroborate the overall forms of the curves and support the distributions shown in figure 6.4 of spread increasing more quickly at lower ages and starting to plateau through the teens as BMI continues to slowly climb.

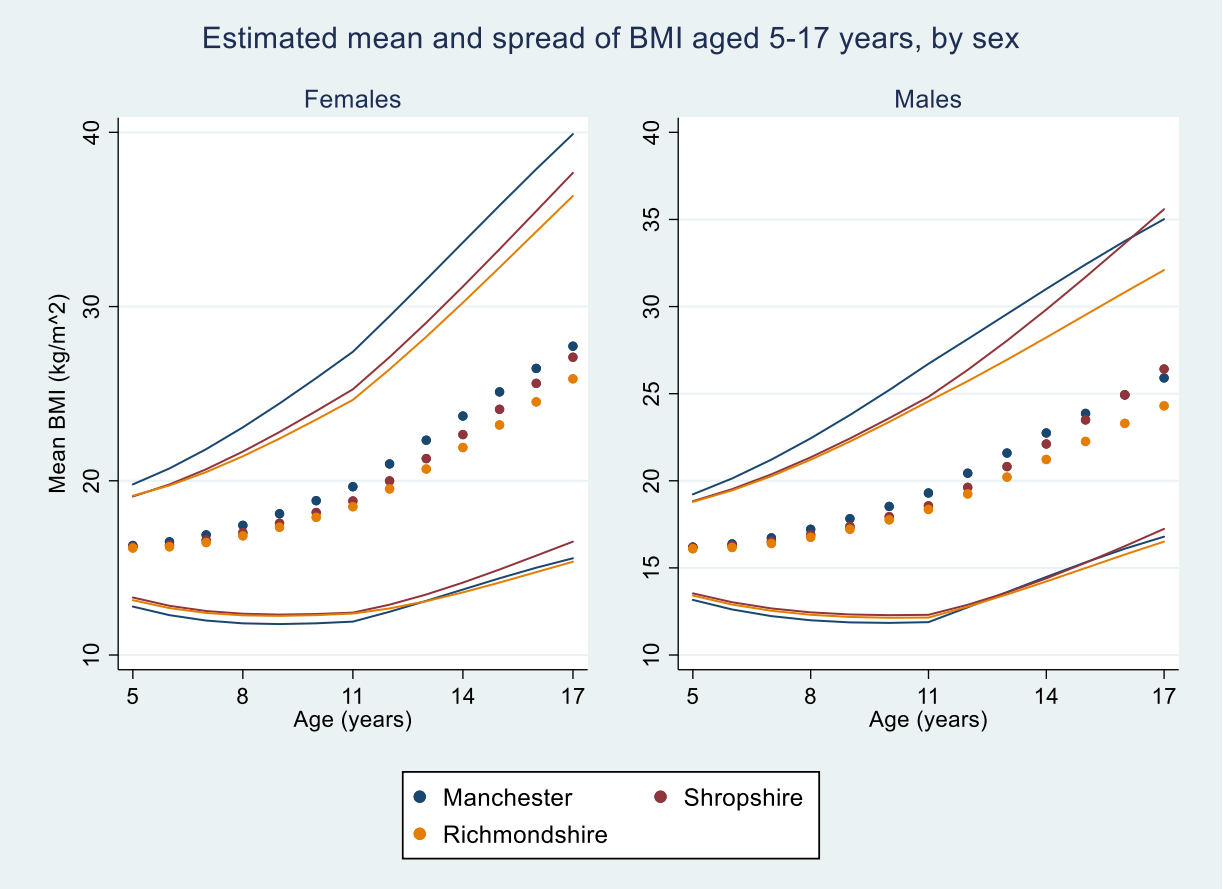


Figure 6.4: Estimated mean BMI (markers) with 95% spread of BMI distribution (lines above and below the means) for females and males by, selected areas.

Discussion

Findings

The process produced estimates of mean and SD BMI and mean height, by age 5-17, sex and lower tier local authority area consistent with locally-collected microdata where available. These could be expected to be broadly consistent with the centile curves published first in 1991 as the *normative* description of children's growth, but as these are not updated over time so no longer descriptive, the 30 years of increasing childhood obesity would be expected to show up as differences between estimated distributions and those plotted on the centile charts. Estimated height and BMI follow approximately the lines dictated by the charts. Here, specifically tracking the 50thcentile line (median) as an approximation to the mean height, we see female heights linearly increasing up to age 13-14 years then starting to plateau. The same is true of males, with the plateau slightly slower and later age at 14-15 years. BMI follows somewhat an opposite pattern, with initially slow rates of change increasing across the age range to a more linear increase. This trend is picked up well, noting that the slight slowing of the increase in female BMI on the growth charts happens largely after age 17 years.

Strengths and limitations

There are a few limitations to note to this method. One major limitation to this approach is inherent in the basic approach of interpolation that intermediate points between those measured can be estimated but are not in fact known. Further, the data points at age 17 are themselves modelled estimates (from Chapter 5) as measured BMI and height data by local

area do not exist. The linear interpolation of SD BMI and the BMI z-scores was a necessary simplification as the age categories in HSE are too wide to use directly, in the case of SD BMI, while there are no other area-specific data to inform the paths of interpolated z-scores.

The NCMP microdata is affected by data censorship, as described above. In brief, some areas have their BMI values suppressed for the underweight and healthy weight categories, rendering those areas' data unusable for these purposes (Appendix 1e). After aggregating the microdata from two survey waves, six local authority areas had censored (unusable) data in both reception year and year 6, while the vast majority had usable data in both school years. None had no usable data for only reception year and 43 had unusable data for only year 6. The NCMP is a survey of only state-funded schools, which achieve a participation rate of 95%, albeit with some variation. Although private schools are encouraged to measure and submit data to the NCMP, these schools educate 5% of primary school children(155) but make up only 0.3% of total submissions. These missing data for privately educated pupils are highly non-random, with a great majority from the richest 10% of households,(155) it is likely these children would be less likely than the main NCMP sample to have above healthy-range BMI. As private schools are also more concentrated in and around London,(155) this may slightly bias some local authorities in these areas to appear to have a lower distribution of BMIs than the 'true' distributions.

The limitations to the HSE are described above (in Chapter 5) but here the well-designed and executed survey is not likely to have major implications other than that the age categories are wider than the ideal annual spacing, meaning that these individual years too had to be interpolated as predicted values from OLS regression modelling. This is unlikely to

bias the result on the assumption that the sample size of 105-194 observations for each of the five age-groups for each sex is sufficient. There are slight differences in the ages that children have their BMI measured in the year, with the SD of local authority mean age at measurement of 0.123 years for reception year and 0.160 years for year 6 (1.5 months and 1.9 months respectively). However, given the breadth of BMI both within and between local areas, these differences in growth related to age are likely to be extremely small.

The collective strengths and limitations of the two datasets had to be managed. The NCMP provides almost complete data on two school years, while the HSE provides the full spectrum of growth through childhood but with much smaller samples, aggregated to the England level. As the NCMP provides a very large sample, it was chosen to assume this was the 'correct' measurement for those two age groups and adjusting HSE estimates to allow the two to be compatible.

Finally, we cannot yet know the long-term impacts of increased child BMI related to the COVID-19 epidemic on longer-term trends; for example, if subsequent cohorts will also be of higher BMI, if future cohorts will return to the previous baseline, at the same or different trends, or if these differences from previous years will themselves disappear by adulthood with minimal difference for those children's long-term health. Knowing none of these things, we assume that future cohorts of children will have access to the same nutritional and PA environments as the 2017-19 cohort and their BMIs will be closely correlated with these previous years.

Conclusion

This work uses a method of interpolation to estimate the BMIs of children aged 6-10 and 12-17 by taking the local measured NCMP data for reception class and year 6, synthetic estimates of local adult BMIs (in Chapter 5) and the nationally-representative HSE BMI trajectories. The forms of the curves reflect those seen on centile charts from 1991, though the time since then makes comparison of the absolute values difficult. These local-specific estimates of child BMI distributions will allow PRIMEtime-local to reflect the level of BMI-related risk with a high degree of accuracy. The use of good quality locally-collected data is a major strength to the method and the local model. Their quality confers a high degree of confidence that the contribution of child BMI estimates to modelled differences is robust.

Chapter 7: Estimating local disease-specific costs

Introduction

Modelling costs in PRIMEtime

Introduction to modelling costs in PRIMEtime

PRIMEtime modelling structures have previously translated modelled disease burden into healthcare costs, calculated as the cost of treating a disease per year per patient to the NHS.(47,156,157) These are termed here as the 'Annual Cost of Disease'. As these Annual Costs of Disease have not been published by the NHS recently, methods were developed to estimate them.(156) However, to represent how modelled local level disease burden translates into local healthcare costs, it would be useful to apply estimated locally-specific Annual Costs of Disease, requiring new methods of cost estimation to be developed.

Previous methods for modelling Annual Cost of Disease for PRIMEtime

The previous method for estimating national-level Annual Cost of Disease is explained in full elsewhere.(156) The approach follows a top-down disaggregation of budgets. Three budget streams covering Primary Care costs, Core Service costs (including most hospital and ambulance services) and Specialised Services costs were each estimated separately and then combined, as follows.

Core Services accounted for £84.4bn of the £114bn NHS budget 2018/19 and cover most hospital functions (such as wards, clinics and emergency care), ambulances, mental health

services and community services.(158) These were distributed by NHS England to CCGs via funding formulae (though CCGs have now been abolished and replaced by Integrated Care Systems, ICSs). CCGs then commissioned services to local providers (usually NHS trusts). The amounts allocated to each of 56 Programme Budgeting (PB) categories (groups of ICD-10 (International Classification of Diseases, Tenth Revision) diagnostic codes approximating to clinical specialisms) from NHS England were used to estimate disease-specific costs. For example, osteoarthritis of the hip relates to ICD-10 codes M16-M16.9, which are only a small proportion of the codes making up PB category 15X (problems of the musculo-skeletal system). As PRIMEtime operates on disease burden that is more granularly categorised than the PB categories, it is necessary to calculate the amount of that PB category allocation spent on each disease. To do this, approximate proportions of NHS activity within each PB category were estimated, termed the "admissions ratio", calculated from HES.(159) This is calculated as the proportion of admitted patients in a given PB category that have the ICD-10 codes assigned to a specific disease. For example, of the 1,057,069 patients admitted under PB category 15X, 90,363 were admitted with ICD-10 codes M16-M16.9 for osteoarthritis of the hip, giving a proportion of 0.085. The spend on PB 15X of £4.9bn, multiplied by the admission ratio of 0.085 gives a Core Services spend on osteoarthritis of the hip of £424m.(47,156)

Specialised Services are commissioned centrally by NHS England at the cost of £24.8bn in 2018/19.(158) These services cover rare and complex conditions such as uncommon cancers, genetic conditions and advanced procedures such as thrombectomy for ischaemic stroke. Specialised services were last commissioned locally in 2012/13, before the implementation of the HSCA 2012. These budgets can be used to calculate the proportions of total PB spending made up by Core Services and Specialised Services. By multiplying up

the Core Service spending on a given disease by the appropriate proportion from 2012/13 local PB spending, this accounts proportionally for local Specialised Services spend. For example, the Specialised Services budget for osteoarthritis of the hip in 2012/13 equated to 0.13 of the Core Services spending, so 2018/19 Specialised Services spending on osteoarthritis of the hip was 0.13 * £424m = £56m. For cancers, the 2012/13 local PB breakdowns don't allocate to cancer subtype, so the admission ratio across all cancers is used for each specific cancer cost calculation.(47,156)

Primary Care services (GP practice services) are commissioned centrally with a budget of £11.8bn. This budget is not broken down by PB category, but the amounts spent on prescribing in Primary Care are published by PB category. By calculating the proportion of total Primary Care prescribing spending to each PB category (the "prescribing ratio") and assuming that total Primary Care spending mirrors the proportions spent on Primary Care prescribing, total Primary Care spending on a PB category can be estimated as the total Primary Care budget multiplied by the prescribing ratio for that PB category. To disaggregate from PB-level spending to disease-level, the admissions ratios used for Core Services were again applied. For example, spending on PB category 15X made up 0.036 of all Primary Care prescribing, and the admissions ratio was 0.085, so the total Primary Care spend on 15X (osteoarthritis of the hip) is estimated as 0.036 * 0.085 * £11.8bn = £36m.(47,156)

Spending on each disease on Core Services, Specialised Services and Primary care were then combined, and finally divided through by the number of cases from the GBD 2019 to give a cost per case per year.

Estimating costs of non-modelled diseases

Costs to the system per person per year of non-modelled diseases were also calculated in this previous work. These represent the average costs to the NHS of all other conditions apart from those modelled. This can allow the option of accounting for future unrelated medical costs in HEE, if desired (whether or not they can and should be included is debated,(160,161) though usually they are not(162)). To do this, the aggregate costs of each disease across Core Services, Specialised Services and Primary care were combined and netted off the total NHS budget (excluding PB category budgets for healthy individuals, social care and other) and divided though by the population.

Aim

For local modelling, the aim was to adapt the approach described above, taking local-level inputs (where feasible) to estimate local-level healthcare costs. The potential approaches for doing this were explored, starting with understanding the underlying NHS cost data, which is complex and of variable quality.

NHS cost data literature review

It is helpful to understand the context of NHS cost data before using them, as their quality is known to be variable and the way they are collected has been shaped in part by politics. This literature review will explore the history and evolution of these data, and provide a critique. The remainder of this introduction is dedicated to this literature, which informed

the selection and interpretation of cost data for PRIMEtime_local, before returning to the methods for how these data were used to estimate local healthcare costs.

Background to NHS costing

The collection of NHS cost data has slowly evolved over time. Accurate cost data was not historically estimated or collected in the NHS as there is no well-developed private insurance market in the UK driving a need for it, unlike countries such as Australia. This resulted in weak relationships between need, activity and income.(163) Through the 1990s, increasing focus on public service efficiency led to the realisation that NHS provider costs and efficiency varied widely. Levers were put in place to attempt to reward efficiency for both purchasers and providers in the internal market (a landmark healthcare reform creating competition between providers in the NHS). Neither reported costs nor variation reduced(163,164) and by the end of 1997, the new government agreed these levers had failed to drive efficiency, and instead that perverse incentives were leading to fragmentation and unfairness.(163,165) While much of the infrastructure of the internal market was maintained after 1997, the old measure of efficiency, the 'Purchaser Efficiency Index' was abolished(164) and new approaches aimed at rewarding performance and collaboration were developed.(165) 'Reference Cost' collections began from 1997/98,(164) intending to gain an understanding of baseline efficiency and track improvement (or lack of).(165) Trusts (NHS providers) were mandated to publish their costs to enable comparison. The NHS Reference Costs Index (RCI) was developed, representing organisation-wide average costs, with similar services grouped as Healthcare Resource Groups (HRGs), relating to diagnosis, treatment and cost implication. The RCI represents comparative costs, with

100 representing the average, 90 representing 10% lower costs than average and 110 as 10% higher.(166,167)

Reference Costs

What Reference Costs are

Every year, all NHS trusts in England must submit estimates to the 'National Cost Collection' (NCC) of what spending has been undertaken, related to what healthcare activity. These were traditionally done as Reference Costs, which NHS England defined as the "average unit cost to the NHS of providing defined services to NHS patients in England in a given financial year". Although this definition still applies to newer methods of cost collection, the term itself is generally no longer applied to the dataset, instead being referred to as the NCC Index. Average NHS costs continue to be calculated annually, producing a database of local NHS cost data. This is the richest source available on spending and variations in unit costs across the NHS in England.(166,167) A Reference Cost is calculated for each HRG code, defined by the National Casemix Office, defined by ICD-10 code (for diagnosis and other clinical features) and OPCS-4 code (Classification of Interventions and Procedures, version 4, for medical procedures).(166,167) Total and unit costs can then be assessed.

Reference Costs are used as inputs to many areas of NHS financing such as planning national and local budgets, the National Tariffs, commissioning, adjusting local pricing, and informing business cases for service development,(168,169) increasing transparency to the public and parliament, academic research and ONS economic statistics.(170–173)

As mentioned, one of the main purposes of Reference Costs is to compare the apparent efficiency of providers. It is not valid to directly compare crude cost per case between trusts as casemix (differences in diagnosis, approach and clinical complexity) will drive warranted variation.(166,167) Other justifiable variation may arise from differences in resources, technology, input costs, difference in priority, service quality or outcomes. Accounting practices and random volatility may also contribute to variation.(164) To improve comparability, reported costs are aggregated as service-casemix categories called "currencies". In England, HRG system groups currencies as "chapters" (specialties) for hospital care (broken into Admitted Patient Care (APC), outpatient care (OPD), and accident and emergency (A&E) care).(166,167)

How Reference Cost submissions are collected

Traditional Reference Costs have been built top-down at the trust level on a full absorption basis and associated with the activity that generated it, that is, collectively exhaustive of spending on a patient activity and avoiding cross-subsidisation of services.(166,167,173) Budgets spent on total direct costs (the requirements of patient care itself eg. front-line clinical staff), indirect costs (necessary to patient care, but not easily attributable to a given patient, eg. power or bed linen) and overheads (organisational costs not easily attributable to a single patient, such as the HR department) are summarised,(163,166,167) then amounts of each budget spent on each service line (specialty) disaggregated via different methods. 'Actual use' costs are those that can be linked to specific activities, such as surgical devices with fixed costs. 'Weighted use' includes factors that require some clinical judgement, such as how much nursing time may be required on average for a group of

patients, estimated as a factor of patient need and the number of nurses on a ward. 'Apportioned costs' are those that are divided though crudely, such as the total lighting bill being divided through by the floorspace of each unit.(174) The Healthcare Financial Management Association (HFMA, the professional organisation for UK healthcare finance) publish detailed guidance that is approved by NHS England(174,175) on how to approach these for different types of unit (eg. medical wards, theatres, mental health units). Each Trust aggregates its unit costs to nationally-standardised service lines (specialties), then divides that by the appropriate number of bed-days, giving a crude cost per bed per day, by specialty. A 'care profile' is then often used to estimate HRG-specific costs from specialty averages, accounting for factors such as staff time and investigations, supported by a series of detailed workbooks.(163,166,167)

Critique of Reference Costs

Problems with the first Reference Costs database quickly became apparent. For example, it was observed in the 2000 dataset that there were gross orders of magnitude difference between providers, for example variation of 20 times between most and least efficient providers for a hip replacement, or if the lowest cost submitted (£480) could even pay for the device alone.(176) Limitations of traditional Reference Costs include variable interpretation of the cost collection guidance, substantial exclusions of certain sources of costs leading to incomplete data, and the aggregation of costs to HRG averages.(177) Concerns over data quality continued, so in 2009 Monitor (the regulator of NHS trusts financial performance, now part of NHS Improvement(178)) commissioned a major investigation into NHS costing that was published by Lord Carter in 2016.(179) This was

critical of Reference Cost data quality, noting “huge inconsistency in costing and budgeting approaches … impairing our ability to compare data across the service” (page 72) and making a series of recommendations.

Alongside the Carter Review, work on Reference Cost data quality was commissioned to Capita and Deloitte. The report from Capita in 2014 audited 50 acute trusts and found financially *at-risk* trusts were much more likely than others to have inaccurate data submissions: 47% of at-risk trusts had overall low-quality submissions compared to 20% of randomly audited trusts. Basic errors were being made such as the inappropriate inclusion or exclusion of activities, double counting impairments, weak internal processes, and inaccurate input data. Very few trusts were found to even have appropriately accounted for nursing need. They concluded guidance had improved, but it was often not being followed.(180)

Deloitte undertook a more overarching approach to assessing Reference Cost quality across 178 trusts (published 2014). They found that one in eight trusts’ submissions 2010-13 contained materially-important errors, for example with approximately 50% of HRGs aggregate submissions being non-normally distributed, and 25 trusts had >50% of their HRG submissions outside the range of 50-150% of mean for that HRG. Duplicate costs were a common issue (ie. exactly the same unit cost given for multiple HRGs) though trusts reporting >40% duplicate costs dropped from 75 to 40 across the three years (of over 150 trusts). They felt excessive use of judgement rather than following guidance was leading to erratic reporting.(181)

Patient-level costing

The move to patient-level costing

Prior to the introduction of Reference Cost collections, through the 1990s an interest arose in understanding the complexity of cost and quality of healthcare services, accepting that variation was lower than in many industries and that aiming to drive down cost per unit was simplistic.(164)

Patient-level costing was seen as a potential route to develop this understanding. Calls for using patient-level data for costing had started even by the time the first Reference Costs were being collected in 1997/98(163,177,179,182) but problems such as the heterogeneity of systems, difficulty in unifying guidelines, reconciling the submissions and supporting trusts to implement systems all took time.(182)

Monitor proposed mandatory patient-level costing submissions (rather than traditional top-down Reference Costing) in 2012(168) and the 5YFV 2014(22) described at length the desire to better understand cost drivers. In the context of a frozen cash-terms budget against rising demand, the 5YFV identified a forecast shortfall in NHS budget of almost £30 billion/year in 2020/21, so an aim was set to improve efficiency – with accurate costing seen as an essential part.(22,177)

The Care and Quality Commission (regulator of healthcare and social care providers in England) applied further pressure on providers to improve their submissions with its report The State of Health Care and Adult Social Care in England 2014/15. This asserted that producing accurate cost data was a quality benchmark of competent providers.(177,183)

In common with Lord Carter, making patient-level reporting mandatory was a key recommendation from both the Capita and Deloitte reports.(179–181) Monitor's Costing

Transformation Plan 2015(184,185) started the process of reforming costing, including new reporting standards, software requirements and engaging providers.(168,184) Even by the time of this report, in the NCC round 2015/16, 149 of the 237 submissions included some degree of patient-level costing.(166,167) Mandatory patient-level submissions started for acute services in 2018/19 then were rolled out across other Core Services.(177)

Cost data quality have gradually improved due to the Carter Report,(179) ongoing work by government and the support of the HFMA.(175) Improvements have been related to developments such as self-assessment checklists, annual updates on standards and targeted external assurance process.(166,167)

What patient-level costing is

Patient-level costing is a different approach to estimating activity costs in the NHS. Rather than gradually disaggregating budgets top-down, costs for each patient are estimated based on their individual consumption then those patient pathway costs aggregated to a given level.(168,182) There has been a longstanding ambition to move from top-down costing to patient-level bottom-up costing on the basis of allowing a better understanding of cost drivers.(182)

In the NHS, the systems used to calculate these is referred to as the Patient-Level Indicative Costing System (PLICS).(177) NHS England defines PLICS(166) as the "IT systems which combine activity, financial and operational data to cost individual episodes of patient care … where an organisation records individual interactions and events that are connected with a patient's care" (page 16). As with Reference Costs, PLICS takes a full absorption approach of all required inputs and overheads from admission to discharge. PLICS standards(177)

"require providers to capture better and more accurate cost information at each stage of a patient's journey. The data should accurately reflect the 'causality of costs' in the system; tracing why costs are being incurred, who is incurring them, by doing what type of activity and for which patient" (page 7).

To support the consistency that makes PLICS valuable, NHS England have published extensive guidance covering each type of service provided.(182,185–187) Five Standards act as the broad framework. They start with 1, ensuring the *General Ledger* (outlay) is defined appropriately, then 2, aggregating outlays into predefined clinically-relevant categories (the *Cost Ledger*), for example accounting wages, National Insurance and pensions as 'staff costs' then disaggregating them by role. Standard 3 details how to map Cost Ledger as direct costs or overheads and link these *resources* (eg. staff, equipment, consumables) to specific cost-generating *activities* (eg. a particular operation or pathology test) using *relative weight values* (eg. what proportion of a member of staff's time is spent on given work). Standard 4, assigning these resources to specific patients (*matching*) is a critical step,(174,188) that is, the clinically-appropriate association of costs to patient episodes, relating to the causality of costs. This is informed by extensive rules and data known as 'feeds' that quantify how much resource is used by patient episode. The aggregated costs then must be reconciled against accounts to check they remain consistent.(174,175,187,188)

Potential benefits and risks from PLICS implementation

Groundwork by NHS Improvement in 2016 identified some trusts that were already implementing PLICS had achieved large benefits from small cost investments.(177) Examples included identification of mis-coded procedures leading to lost income of £160,000 per year,

omission of claims for Cancer Drugs Fund income, and more efficient running of ophthalmology theatres to the order of £211,000 per year. Potential benefits included increased detail, improved comparability, identification of areas of inefficiency, increased long-term stability of providers, and informing local commissioning and tariff variation. It was felt other non-cost benefits were also achieved, such as increased clinician engagement with finances and cost implications of different choices.

There are also potential downsides. The use of PLICS increased the risk of inaccurate reporting. In Capita's 2014 report mentioned above, PLICS were felt to add complexity and bureaucratic burden. It was realistic that "Non-admitted patient care services still present the same challenges to cost with or without PLICS".(180) It was also acknowledged that larger, broader and more complex providers with better upstream cost systems would find it more expensive to implement PLICS and that any successful implementation relied on good senior leadership and buy-in across management and clinicians, not only for implementation but for the ongoing embedding of cost data in decision-making, requiring a complex managerial, financial and clinical skill-mix that may be difficult to access.(177,185)

Other considerations on NHS cost data

Budget allocation formulae

The geographic distribution of NHS resources is determined by a series of slightly varying funding formulae (for relative need, not absolute), written by the Advisory Committee on Resource Allocation (ACRA), last updated in 2019. The formulae intend to provide fair funding, in terms of following need and allowing the same standard of care to be provided regardless of input costs.(189)

A different formula governs each funding stream. For Core Services, the Nuffield Trust designed a formula with five factors: 1, age-related need; 2, non-age-related need; 3, unmet need and health inequalities (Standardised Mortality Rate <75 years); 4, input cost variation (using the Market Forces Factor (MFF), discussed below); and 5, additional costs of providing services to sparse or unavoidably small populations (for small hospital providers, core services and ambulance budgets only). How these factors are weighted is shown in figure 7.1.(189)

The Primary Care allocations formula was developed by ACRA, based heavily on the previous Carr-Hill formula. This is based on medical record opening times (from the Clinical Practice Research Datalink, CPRD), adjusted for age-sex group, new registrations and IMD.(189)

As mentioned, Specialised Services were commissioned centrally after 2014 rather than resources being transferred to CCGs. The dataset detailing specialist services usage (the Prescribed Specialised Services dataset) had 54% missing data, rendering it unusable for determining how much resource was required for some services by ACRA. Therefore, the previous historic pattern was used to set these services' budgets instead.(189)

The unmet need budget adjustment required judgement to be applied by ACRA as nationally-comparable data on need are not available. The consequent budget allocations exponentially increase across the range of unmet need, with Core Service CCG budgets up-weighted by up to 10%, Primary care 15% and Specialised Services 5%.(189) Weightings are applied to account for the supply-side costs of providing services (the MFF) and the added costs to sparsely-populated areas, with neither specifically relating to expected demand.(189,190)

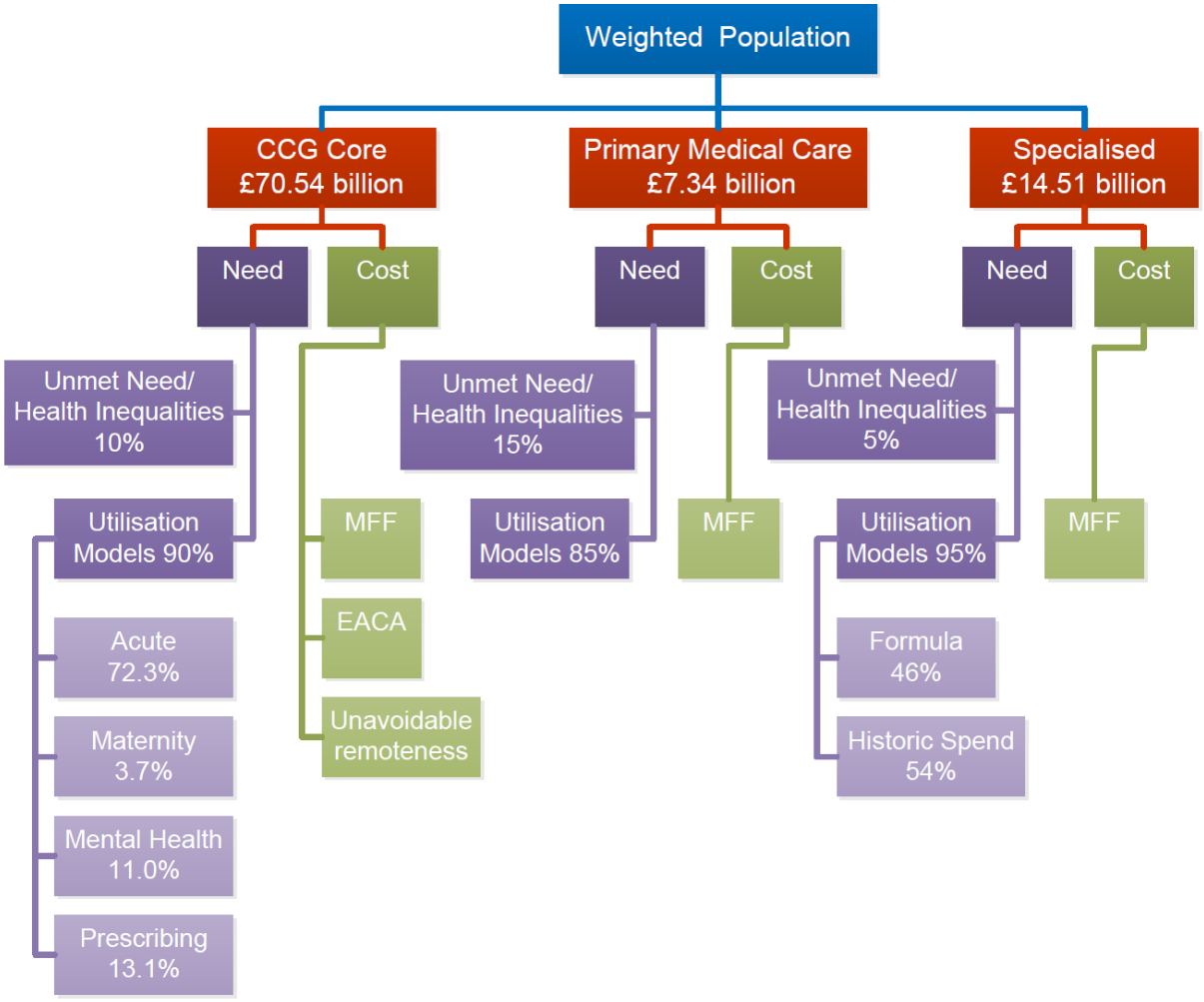


Figure 7.1: Summary of the CCG funding formula and adjustments, from NHS England Allocations and pace of change 2016/17-2020/21(189) reproduced under the Open Government License. MFF = Market Forces Factor, EACA = emergency ambulance cost adjustment.

The Market Forces Factor

The MFF is an index published by NHS England intended to estimate "the unavoidable cost differences between healthcare providers".(190) NHS resource allocations are adjusted(189,191) to account for these inherent differences "so that patients are neither

advantaged nor disadvantaged by the relative level of unavoidable costs in different parts of the country" and "to ensure they can afford the same level of services, relative to need" despite varying underlying costs (page 2).(190) The MFF is usually based with a minimum of one, with decimal increases representing proportional percent increases, ie. a MFF of 1.02 represents an area with 2% higher costs than the lowest cost area.(190)

NHS England defines "unavoidable costs" as those that providers cannot "influence significantly", grouped into five categories: non-medical and dental staff, medical and dental staff, land, buildings, business rates and other. Each of these is calculated independently then a weighting is applied to combine them, shown in figure 7.2.(190)

Non-medical and dental staff costs are calculated at the travel-to-work-area level and based on those in the private sector, on the basis that staff turnover and agency worker use follow local private sector wages (not that their pay varies particularly). The 'other' category includes all other, non-varying, costs, so that the index represents all costs of care so can act as a valid multiplier on prices.(189,190)

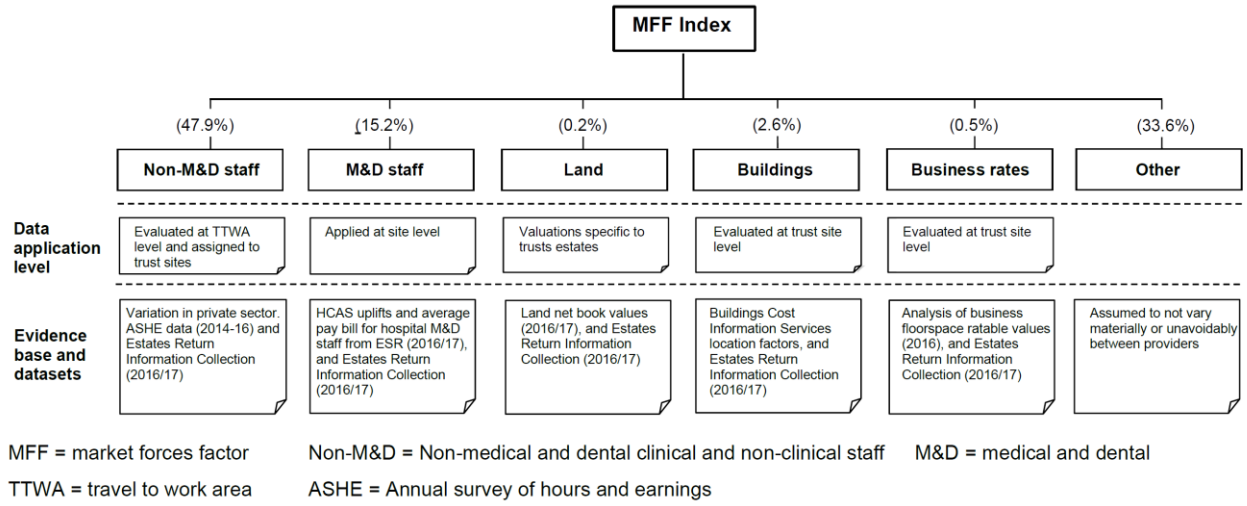


Figure 7.2: Composition of the MFF. From Consultation on 2021/22 National Tariff Payment System Guide to the Market Forces Factor, Appendix A, NHS England and NHS Improvement, reproduced under the Open Government License.(190)

Methods

Overview of the approach for local level disease costs:

A method was developed based on that described by Briggs(156), also applied by Cobiac *et al*,(47) yet differing, by using local-level data inputs where feasible. Hospital costs, Primary Care costs and average costs of remaining non-modelled diseases are estimated separately. The approach is explained in full below, but in summary:

Hospital costs are estimated by varying the national average disease costs by a factor accounting for local variation. This variation is estimated from PLICS submissions to the National Cost Collection. Primary Care costs were calculated by starting with Primary Care budgets at the CCG level. These were disaggregated to specialty-level by applying the local

'prescribing ratio' to disaggregate from total spending to specialty, then admissions ratio to disaggregate from specialty to disease. Hospital and Primary Care costs were added together then divided through by the number of local cases.

Costs of non-modelled diseases were estimated by taking published NHS Core Budgets and Primary Care budgets for each CCG area, multiplying up Core Budgets to account for average Specialised Services spending, then subtracting baseline spending on modelled diseases before dividing through by local population.

Approach for hospital costs

The national-level mean hospital costs used previously in PRIMEtime by Cobiac *et al*(47) were kindly provided by Dr Cherry Law at the London School of Hygiene and Tropical Medicine along with the spreadsheets used in their development that include costs of hospital services per disease. These followed the method outlined in the introduction to this chapter.

PLICS submissions 2019-20(186) on Emergency, APC and OPD at the trust level were used (from NHS Digital). These provide the spending for all activity done that year and the activity count (the number of times a patient with a given condition and treatment was processed) in each disease group for each trust. As these data include both Core Services and Specialised Services activity, it was possible to apply this method to account for both. From these data, a cost per activity for each specialty for each trust was calculated, averaged across Emergency, APC and Outpatients, weighted for activity count. This cost per activity could then be used to produce an index representing the between-trust variation in cost-per-activity, capturing variation such as the unit's efficiency, technology costs, case mix and

other factors. This index had a mean of 1 with 1.1 representing 10% higher costs than average and 0.9 representing 10% lower. Each disease in the model was categorised by specialty and that specialty's trust-level index applied to the national average hospital cost from Cobiac *et al*,(47) producing an estimate of cost per case per year by trust. For example, Manchester University Foundation Trust cardiology had a weighted average cost per activity across Emergency, Admitted Patient Care and Outpatients of £2,559.50, compared to a national average of £1,839.20, giving a cardiology cost index of 1.392. The national average hospital services cost of IHD in Cobiac *et al* was £416.74 giving a local cost estimate of £416.74 * 1.39 = £579.34.

To move from the trust level to the local authority level, a mapping file(192) was used that provided a proportion of each local authority's COVID-19 admissions at each trust. The COVID-19 admissions data were released by NHS Digital to describe the relationship between local authority of residence and hospital trust of admission, for all COVID-19 admissions in England. These were produced from HES,(159) including all patients with non-missing variables for residence and a confirmed diagnosis of COVID-19. This acts as a proxy for other hospital work, as the distances required to travel for COVID-19 acute care are likely to be similar to other hospital care. This provided a method of quantifying the geographic spread of clinical work for each trust. For a given local authority, a weighted average between trust-specific disease costs was then estimated, with the proportion of each local authority's COVID-19 work to each trust acting as the weight. For example, in Manchester local authority, 76.2% of work is allocated to Manchester University Foundation Trust at £579.34 per case of IHD, 22.7%% to Salford Royal Foundation Trust (at £310.46 per case) and 1.1% to Christie Foundation Trust (£315.72 per case), giving a weighted average cost of £515.87 per IHD case per year in Manchester. Of the 145 acute trusts (excluding

trusts providing only ambulance, community care, children's services and other non-modelled specialist services), eight did not provide relevant PLICS submissions, so proportions of activity were re-calculated including only trusts that did respond. Minimum and maximum costs were set as no more extreme than the exponent of 2 SDs of log-costs for each disease. For example, Harrow local authority had costs of £3,453 per incident case of osteoarthritis of the knee, but this was capped at £3,293.

Primary care costs

Primary care costs were calculated by extending the method described in Briggs *et al*(156) and used in Cobiac *et al*,(47) but using local-level data inputs where available. First, it is possible to calculate total specialty-level primary care prescribing costs in each CCG from Programme Budgeting net expenditure 2018-19 then calculate the proportion of prescribing spending going to each speciality (with specialty equated to Programme Budgeting category). Assuming that total Primary Care spending reflects primary care prescribing spending, total CCG-level Primary Care budgets were multiplied by prescribing ratios to estimate CCG-level Primary Care spending on each Programme Budgeting category. An assumption was made that the proportions of patients admitted to hospital are reflective of the proportions of spending in Primary Care. This 'admissions ratio' used here were taken from the calculations undertaken by Cobiac *et al*,(47) representing the proportion of each Programme Budgeting category's cases in the HES 2018-19 (Admitted Patient Care)(159) with ICD-10 diagnostic codes relating to the given disease. This process produced an aggregate total cost for each disease in each CCG.

Total costs of treating each disease were then translated from CCG to local authority level using a mapping file reporting how many LSOAs in each local authority shared with each CCG. This allowed total budgets to be apportioned to local authorities in line with the proportion of LSOAs sitting in each local authority. For example, Devon local authority has 69.4% of the LSOAs paid for by NHS Northern, Eastern and Western Devon CCG whose total Primary Care cost of treating IHD was £2,875,831. It also is covered by 50.8% of the LSOAs under NHS South Devon and Torbay CCG, with a total Primary Care cost of IHD of £906,466. This gives a weighted total cost of (0.694 * £2,875,831) + (0.508 * £906,466) = £1,995,827 + £460,485 = £2,456,312 spent on IHD by Primary Care services in the Devon local authority in 2018-19.

Total Annual Costs of Disease

Local hospital cost estimates and local Primary Care cost estimates were combined. Cost per case (Annual Cost of Disease) was calculated by dividing aggregate costs by the number of treated cases. The number of treated cases was calculated as the incidence or prevalence of the given disease (whether the disease was treated as an incident or prevalent case) taken from the GBD,(2,97) multiplied by a factor for detection rates. This detection factor was calculated as the proportional difference between the national-level number of cases in HES (used in Cobiac *et al*)(47) and the number of cases in England's GBD data. This was calculated as 91-92% for all included diseases, so the mean of 91.5% (1 decimal place, d.p.) was used.

Whether diseases are treated as incident or prevalent cases depends on their clinical pattern of exerting costs. For some diseases such as IHD, ongoing costs are exerted such as

through angina, heart failure, recurrent acute events. Costs were therefore calculated as costs per prevalent case. Others, such as low back pain, occur more as acute episodes that resolve, so treated as incident. Some diseases, such as osteoarthritis of the hip or cancers, exert the majority of their costs in the form of single treatment event (eg, hip replacement, chemoradiotherapy) regardless of duration of disease, so can also have their costs better modelled as incident cases. Therefore, breast cancer, colorectal cancer, oesophageal cancer, low back pain, osteoarthritis hip and osteoarthritis knee were treated as incident cases, and IHD, stroke, T2DM and AF/ flutter were treated as prevalent cases with ongoing annual costs.

Costs of non-modelled diseases

CCG-level Primary Care and Core Service CCG budgets 2019-20 were taken from CCG Allocations,(193) (excluding costs included under Programme Budgeting categories 21, for healthy individuals; 22, for social care needs; and 23, other). Specialised Services are no longer included under Programme Budgeting. Nationally, Specialised Services equate to 15.5% of the value of total Core Service budgets, so assuming that this proportion is consistent across residents of local areas, these costs were accounted for by multiplying local Core Budgets up by an additional 15.5%. These hospital and Primary Care costs were combined, then mapped to local authority level using the same LSOA-level mapping file used for Primary Care costs. Costs were estimated from 2018-19 to keep consistency with the national disease-level costs from Cobiac *et al*(47) used for estimating hospital costs, so where CCG mergers had taken place between the publication of the local authority-to-CCG mapper in 2016-17 and the budgets in 2018-19, areas were aggregated. Weighted total

budgets coving local authority areas could then be estimated in the same manner as total Primary Care costs, described for Devon, above. These total local authority equivalent budgets were netted of the aggregate disease costs for the included diseases then the remainder (representing the total NHS spend on non-modelled diseases) was divided through by local authority populations. Minimum and maximum unrelated disease costs were capped to be no more extreme than the exponent of 2 SDs of log-costs for each disease.

For example, the Core Services allocation for NHS Barnet CCG was £351,690,000 and Primary Care allocation £53,220,000, plus an allocation for Specialised Services of £351,690,000 * 1.155, giving total allocations of £459,421,950. Barnet CCG LSOAs map onto Barnet local authority exactly so the local authority equivalent was the same. The sum of Primary Care and hospital services costs (Core Services and Specialised Services) for modelled diseases in Barnet was £58,456,650, leaving £400,965,300 spend on non-modelled diseases. Dividing by the population in 2018-19 of 392,139,(109) this estimates unrelated disease costs of £1051.09 per person.

Results

The method produced a mean cost of managing each of the modelled diseases for each of the 150 upper tier local authorities (2018) per year. Costs were assumed to be the same for each lower tier local authority within an upper tier local authority. Table 7.1 provides median, mean, 25th and 75th centiles of upper tier local areas' Annual Costs of Disease for each modelled disease and combined non-modelled diseases (to 1 d.p.). Figure 7.3 shows these distributions as log-costs as they are very highly skewed and not easily legible on the costs scale. This shows there was wide variation in costs between diseases and between local authority areas. The lowest median cost of disease management per person per year was for hypertensive heart disease at £71.3 and the highest osteoarthritis of the hip at £13,684.8. On the log-costs scale, distributions were fairly consistent, with SDs of log costs 0.25-0.32 for most diseases – though on the costs scale spread is more variable, as identifiable in table 1, for example with AF and flutter having an interquartile range of £104 and osteoarthritis of the hip £3362.

Maps in figure 7.4 show the spread of two example disease states – the cost per prevalent case of T2DM and cost per incident case of osteoarthritis of the hip. The maps in figure 7.5 are included for context. These demonstrate no obvious geographic patterning or association of PLICS costs with local RCI, the MFF index or IMD score. Here, PLICS values are calculated as simple crude average across the 12 disease states. Though it is not directly relevant here, a brief quantitative analysis of these relationships is given in Appendix 1f, with graphing, showing a modest relationship between MFF and PLICS costs.

Table 7.1: Median, mean, 25th and 75th centiles of distribution of costs (£ per patient per year for diseases modelled as prevalent cases; £ per incident case per year for diseases modelled as incident cases) between upper tier local authority areas (to 1 d.p.).

Disease	Mean cost (£)	Median cost (£)	25th centile (£)	75th centile (£)
AF and flutter	196.1	171.7	148.4	221.2
Asthma	3610.6	3571.4	3249.3	3865.5
Breast cancer	12321.7	11429.5	9677.7	14340.4
Colorectal cancer	8938.7	7994.2	6812.4	10467.4
Oesophageal cancer	2351.0	2100.9	1791.8	2753.2
Hypertensive Heart Disease	71.3	65.0	57.1	83.9
Ischaemic Heart Disease	572.8	529.3	447.8	673.0
Low back pain	421.9	411.0	356.9	464.2
Osteoarthritis of the hip	13684.8	12554.9	10920.2	15325.2
Osteoarthritis of the knee	1765.4	1619.5	1408.6	1977.6
Stroke	1951.2	1789.2	1511.2	2205.5
T2DM	173.1	163.8	142.7	192.7
Non-modelled diseases	1161.5	1131.9	1080.0	1230.8

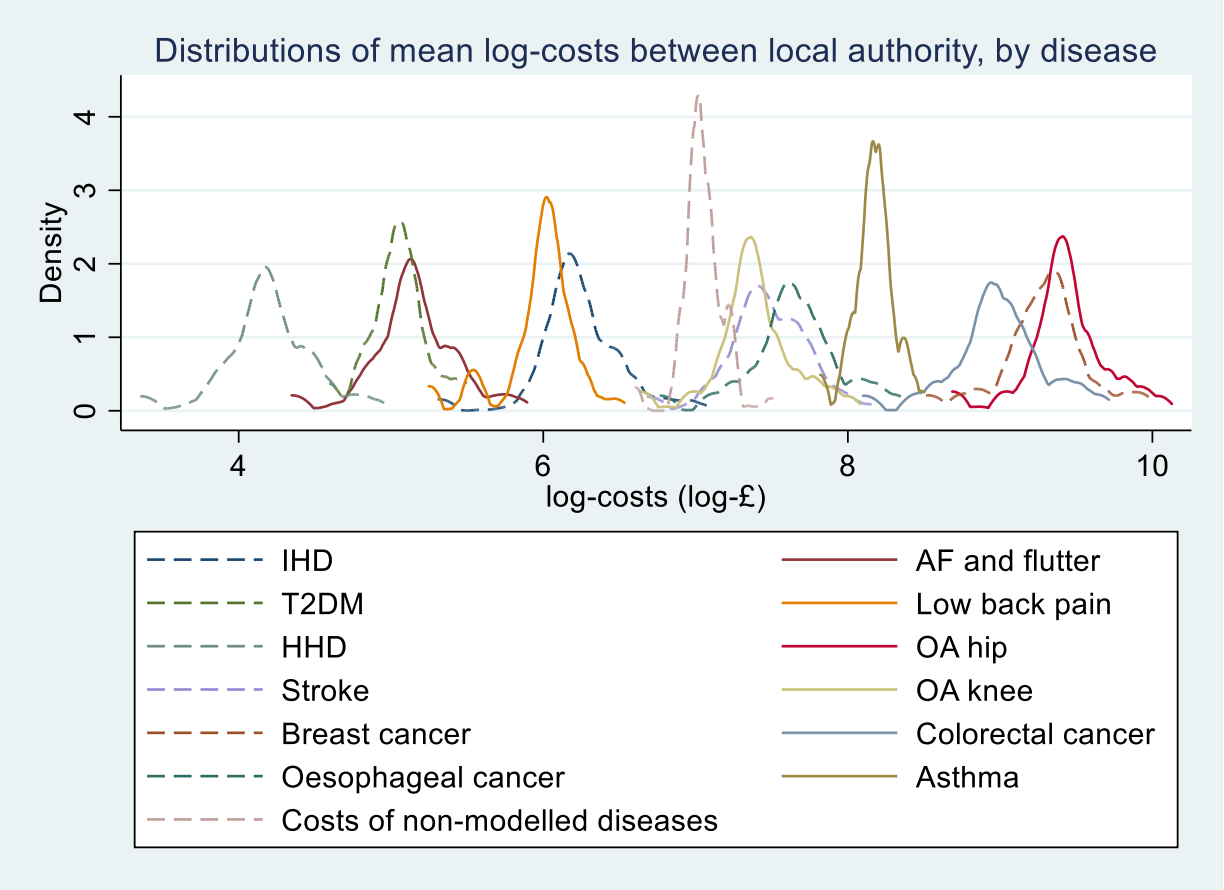


Figure 7.3: Distributions of annual costs of disease per person between local areas, as log-costs.

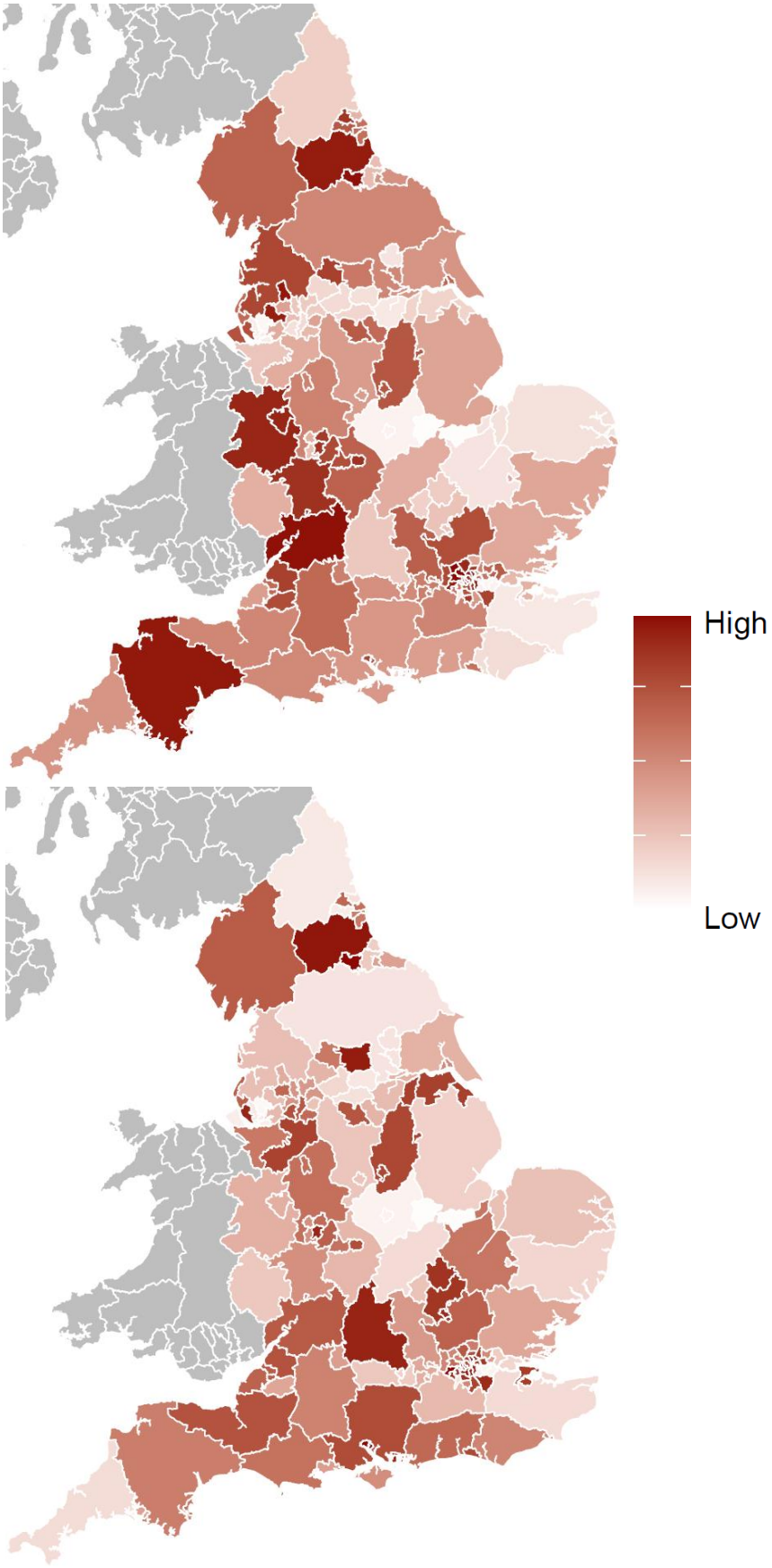


Figure 7.4: Example costs by local authority: T2DM (above) and osteoarthritis of the hip (below).

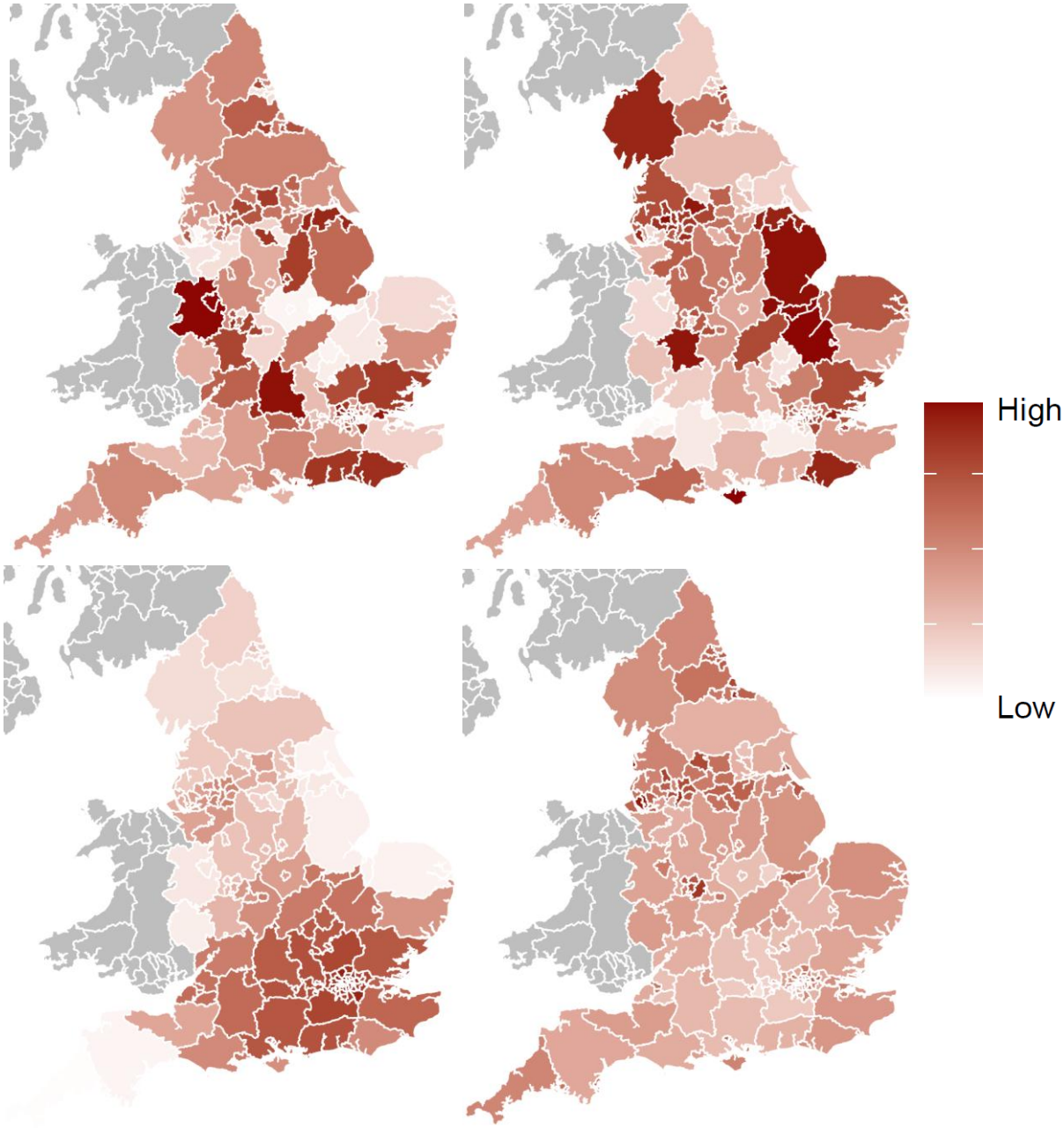


Figure 7.5: Variation in average costs by PLICS submissions (average between diseases) (top left), by Reference Costs Index (average between diseases) (top right), variation in Market Forces Factor score (bottom left) and Index of Multiple Deprivation score (bottom right).

Discussion

Summary of findings and interpretation

Costs of managing the modelled diseases per year varied from £161.0 per person per year, to £12383.5 and diseases with greater mean costs tended to have proportionally wider distributions. No obvious geographic patterning was identified to PLICS costs and they do not closely reflect Reference Costs by crude examination. Costs are somewhat associated with supply-side factors included in the MFF.

Being able to break down the average costs of providing healthcare by area of residency provides new ways of examining the drivers of costs by framing local areas as health geographies to allow differences in the types of care provided between Trusts to be smoothed out. This method allows non-geographic trust-level costs to be linked to local geography of residence, allowing local modelling to include for costs reflective of local variation. These data may also help local ICSs, which align closely to groups of local authorities, to understand their own differences in cost pressures.

Key assumptions

The method used a mapping step quantifying what proportion of a local authority's residents admitted with COVID-19 went to each trust. This relies on a simplifying assumption that COVID-19 admissions are reflective of broader healthcare provision, both inpatient and outpatient. It is possible that trusts with larger or more specialist respiratory centres would admit more COVID-19 patients, creating an undue appearance of accounting

for more of a local area's overall provision than they do in practice. This is a small risk as people and ambulance services generally attend the nearest acute centre. Likewise, provision of outpatient services tends to be closely related to inpatient services as GPs refer to the closest centre and consultants providing outpatient care also provide related inpatient services. It may be theoretically possible to use HES to link local authority of residency with each HRG chapter, to produce weighted average distributions of clinical activity by local area, but this process would be highly onerous and not feasible in this case.

The PLICS methods themselves are certainly a theoretical improvement on previous reference cost collection methods, but are as yet unproven and still have important assumptions and limitations to bear in mind, such as allocating overhead costs to units by floor area (rather than actual consumption) or local trusts following the costing methods guidance accurately.

For primary care costs, the mapping of costs between CCGs and local authorities by LSOA required an assumption that a CCG's LSOAs in different local authorities have the same average population size and composition. This is a fair assumption as LSOAs are a similar size by design, but there remains a risk that the LSOAs that do not conform to this pattern of CCG-local authority overlap do so for relevant reasons, ie. non-randomly. For example, they may have a more remote population than the main population centre of the local authority, which tend also to be older in England. Likewise, there is an assumption that the proportions of care provided by different trusts at the upper tier local authority level is the same in each lower tier, while in some areas they may vary, depending on factors such as geography, transport infrastructure and relationships between the locations of hospital services and populations, as well as demographic and social variation.

There is a judgement required to either model diseases as incident or prevalent, made on the basis of clinical pattern as explained above. In reality, each disease could have greater up-front costs or more ongoing costs and each lies on a spectrum. What is important for the modelling process is capturing what the cost of an individual developing a disease would be to the system, over the typical lifetime. Therefore, all diseases could be modelled as incident, with lifetime costs accruing at the point of the incident case, as the intervention does not vary case fatality so longevity from that disease (and lifetime costs) would remain the same. Treating all disease costs as incident could introduce a bias via changes to longevity related to the prevalence of other fatal diseases, as total population mortality changes, as well as biasing discounted costs incurred earlier rather than later (including indirect costs). Therefore, where prevalent disease costs are appropriate, they should still be preferred, as implemented here.

Each local area's Annual Costs of Disease were capped at +/- 2 SD of log-costs. Capping processes are standard(167,181) to account for the reasonable likelihood that very extreme cases are more likely to be related to reporting practices rather than genuine extreme differences. Consistency between total budgets, the costs of modelled diseases and non-modelled diseases was maintained by using capped costs to net out remaining costs before calculating average population costs to non-modelled diseases.

Strengths and limitations

Strengths

Broadly, top-down and bottom-up approaches have different strengths, with top-down methods being better suited to estimating national average costs and bottom-up methods being better suited for capturing difference between providers or over time.

This method used available inputs to translate good-quality reported costs of providing healthcare in local trusts to area of residency in England. This built on previous methods for estimating disease costs for PRIMEtime models for the local level. Distributing costs geographically in the way described helps to overcome the problem that the types of work done in each trust varies considerably. For example, some trusts have an emphasis on certain activity with greater or lesser costs, eg. cancer care, with varying exceptional personnel or technology costs, economies of scale and balance between Emergency, Admitted Patient Care and Outpatient care. Patient factors such as age and multimorbidity may also make care more or less expensive to provide. It is therefore not possible to directly compare the average cost per case in a specialty between trusts. This method helped to overcome this issue by distributing the caseload of trusts around their surrounding local geographies so that local authority-level costs are reflective of different amounts of provision from different trusts nearby.

Limitations

There are several important limitations to this approach.

The top down methods used require assumptions to be made to disaggregate budgets, so the cost of better capturing all costs is a lack of nuance and granularity between disease states or greater levels of specificity. For these assumptions to be correct also requires budgets to be spent on their intended services, ie. there is no cross-subsidisation of services. How much this is true may be very different across HRGs and individual providers.(185,194)

The admissions ratio is used to disaggregate costs from the HRG to disease level for hospital costs, ie. assuming that the proportion of total spending on a disease within a PB category is in line with the proportion of patients in that PB category admitted with that disease. This assumption is more stretched when it is used for the disaggregation of costs to the disease group in primary care, ie. assuming the proportion of work for each disease in primary care within a PB category is also in line with that of inpatient admissions. It is also a limitation that only national-level admissions ratios were available for use. Using HES to estimate these locally was not feasible and the option of using GBD epidemiology to estimate burden was not possible as the GBD disease groups aren't directly equivalent to PB categories.

A ratio was applied to account for the difference between the numbers of treated cases and the number of extant cases. Nationally, this was found to be consistently 91-92% with a crude mean of 91.5%. Again, this national figure is adequate, but it is likely that detection rates would vary locally, for example with people of lower SES being less likely to present to healthcare services, so in these areas the apparent cost per case may be higher than that estimated due to the lower denominator.

The local distribution of specialised service spending is very difficult to account for, as specialist centres tend to be in large urban teaching hospitals and may have large geographic footprints unalike the catchments for their secondary core services. Where costs

originate (trust spending) are not the same as where costs fall (national commissioning) or where patients live (remote from the trust). NHS England has not published breakdowns of their specialised services spending by PB category since 2012/13 and do not publish spending by trust. This introduces two problems. First, the distribution of specialised service costs across geography may not be fully smoothed out by the method described above using the COVID-19 admissions as the proxy for the distribution of other services. Likewise, distributing the ambulance service component of Core budgets is very challenging as there is no mapping method between ambulance trusts and local authority of residence, so the same distribution of costs was assumed as the rest of Core Services.

Conclusion

NHS cost data has progressively improved over the last three decades and is meeting new standards through mandatory PLICS submissions. These provide both greater accuracy in reflecting the origin of costs in the system and greater consistency in approach between providers.

Existing methods were updated to use local-level inputs where feasible and novel approaches to distributing provider-level costs to local areas was applied. Simple analysis did not identify an association between PLICS costs and previous Reference Cost submissions or deprivation, but there was some association with MFF. This provokes more questions about how cost variation arises, how to fairly reimburse providers and what strategies may be appropriate to improve efficiency. The use of good quality local data in

the form of the PLICS submissions in the methods provides confidence that between-area variation in healthcare cost estimates from PRIMEtime_local will be appropriate to reality.

The literature review section of this chapter is currently under peer review, as listed in Appendix 3.

Chapter 8: Modelling the potential health, cost and equity implications of increasing restrictions to television advertising of high fat, sugar and salt foods

Introduction

Obesity rates have been progressively rising in the UK(195) with particular concern around excess weight among children, with its associated lifelong disease burden. Up to the COVID-19 pandemic, inequalities in child obesity rates were increasing, with obesity rates increasing in the most deprived groups while they fell in the least deprived.(7,151) The pandemic has substantially worsened matters, with rates of obesity in reception class (aged 4-5 years) increasing by 4.5 percentage points on average between the 2019-20 and 2020-21 surveys – while dramatically increasing inequalities overall: an increase of 7 percentage points for the most deprived and just 1.8 percentage points for the least deprived. The increase in inequalities for Year 6 children (aged 10-11 years) was marginally smaller, at 6.3% and 2.4% respectively, though partly as rates in the least deprived groups increased faster in year 6 than in Reception class.(7,151) For Reception class, this overall rate of increase was over 25 times higher than the average seen over the five years since the previous nadir in 2014-15 and for Year 6 it was over 11 times.(152)

Recent UK governments have published a series of strategies to address obesity for both children and adults, most recently “Tackling obesity: empowering adults and children to live

healthier lives”(11) in 2020, following “Childhood obesity: a plan for action, Chapter 2”(12) in 2018, specifically aimed at children. In spite of the rapid deterioration in recent years, in June 2021 the UK government reiterated its target to halve childhood obesity rates by 2030.(196)

Chapter 2 included an announcement to consult on further restrictions to the television advertising of HFSS products to children, specifically aimed at those aged 5-17 years. This would bar the advertising of products meeting certain nutritional standards from the Food Standards Agency 2004-05 Nutrient Profile Model(197) from being advertised on television from 5.30am-9pm. This proposal is built on the evidence that advertising of food products to children changes preferences and has an impact on the quantity of advertised products consumed.(198) This follows earlier (2017) bans on the advertising of HFSS products on media targeted specifically at children both on broadcast (television and radio) and broader advertising guidelines not to promote unhealthy lifestyles or poor diet to children, or the use of ‘pester power’, irrespective of time of day.(12) However, children remain exposed to 2.9bn HFSS advert views on television per year as of 2019 (and a further 11bn online)(199,200) because the previous regulation leaves a large gap in limits to HFSS advertising, as 50% of all children’s viewing is commercial programming aimed at adults, peaking at 6-9pm, as shown in figure 8.1, below.(12) It is also known that children in more deprived areas consume more HFSS products are more likely to be overweight/ obese, and watch more television than children in less deprived areas, indicating a potentially useful route to addressing obesity-related health inequalities.(199–201) This situation comprises the business-as-usual scenario.

These advertising restrictions were planned to be implemented in January 2023 after the ban on sales promotions for HFSS products in October 2022. However, with increased inflation in 2022, particularly related to food prices and particularly for the poorest, the government delayed implementation of these proposals for 12 months, at the risk they prove pro-inflationary for those who could least afford it.(202)

Previous work by Mytton *et al* estimated the impact of these proposed television advertising restrictions on obesity rates and disease burden at the national level, concluding that if all target advertising was withdrawn then children would consume an average of 9.1kcal/ day less, leading to 40,000 fewer obese 5-17 year-olds (equating to 4.6% of baseline of that group) and averting 240,000 DALYs. Effects on obesity rates varied by social grade, with 2.5 times greater impact for the lowest social grades (DE) than for the highest (AB), but it was not possible to model in health and healthcare cost consequences on inequalities.(203) Alongside these benefits, this intervention may have cost impacts for the government and private sector. The government's Impact Assessment(198) estimated the government would face £1m setup cost for enforcement plus £9m/ 25 years maintenance (specifically for OfCom, the broadcast media regulator, to police adherence), while industry would be exposed to a £1m setup then £5m (£3-7m)/ 25 years ongoing for monitoring, plus £518m/ 25 years lost in returns. Their health assessment through the Calorie Model (Department of Health and Social Care) estimated £18m would be saved to the NHS over 100 years, £15m to social care, £750m in health saved (valuing 1 QALY at £60,000 as recommended in the Treasury Green Book(204)) and £44m of economic output due to lower mortality.

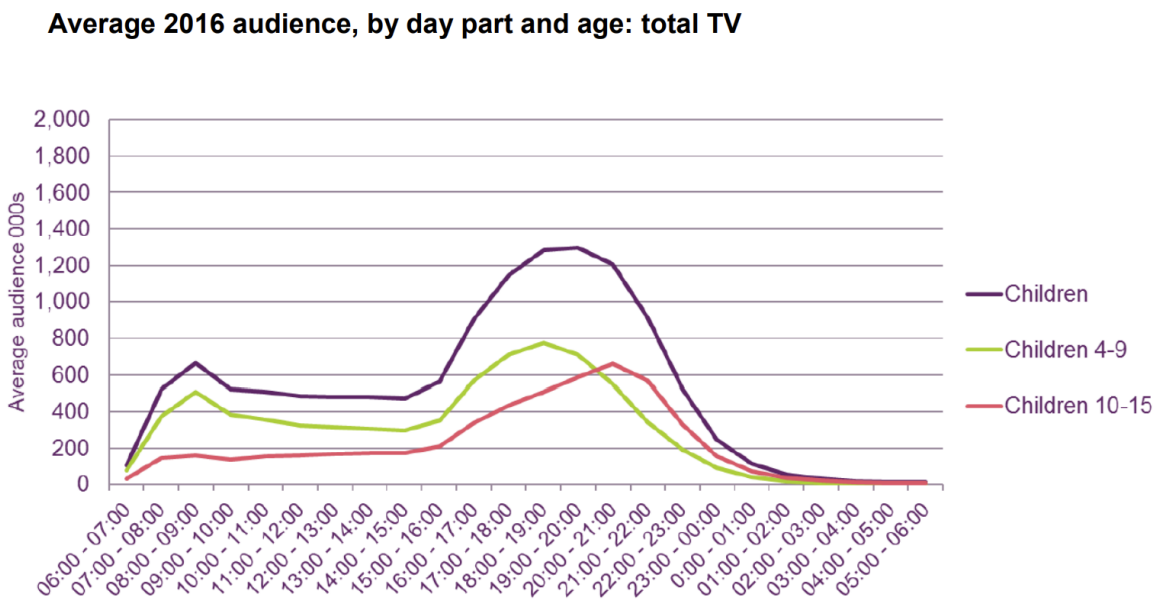


Figure 8.1: Average audience size over time of day, by age group. From Broadcast Audience Research Board 2016, cited from OfCom Children and Parents: media use and attitudes report, 2017. From Childhood obesity: a plan for action, Chapter 2,(12) reproduced under the Open Government License

A wide variety of other programmes were proposed in Childhood obesity: a plan for action, Chapter 2 were.(12) Tackling obesity (2020) advanced some of these proposals and a range of others aimed at adults.(11) It is important to consider the effectiveness, efficiency and equity of each of these interventions so we can implement these complex policy programmes with a reasonable idea of how likely they are to achieve their stated aims. In this Chapter, the effectiveness, cost-effectiveness and equity of restrictions on television advertising restrictions is modelled at the local level on disease burden and healthcare costs. This follows a previous paper(203) that estimated effects of the proposed additional restrictions on overweight and obesity rates by household SES, but health and cost implications were not estimated as appropriate modelling capabilities were not available. The alternative of evaluating potential restrictions on outdoor advertising, as implemented on the London public transport system,(205) was not pursued as the specifics of each local

area's geography and distribution of outdoor advertising space would involve a much greater level of geographic modelling that would not be feasible for this chapter.

Methods

Locally-specific effects of television advertising restrictions

Overview

The intervention was the banning of all HFSS product advertising on broadcast media between the hours of 5.30am to 9pm, against the business-as-usual scenario of advertising ban on broadcast media that is aimed primarily at children. The effect of these additional restrictions was estimated for children aged 5-17 years, consistent with the intervention being targeted at children, the evidence for impacts of advertising on food consumption and the validated population for the Hall weight change equations(93) (described in Chapter 3). A locally-specific effect for each local authority was estimated based on the approach reported in Mytton *et al*.(203)

The approach by Mytton *et al* can be summarised with the equation:

$$\bar{x} = \sum g_i t a_i e_w$$

where $\bar{x}$ is the estimate of the effect, i is social grade (in three groups AB, C, and DE), g is the proportion of people in that grade, a is the number of adverts seen per day per child in that social group,(206) t is time in minutes of each advert(207) and e_w is change to energy consumption per minute of adverts watched (estimated for each area, as explained below). This equation was also used to estimate uncertainty on each local area's $\bar{x}$ using the Monte Carlo analysis: taking uncertainty from original sources for time(207) and energy,(208) and could be estimated for proportions of social grade (described below as part of the Monte Carlo analysis), but no estimate of uncertainty was available for the number of adverts seen.(206) These social grade groups were taken from the original paper.(206) Social grades AB are is defined as 'managerial and professional occupations', C as 'intermediate occupations', 'small employers and own account workers', and DE as 'lower supervisory and technical occupations' and semi-routine occupations'.(209) The value for advert duration was taken from Mytton *et al*(203) and values for the impact on energy consumption per minute of adverts and the number of adverts seen per day were re-calculated for each local authority separately as follows:

Energy change per minute of HFSS adverts

Meta-analysis by Russel *et al*(208) found that the impact of advertising on energy consumption was greater for children with obesity than for children with a healthy weight (28.5kcal increase to consumption per day, per minute of HFSS advert viewing per day (SE 1.30) and 18.2kcal/day /min (SE 0.676kcal) respectively), so change to energy consumption for the Monte Carlo analysis was calculated separately as a weighted average of these values, based on the proportions of children of healthy and above-healthy weight in each

local authority (taken from those estimated in Chapter 6 above). The SE of the mean was used to create uncertainty on these weighted calorie figures, calculated from the data in Russel *et al*.(208) Random draw from these values was used to produce 3000 uncertain weighted averages for the Monte Carlo analysis.

Adverts seen per day

The number of adverts seen per day was also estimated as a weighted average in each local authority, based on the proportions of children in households of each social grade (determined as living in households whose survey reference person was in the given social grade). These proportions were estimated from the Active Lives Survey, aggregating waves 3, 4 and 5 (2017-18, 2018-19 and 2019-20) to increase numbers in each local area. This was done by converting the variable for 7-level National Statistics Socioeconomic Status to social grade using the approach suggested by the ONS,(209) as level 1-2 = grade AB, levels 3, 4, 5 = grade C, and levels 6-7 = grades DE. This provided point estimates of proportions of people living in each area in grades AB, C and DE. Uncertainty on these proportions was estimated as the SE of the sample proportion (as $SE = \sqrt{(\hat{p}(1-\hat{p}))/n}$ where $\hat{p}$ is the sample proportion and n is sample size). This estimate applies to one proportion alone, with the remaining fixed at certain values, so to allow all three proportions to vary, they were estimated in turn.

1. First, the SE of the proportions in grades DE were estimated for each local authority in the ALS, then this was used to produce 3000 uncertainty estimates of the estimate of the proportion in each area using a random draw.
2. Then, 3000 proportions in grade C were calculated by subtracting each uncertainty estimate of DE and the certain estimate in group A from 1. A SE was then calculated

for each of these estimates and again 3000 uncertain values of the proportion in Grade C were drawn using a random draw, producing estimates consistent with their equivalent estimates for the proportion in grades DE.

3. Finally, the uncertainty estimates of groups DE and C were subtracted from 1 to give 3000 uncertainty estimates of the proportions in grade AB for each area, consistent with those for DE and C.

As the number of adverts seen per day varies by social grade,(206) these 3000 estimates could then be used to calculate uncertain weighted average numbers of adverts seen per day in each local authority.

Effect of advertising exposure on calorie consumption

Finally, 3000 random draws on the Monte Carlo analysis were run to produce uncertainty estimates of the absolute impact of HFSS advertising on child energy consumption. The same random draw uncertainty was applied to each area in each iteration, so that different random draws did not drive differences between areas.

Point estimates of disease case, QALY and healthcare cost outcomes

Point estimates of the intervention on change to QALYs and healthcare costs was modelled using PRIMEtime_local, described in Chapter 3. This produced a separate estimate of the effect on females and males for each lower tier and unitary authority in England (except the City of London and Isles of Scilly, due to size).

Outcomes were measured in terms of QALYs, costs, and disease cases (AF and flutter, asthma, breast cancer, colorectal cancer, oesophageal cancer, hypertensive heart disease, IHD, low back pain, osteoarthritis of the knee, osteoarthritis of the hip, stroke and T2DM). Where per-person values were calculated for these, the denominator was the number of children targeted in the intervention, ie. the number of children aged 5-17 years in each local area's 2018 population. The start of the simulation was set to 2018 and the time horizon was set at 100 years (lifetime) using a closed cohort. The intervention effect was only applied to children aged 5-17 years but they were then modelled across the entire lifetime. Trends to incidence and case fatality were applied for 10 years, and lags between the change to risk exposure and the change to disease incidence of 5 years for IHD, stroke and T2DM, 20 years for cancers, and 1 year for other modelled diseases. Discount rates were applied to costs and QALYs at the NICE recommendation for public health interventions of 1.5% for health and 3.5% for costs.(210) The model was run to produce a point estimate for the effect of the intervention in each of the 315 local authority areas.

Probabilistic Sensitivity Analysis on QALY and healthcare cost outcomes

To estimate the geographic distribution of effects of television advertising restrictions, PRIMEtime_local was used. The model was parameterised as described above in Chapters 3-

7. Uncertainty was estimated for QALY and healthcare cost outcomes, with the same analytic choices listed for the point estimate outcomes, above.

Due to prohibitive run time, PSA was performed for 10 areas, selected from the range of deprivation, RUC and geographic For each area, 2000 runs were performed, drawing uncertain values for BMI-disease RRs, BMI-mortality RRs, utilities values, costs and the effect of the intervention. The probability density functions for uncertainty around Relative Risks (between BMI and disease incidences, BMI and mortality rates and T2DM and disease incidences), utility values and the effect estimate, were assumed to be gaussian. Variation in costs was assumed to be log-normally distributed with mean and SD of log-costs taken to represent a full range of variation, rather than uncertainty, in the cost estimates. As costs are a function of the other inputs and not the other way around, this exploration of variation does not bias uncertainty estimates of the other parameters.

Health economic evaluation

Modelled impact of the intervention on lifetime QALYs and costs were used in combination with government estimates of the cost of the intervention from the Government's Impact Assessment,(198) to calculate cost-per-QALY and NMB compared with a business-as-usual scenario, for each local authority. A government healthcare (NHS) perspective was taken, with intervention costs due to legislating, implementing and enforcing the regulations, netted of modelled NHS savings due to lower disease incidence. Cost-per-QALY and NMB were examined across the spectrum of deprivation. Government implementation costs were £0.8m for setup over the course of 2022-23 and ongoing enforcement costs of

£370,000 per year.(198) The Government's Impact Assessment(198) gives estimates of uncertainty only at more aggregate levels of total costs, though are equivalent to +/-40% of setup costs and +/-11% of ongoing enforcement costs, so these were applied to the stated mean values. For the purposes of uncertainty in the CEA, these were treated as 95% CI, giving estimates of £0.48m-£1.1m for transition costs and £330,000-£410,000/ year for enforcement. To implement this, first, 3000 random draws were taken from these distributions, then each of these uncertainty estimates translated into a local cost allocation by weighting to the proportion of children in each local area (and accounting for England having only 84.9% of the UK's 5–17-year-olds in 2018.(109) These intervention costs were then paired with the 2000 uncertainty estimates of QALY and healthcare cost outcomes from PRIMEtime, to estimate NMB. This used a willingness to pay threshold of £30,000/QALY against a business-as-usual scenario. As the intervention is only framed to be effective during childhood years and a closed cohort is modelled, the total intervention-related cost was calculated over 13 years, ie. for the whole cohort of 5-17 year-olds to reach 18 years of age. Any health care costs unrelated to the intervention or conditions of interest were excluded, in line with NICE guidance.(210,211)

Sensitivity Analyses

Isolating local input effects

The differences in modelled QALY outputs between the different local authorities arise from differences in input values – specifically, age structure (the numbers of people of different ages), BMI distribution, epidemiology and the effects of the intervention on BMI all vary by local area. Costs could be broken down in a similar manner, though QALYs were chosen as a

single example of these differences. Equivalent outcomes for costs would look very similar to those for QALYs The scenario was explored in terms of identifying which of these factors were driving variation in QALYs saved between areas, by standardising the between-area variation in age structure, BMI distribution, disease epidemiology and the intervention effect, leaving only one of these parameters in turn as local estimates. An average population was estimated by simply drawing mean population size by age and sex across the 315 areas. An average BMI distribution and disease epidemiology by age and sex were estimated as population-weighted means across the 315 areas and the average effect was taken as a crude mean across the areas.

Deterministic sensitivity analyses

Four deterministic assumptions to the intervention were also tested, two around the 'structural' assumptions of the logic model linking intervention to effect, and two on discount rates:

1) That adverts would be displaced until after 9pm rather than being removed from viewing altogether. The calorie impact of this intervention was tested by re-calculating the impact of the intervention on child calorie consumption then modelling these health and cost consequences.
2) That the impact of the intervention was the same for obese as healthy-weight children. Four of the 11 studies in the meta-analysis by Russel *et al* separated these groups. Assuming no difference in effect between healthy weight and overweight/ obese children, and using the result of the meta-analysis including all 11 studies also resulted in a lower mean estimate of 14.2kcal/ minute exposure.(208)

3) Discount rates were varied

   a) from the NICE recommendation of 1.5% for health and 3.5% for cost impacts for public health interventions to the UK Treasury Green Book recommendations of 1.5% and 3.5% (respectively) between 0-30 years, falling to 1.29% and 3% between 31-75 years and 1.07% and 2.5% thereafter.(204)

   b) to zero for health impacts, avoiding biasing disease experienced earlier in life as more important.

Results

Effect of intervention on calorie consumption

Table 8.1 shows the input values for the Monte Carlo analysis estimating the effect of the intervention on energy consumption for each local authority. Numbers of adverts seen per day by social grade describe, for example, that children in families of social grades DE (low) are estimated to watch over double the number of television adverts per day as those in grades AB (high).

Table 8.1: Input parameters for Monte Carlo analysis estimating uncertainty and local variation of television advertising restrictions on calorie intake

Input	Group	Mean	Uncertainty
HFSS adverts seen per day(207)			
	Social Grade AB	0.97	(No uncertainty)
	Social Grade C	1.42	(No uncertainty)

	Social Grade DE	2.06	(No uncertainty)
Seconds per advert(203)			SD
	All	25.9	11.9
Effect on calorie intake per minute exposure(208)			SE
	Healthy weight	18.2	0.676
	Overweight or obese	28.5	1.30
	Whole population (sensitivity analysis 2)	14.2	0.241
Proportions by Social Grade(114)		Mean of area proportions	Mean SE of area proportions
	AB	25.0%	0.318%
	C	62.5%	0.405%
	DE	12.5%	0.236%

Figure 8.2 shows the distribution of point estimates of average calorie reduction per day, for areas in each IMD 2019 quintile, as kernel density plots. This shows greater average impacts on calorie consumption for more deprived areas. This pattern is consistent with the input values to the Monte Carlo analysis, in that more deprived areas tending to have higher rates of overweight/ obesity and higher proportions of people in lower social grades. There was a great deal of overlap between deprivation quintiles, with the smallest effect in each of the four more deprived quintiles being around the same as the modal effect of the least deprived quintile.

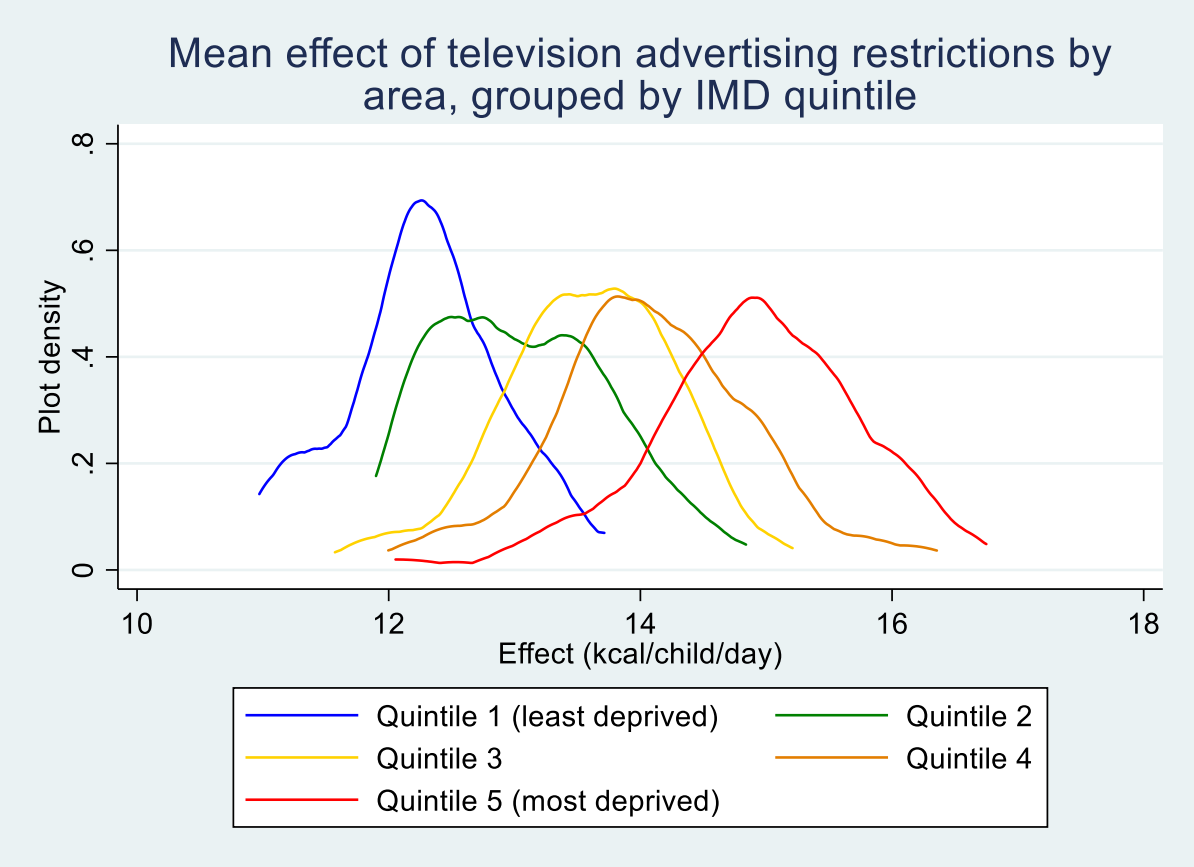


Figure 8.2: Variation in mean estimates of impact of television advertising restrictions on child calorie consumption, by local area Index of Multiple Deprivation quintile.

Table 8.2 gives the mean calorie change effect for the main scenario across all 315 area point estimates. The SEMs of each area's mean estimate are also averaged, providing a sense of the uncertainty used un the PRIMEtime Monte Carlo analysis. The SD of the areas' mean calorie change estimates is also given to demonstrate the spread of effect.

Table 8.2: Estimated calorie change associated with proposed television advertising restrictions on HFSS products

* Effect of exposure to advertising on calorie consumption does not vary by weight category.

** Adverts that are restricted from being shown before 9pm are assumed to be shifted to after 9pm.			
Analysis	Mean of area point estimates	Mean SEM	SD of area means
Main analysis	13.6kcal/day	0.0993	5.44
Sensitivity analysis 1*	8.70 kcal/day	0.0635	3.48
Sensitivity analysis 2**	4.54 kcal/day	0.0329	1.80

Modelled outcomes – point estimates

Table 8.3 shows the mean and SD of point estimates of QALYs and costs saved per capita by area IMD quintile and sex. These are then displayed as continuous data in figures 8.3 and 8.4 respectively, as a histogram (all areas) overlaid with kernel density plots of these effects categorised by area IMD, by females and males. There were broad overlaps of effects between deprivation groups, with larger and more widely varied impacts for more deprived areas. Males had slightly greater QALY and cost savings than females for each IMD quintile. The range of QALYs saved ranges from 0.0222 QALYs per person in Richmond upon Thames (London) to 0.0521 QALYs per person in Lincoln. For costs, the smallest saving was £11.7 per person in Blaby (Leicestershire) and greatest £42.8 per person in Nottingham.

Table 8.3: Modelled impact on incremental QALYs and incremental healthcare cost (2018/19 prices) saved per person, for television advertising restrictions, by IMD quintile and sex

IMD quintile	Females		Males	
	Mean	SD	Mean	SD
QALYs:				
1	0.0279	0.00224	0.0314	0.00267
2	0.0285	0.00217	0.0344	0.00302
3	0.0331	0.00226	0.0415	0.00328
4	0.0353	0.00254	0.0409	0.00356
5	0.0363	0.00291	0.0490	0.00405
Costs:				
1	-25.6	4.46	-16.3	2.59
2	-26.5	4.22	-17.2	2.57
3	-29.2	4.47	-19.3	2.70
4	-30.3	5.59	-20.6	3.56
5	-34.2	7.58	-23.1	5.03

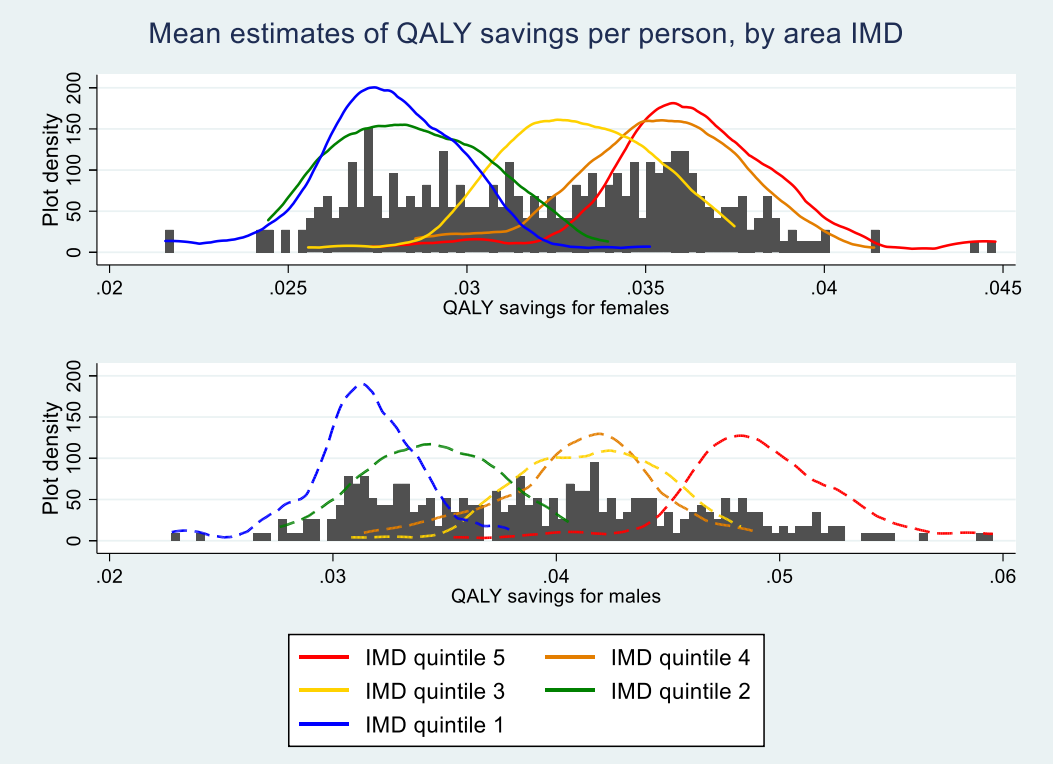


Figure 8.3: Modelled effect of television advertising restrictions scenario on QALYs saved per person for females and males, by area IMD quintile.

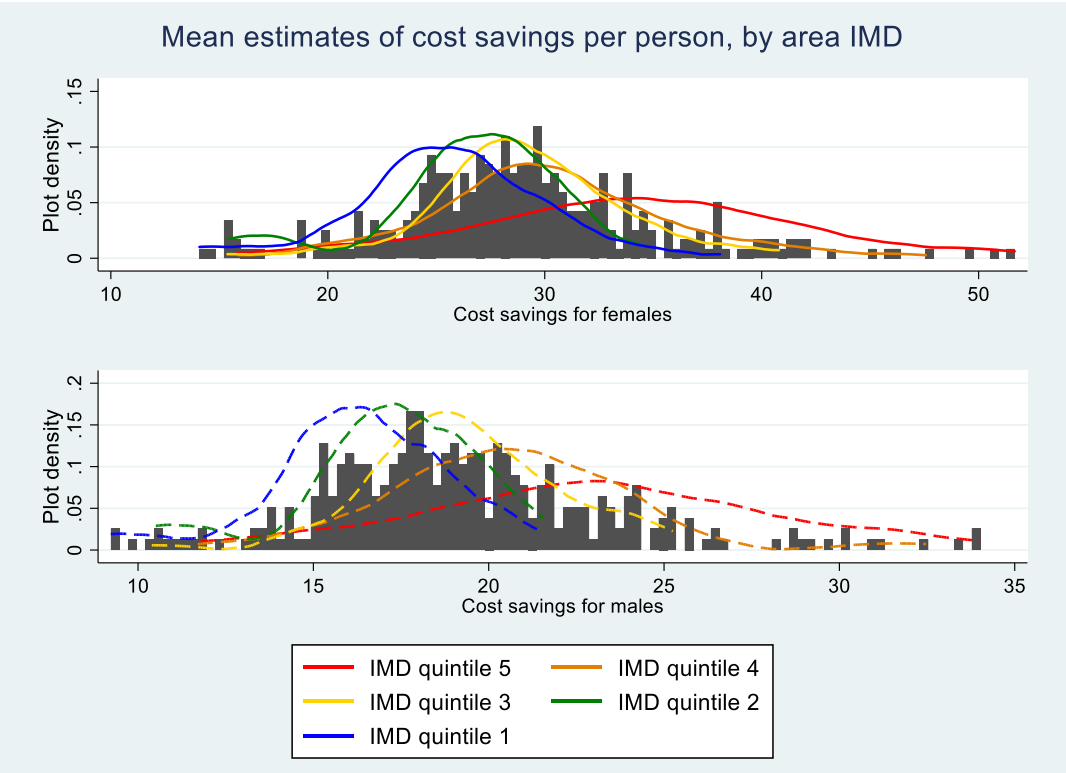


Figure 8.4: Modelled effect of television advertising restrictions scenario on costs saved per person for females and males, by area IMD quintile.

Estimated impacts on numbers of disease cases are presented in table 8.4 (per person). Orders of effects across IMD quintiles show comparable patterns to those for QALYs or costs, with greater benefits for more deprived areas, though for colorectal cancer there is an increase in cases. Because magnitudes of the intervention's impact on cases increases with deprivation, the most deprived areas have the greatest increases in colorectal cancer cases, as well as benefitting the most where cases decrease.

Table 8.4: Incremental change to numbers of cases prevented of modelled diseases (per person), by area IMD (undiscounted, to 3 s.f.). Negative numbers represent reductions in numbers of cases and positive numbers represent increases.

	Females		Males	
IMD quintile	Mean	SD	Mean	SD
	AF and flutter			
1	-0.000715	0.0000555	-0.00111	0.0000880
2	-0.000785	0.0000632	-0.00109	0.0000931
3	-0.000781	0.0000604	-0.00111	0.0000918
4	-0.000760	0.0000739	-0.00106	0.0001019
5	-0.000811	0.0000862	-0.00112	0.0001134
	Asthma			
1	-0.00332	0.000216	-0.00176	0.000137
2	-0.00351	0.0002383	-0.00185	0.000142
3	-0.00365	0.0002442	-0.00194	0.000141
4	-0.00375	0.0002349	-0.00205	0.000139
5	-0.00381	0.0002946	-0.00204	0.000170
	Breast cancer			
1	-0.000168	0.0000146	-	-
2	-0.000180	0.0000178	-	-
3	-0.000164	0.0000153	-	-
4	-0.000168	0.0000216	-	-
5	-0.000192	0.0000258	-	-
	Colorectal cancer			
1	0.000116	7.79E-06	0.000218	0.0000176
2	0.000130	9.75E-06	0.000302	0.0000406
3	0.000149	0.0000102	0.000436	0.0000482
4	0.000161	0.0000103	0.000430	0.0000347
5	0.000160	0.0000121	0.000416	0.0000340
	Hypertensive heart disease			
1	-0.000214	0.0000168	-0.000268	0.0000196
2	-0.000215	0.0000185	-0.000270	0.0000207
3	-0.000211	0.0000154	-0.000260	0.0000172
4	-0.000210	0.0000163	-0.000278	0.0000192
5	-0.000210	0.0000188	-0.000280	0.0000231
	Ischaemic heart disease			
1	-0.00172	0.000125	-0.00295	0.000228
2	-0.00183	0.000143	-0.00339	0.000292
3	-0.00198	0.000134	-0.00312	0.000275

4	-0.00205	0.000161	-0.00357	0.000306
5	-0.00214	0.000199	-0.00334	0.000286
	Low back pain			
1	-0.0117	0.000984	-0.00857	0.000842
2	-0.0124	0.00112	-0.00908	0.00100
3	-0.0125	0.00101	-0.00895	0.000917
4	-0.0128	0.00114	-0.00968	0.000956
5	-0.0133	0.00129	-0.00988	0.00102
	Oesophageal cancer			
1	-0.000176	0.0000179	-0.000236	0.0000292
2	-0.000173	0.0000229	-0.000285	0.0000405
3	-0.000220	0.0000263	-0.000260	0.0000458
4	-0.000224	0.0000305	-0.000347	0.0000581
5	-0.000252	0.0000379	-0.000388	0.0000719
	Osteoarthritis hip			
1	-0.000160	0.0000120	-0.000136	0.0000110
2	-0.000169	0.0000137	-0.000142	0.0000124
3	-0.000170	0.0000116	-0.000144	0.0000113
4	-0.000177	0.0000141	-0.000155	0.0000123
5	-0.000182	0.0000170	-0.000156	0.0000138
	Osteoarthritis knee			
1	-0.00425	0.000310	-0.00180	0.000134
2	-0.00449	0.000350	-0.00187	0.000151
3	-0.00457	0.000299	-0.00187	0.000132
4	-0.00475	0.000361	-0.00200	0.000147
5	-0.00488	0.000436	-0.00202	0.000170
	Stroke			
1	-0.000623	0.0000461	-0.000267	0.0000267
2	-0.000679	0.0000535	-0.000417	0.0000389
3	-0.000664	0.0000454	-0.000378	0.0000349
4	-0.000715	0.0000552	-0.000441	0.0000405
5	-0.000741	0.0000684	-0.000468	0.0000463
	T2DM			
1	-0.0179	0.00121	-0.0150	0.00103
2	-0.0186	0.00145	-0.0162	0.00124
3	-0.0211	0.00148	-0.0186	0.00136
4	-0.0233	0.00159	-0.0208	0.00152
5	-0.0224	0.00169	-0.0214	0.00173

Figure 8.5 shows the point estimates of cost-per-QALY for each local authority (using costs as healthcare cost impacts net of intervention costs), demonstrating that mean modelled incremental cost per QALY was very consistent across the range of IMD scores. Despite the negligible slope of cost/ QALY across deprivation, there was wide variation in cost per QALY at any given level: across all areas, mean healthcare saving per QALY was £675 with a SD of £20.1. In all areas, point estimates indicated the intervention was cost-saving, and therefore dominant compared to the business-as-usual situation.

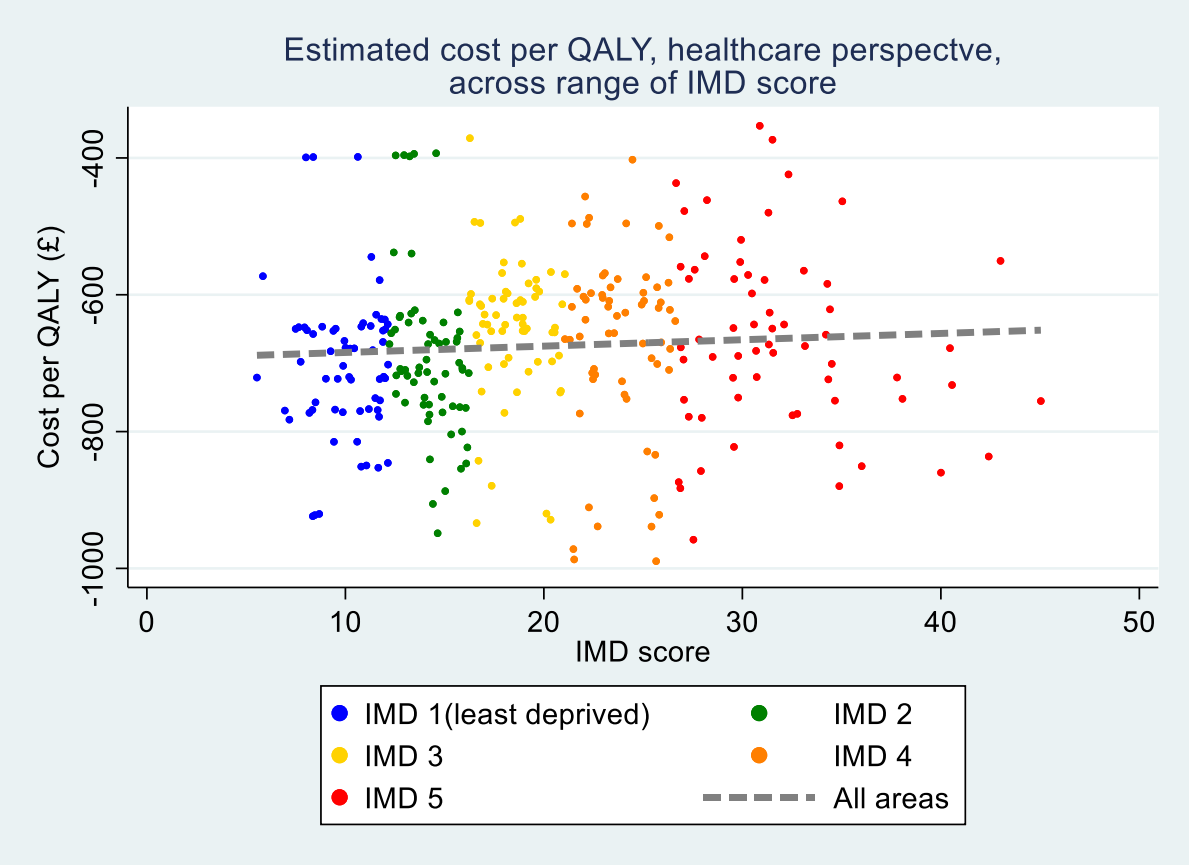


Figure 8.5: Scatter plot of cost per QALY from healthcare perspective against Index of Multiple Deprivation Score.

Probabilistic Sensitivity Analysis

The PSA shown in figure 8.6 and table 8.5 demonstrates uncertainty tended to be proportionally related to effect, for both QALYs and healthcare cost savings. In no cases were healthcare costs anticipated to rise. Best fit lines for each area in figure 8.6 are largely parallel, indicating the uncertain ranges of cost-effectiveness between areas are likely to be comparable. For legibility, equivalent charts separating the scatter plots by area and demonstrating the same confidence ellipses are shown in Appendix 1g. Each area's RUC code is included in table 8.5. Table 8.5 indicates there are no very obvious trends in effect of RUC on QALYs or costs, though there are not enough cases here to judge whether there are broader trends in uncertainty between types of area or perform a formal analysis.

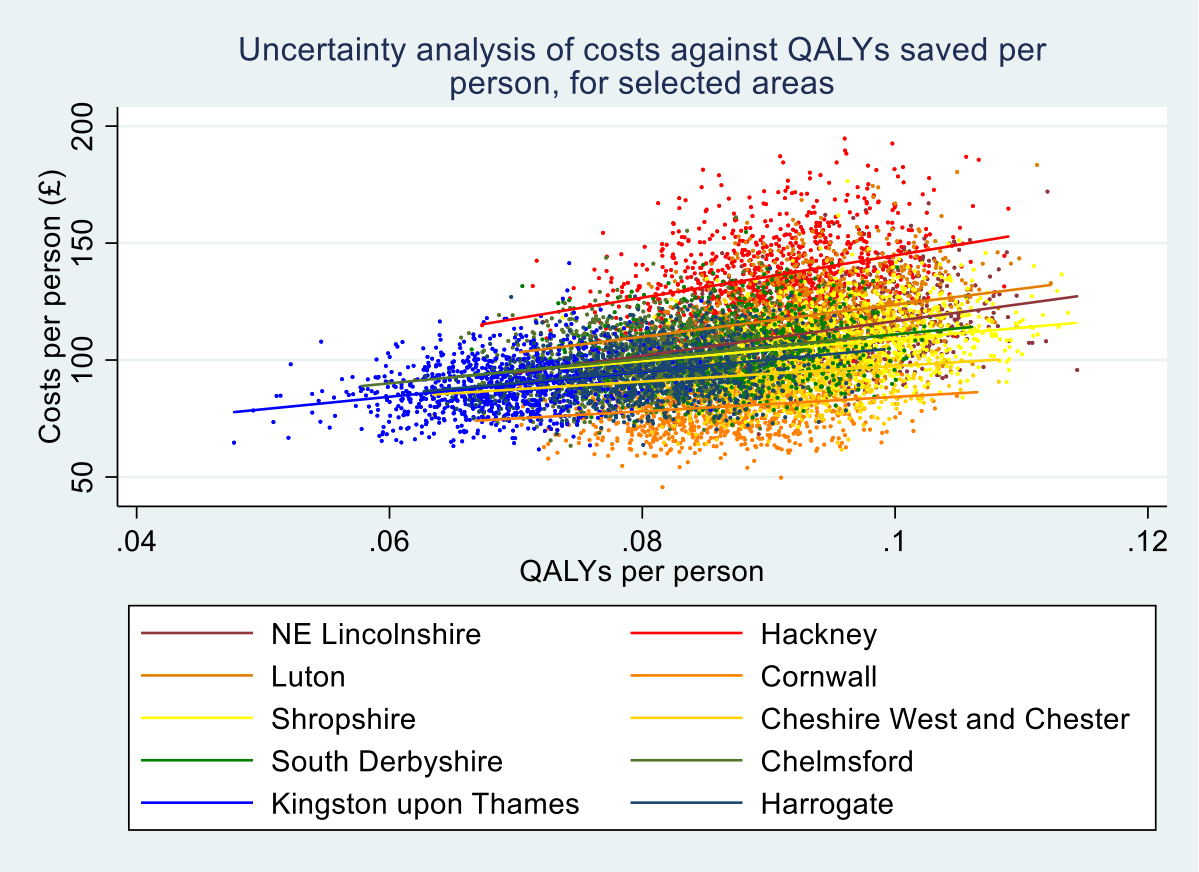


Figure 8.6: Scatter plot of Probabilistic Sensitivity Analysis of cost savings against QALYs for 10 selected areas.

Table 8.5: Mean and SD of uncertain incremental QALYs and costs from PSA, by area.

Location	IMD quintile	RUC code	QALYs per person: Mean	QALYs per person: SD	Costs per person (£): Mean	Costs per person (£): SD
Harrogate	1	3	.0813	.00595	-95.8	11.5
Kingston upon Thames	1	6	.0693	.00547	-89.3	11.4
Chelmsford	2	4	.0802	.00572	-100	12.5
South Derbyshire	2	3	.0861	.00620	-104	13.0
Cheshire West and Chester	3	3	.0905	.00637	-94.3	13.9
Shropshire	3	2	.0959	.00655	-108	11.8
Cornwall	4	1	.0864	.00615	-80.1	10.3
Luton	4	4	.0934	.00621	-119	15.6
Hackney	5	6	.0905	.00638	-112	14.7
North East Lincolnshire	5	4	.0939	.00614	-136	18.0

The uncertainty estimates produced from the PSA were then used to explore uncertainty in the HEE. Incremental net government costs (intervention costs net of healthcare cost savings) are plotted against incremental QALYs as a cost-effectiveness plane in figure 8.7. This demonstrates the orders of magnitude of incremental costs and QALYs relative to the willingness to pay threshold of £30,000/ QALY. NMB is calculated for each of the ten selected areas, with uncertainty, shown in table 8.6. This PSA indicates that for none of these areas' uncertainty crosses the break-even point, indicating that after uncertainty is taken into account, the intervention continues to be likely to be dominant compared to business-as-usual, at least in these selected areas.

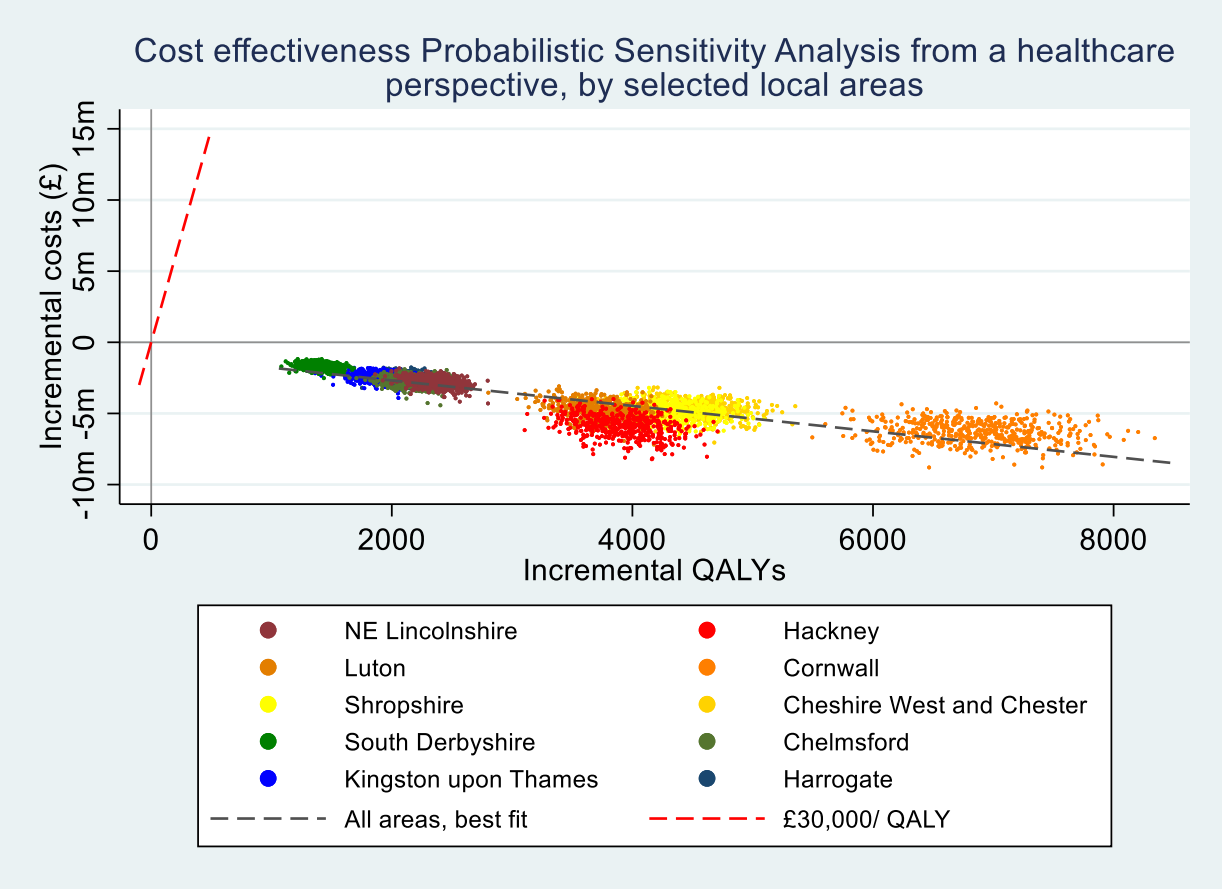


Figure 8.7: Cost effectiveness plane demonstrating distributions of Probabilistic Sensitivity Analysis around incremental QALYs and costs, with a best-fit line across all uncertainty

estimates across all areas and line indicating a willingness to pay threshold of £30,000/QALY.

Given only ten of the 315 areas are fully modelled here, it would be useful to estimate approximate uncertainty on cost-effectiveness for other areas. Using the coefficient of variation (CV, as CV = SD/ mean) (in table 8.6), variation in NMB can be approximated for other areas. The area with the lowest NMB per person was Richmond on Thames (London), at £683. The highest CV of the ten areas (given in table 8.6) is 0.0772, so using this would pessimistically infer an inferred lower uncertainty interval for Richmond upon Thames of: £683 – (0.0695 * £683 * 1.96 = £580 per person, making it unlikely that any area's fully modelled uncertainty intervals would cross the break-even point on this economic perspective. By coincidence, Kingston upon Thames (full PSA performed) had the second lowest point estimate of NMB per person of any area, which also was not close to crossing the break-even point, corroborating this likelihood.

Table 8.6: Mean and 95% Uncertainty Intervals on Net Monetary Benefit estimates, with coefficient of variation (CV = SD/mean) (to 3 s.f.).

Area	IMD quintile	Mean	Upper UI	Lower UI	CV
Harrogate	1	63,700,000	72,700,000	54,800,000	0.0718
Kingston upon Thames	1	60,000,000	69,100,000	50,900,000	0.0772
Chelmsford	2	68,800,000	78,300,000	59,400,000	0.0699
South Derbyshire	2	43,700,0000	49,700,000	37,600,000	0.0706
Cheshire West and Chester	3	139,000,000	157,000,000	120,000,000	0.0689
Shropshire	3	133,000,000	151,000,000	116,000,000	0.0671
Cornwall	4	212,000,000	241,000,000	183,000,000	0.0699
Luton	4	116,000,000	131,000,000	101,000,000	0.0654
Hackney	5	123,000,000	140,000,000	106,000,000	0.0695
North East Lincolnshire	5	73,200,000	82,500,000	63,900,000	0.0647

Sensitivity Analyses

Sensitivity analyses – isolating local input effects

To explore which of the local-level inputs was having greater effects on driving difference between the point estimates for each area, the modelling was repeated but only allowing one of these factors at a time to vary by local area, with the remainder fixed to an England average. The health effects of differences in population structure, BMI distribution and intervention effect alone are shown in turn in figure 8.8 (aggregated by sex). The isolated effect of disease epidemiology is shown in figure 8.9 and effects separated out by sex as the results by sex were very different, while for the three inputs in figure 8.8 they were very similar. For each sex, the order of effects remained broadly with the greatest effects in the most deprived areas. Males also had greater effects than females when only accounting for BMI distribution or epidemiology. The effects of population are highly skewed, in essence following the size of the populations. BMI distributions and the effect of the intervention follow increasing patterns in line with deprivation, albeit with large overlaps. The interquartile range was 758 for the analysis allowing only allowing the population to vary, 59.8 for BMI distributions, 139 for the effect and 76.4 for epidemiology (though this is not very meaningful).

The effect of the underlying epidemiology is unlike that of the other factors. Epidemiology is modelled at the IMD quintile level, so there is no variation between areas within the same IMD quintile. The benefit for the least deprived was less overall and more deprived quintile

more overall, though the pattern is more disordered, increasing in the order 2, 5, 1, 3, 4 for females and 1, 2, 4, 3, 5 for males.

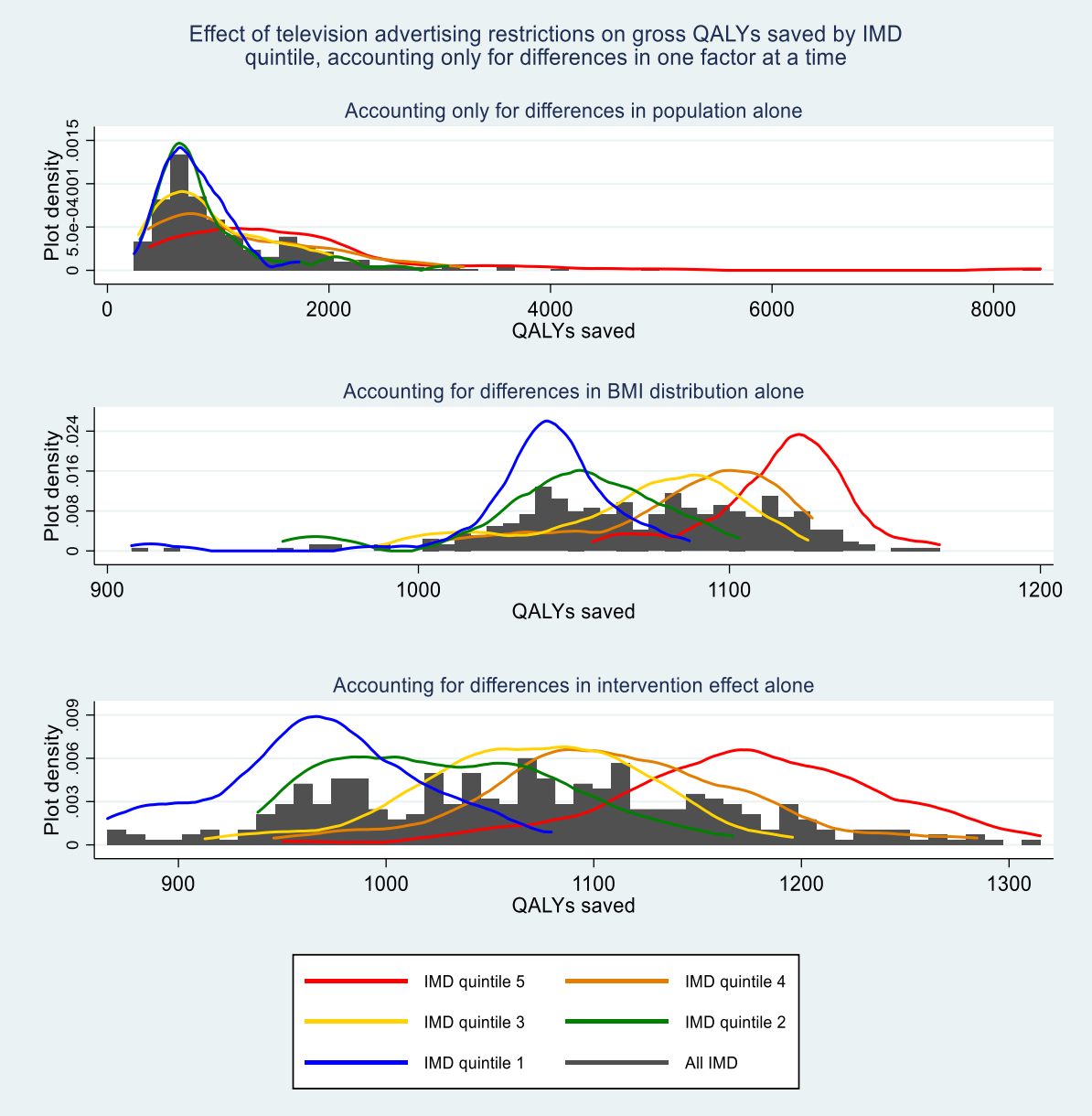


Figure 8.8: Distributions of QALYs saved accounting only for one locally-varying factor at a time, removing between-area variation in the PRIMEtime modelling for others.

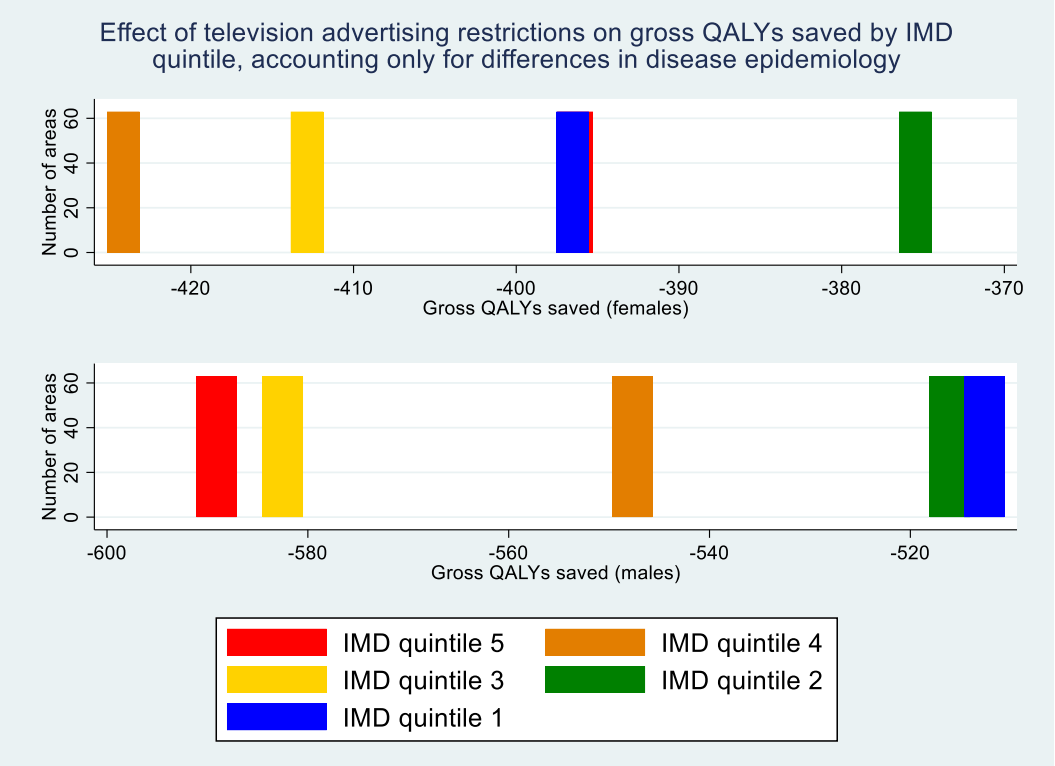


Figure 8.9: Effect of television advertising intervention on local areas on QALYs saved, accounting only for differences in disease epidemiology.

Sensitivity analyses – deterministic sensitivity analyses and discounting

Testing the assumption that advertising will be delayed to later in the evening rather than cancelled reduces the scale of the effect by around two thirds, as shown in figure 8.10 and table 8.7, in line with this sensitivity analysis in Mytton *et al*.(203) Assuming that the effect is the same for healthy weight children as overweight/ obese children also results in attenuated estimates of impact to a lesser degree, bearing in mind that this sensitivity analysis involved a reduced mean effect as well as less between-area variation, as detailed in table 8.1. Reducing discount rates also predictably increases estimated effects.

Table 8.7 shows the quantitative differences between sensitivity analyses in terms of their estimates of QALYs saved and their percentage difference from the main analysis. This shows that effects are smaller for the sensitivity analyses on structural assumptions and greater for sensitivity analyses varying the discount rates, as one would predict. The undiscounted health benefits are over double that of the main analysis. Variation follows a comparable pattern, with SD increasing or decreasing by similar proportionate change to the means, with again with the SD of the undiscounted analysis over double that of the main analysis.

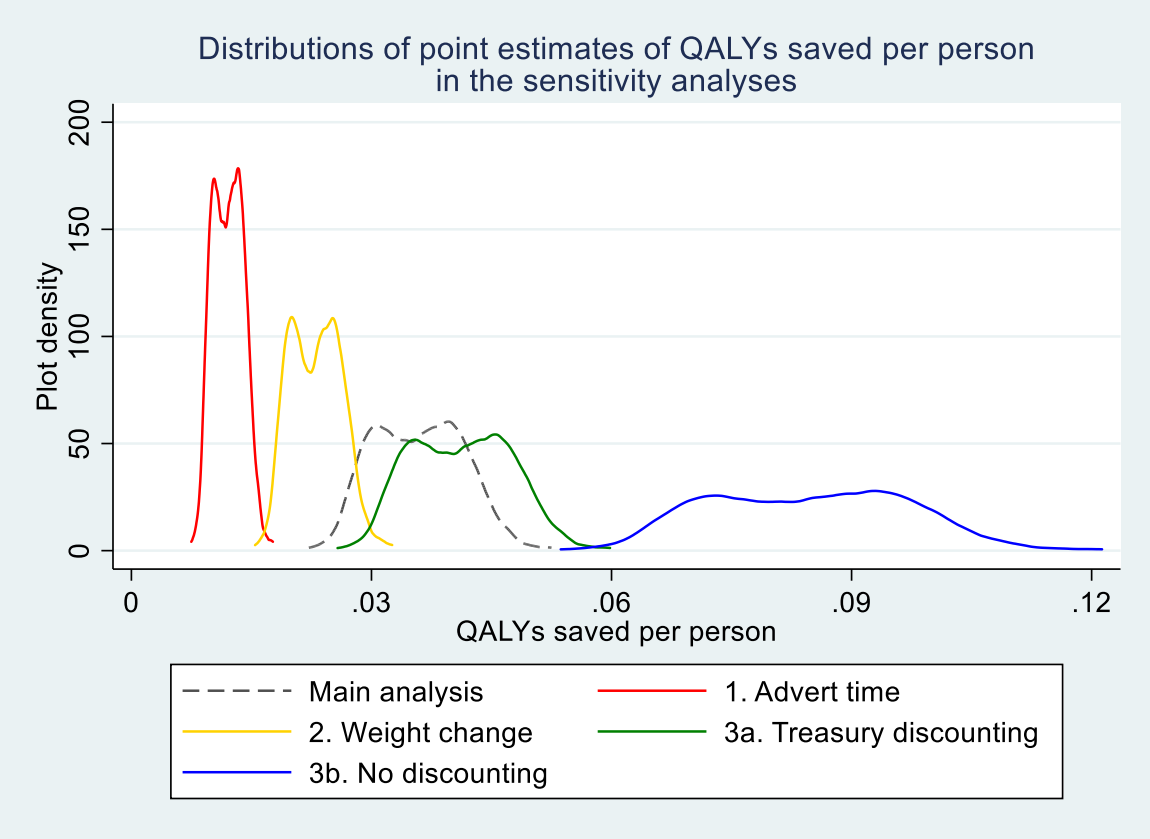


Figure 8.10: Distributions of sensitivity analysis estimates of QALYs saved per person, with main analysis distribution for reference.

Table 8.7: Crude mean and SD of effects of the sensitivity analyses on QALYs, and their percent difference from the main analysis.

	QALYs saved per person		% of main analysis outcome	
Sensitivity Analysis	Mean	SD	Mean	SD
(Main analysis)	0.0359	0.00555		
1. Advert time	0.0121	0.00187	33.7	33.7
2. Weight change	0.0231	0.00315	64.4	56.8
3a. Treasury discounting	0.0413	0.00623	115	112
3b. No discounting	0.0847	0.0123	236	222

Discussion

Summary of findings

This analysis produces estimates of health, QALY and cost impacts of the proposed restrictions to HFSS advertising on television for each lower tier and unitary local authority in England. The mean effect of the intervention is estimated at 13.6kcal per person per day (SD across area means 5.44kcal), with greater benefit estimated on average for areas in the more deprived IMD quintiles, but with large overlaps between quintiles.

This modelled analysis estimates that proposed restrictions to HFSS advertising on television could save approximately 0.221-0.521 QALYs per person, making up £11.7-£42.8 savings per person to the NHS. Men consistently benefitted more. Between-area variation is greater for the more deprived IMD quintiles and more for costs than QALYs. We would expect costs to vary more than QALYs as both are modelled to arise from cases, in this analysis the utilities

that convert cases to QALY changes do not vary between areas, while the unit prices that convert cases to costs do vary by each individual area, with greater variation between areas' costs in more deprived areas, explaining the greater overall spread. Although both cost savings and health benefits tended to be higher in more deprived areas, there was no overall difference in cost-effectiveness across the range of IMD score.

The impact of the intervention on the numbers of disease cases varied widely between diseases as well as areas. For example, there were 0.0179-0.0224 cases of T2DM prevented per female child, while cases of colorectal cancer increased by 0.000218-0.000416 cases per male child. That T2DM has baseline high rates and a high Relative Risk from BMI may explain why it had the greatest decrease in cases. Cases of colorectal cancer may be modelled to rise as a consequence of age-related cases being greater than the reduction due to BMI-related cases.

The PSA for 10 areas showed a great deal of overlap between areas' estimates. Areas in the most deprived quintiles showed larger effects on QALYs and costs saved per person than in the less deprived quintiles, though cost per QALY was fairly consistent between IMD quintiles at around -£675/QALY, with a SD of £20.1. From a healthcare perspective, accounting for government input costs and healthcare savings equivalised from the 25-year horizon, the intervention was found to be highly cost-effective against the £30,000/QALY threshold, indeed, cost-saving.

When isolating the influence of key factors on variation in areas' outcomes, it was found that BMI distributions, baseline epidemiology and (most clearly of all) the intervention effect all contribute to the tendency for greater effects in the most deprived areas. The disordered nature of the quintiles seen when the isolated effect of the disease epidemiology

is modelled (in figure 8.9) is reminiscent of the disordered patterns shown in the Chapter 4 detailing the process for estimating baseline disease epidemiology, though as discussed, these data arise from the GBD estimates, which themselves largely rely on measured data for local areas of England, but methods are not entirely transparent and reasons behind these patterns cannot be definitively explained here. The effect appeared to drive more variation than the BMI distribution or epidemiology.

Taking a single factor alone, there is not a clear sense that the variation within quintiles increases with increasing deprivation as we can see with the main analysis. The sensitivity analyses varying structural assumptions made in the estimate of the effect both reduced the size of the modelled outcomes, and those varying discounting to lower rates increased effects, as we would expect.

Key assumptions

Theoretical model

The theoretical model of the intervention on its effects involves some important assumptions. It is assumed that children living under the proposed restrictions would reduce their energy consumption by the estimated amounts each day. Under the body weight change equations by Hall *et al* these changes take 3 years to accumulate to a new steady state,(92,93) which is built into PRIMEtime. Children would continue to put on excess weight but to a lower level. Energy intakes reduce as a consequence of the reduced exposure to advertising and it is not assumed that there are additional benefits with longer durations under the restrictions beyond the three years taken for the original energy intake

to finish impacting body weight. The increment on weight gain that people end childhood at is assumed not to be lost later by other means, so the lower body weight and BMI confer health benefits throughout adult life.

It was necessary to assume that there is a linear association between the number of seconds of adverts and the change to energy consumption, which is not formally evidenced to be either true or false. It was also assumed that the direct short-term impacts of advertising exposure on energy consumption captured in the Systematic Review and meta-analysis of experimental studies(208) then translates through to longer-term and more general impacts on consumption. The impact of advertising on calorie consumption was different for children of healthy weight than those above a healthy weight. This was implemented as a single categorical difference as this relationship was reported in this way in the original meta-analysis,(208) but we would expect a continuous relationship in the real world. Without knowing the distribution of how energy consumption changes continuously with BMI, it is impossible to comment on the impacts of this on the results.

Here we restrict the population to children in line with the government's targeting at children, and the weaker evidence that adult choices are shaped by HFSS advertising. If adults are affected by this advertising or not is a question debated elsewhere.(212–214)

Health economic perspective

Economic perspective involves a complex judgement. What resources to count as used or spared is a political choice, and there is a risk in this context that distributing national-level input costs to local areas to calculate measures of efficiency stretches its conceptual

coherence. The figures required to make economic evaluations may also require judgement on the quality of each of the choices, for example, we may want to include local-level social care costs – and have to choose between unreliable estimates of local costs, more reliable national-level estimates, or excluding these costs altogether. This HEE took a government healthcare (NHS) perspective, treating the cost of legislating, implementing and enforcing the regulations as the input costs. Net costs are estimated using modelled NHS savings due to lower disease incidence to offset input costs. This framework invites a range of questions. The government's Impact Assessment(198) highlights that industry bears most of the cost of the intervention, but these costs are not dealt with here as NICE does not recommend including them.(211) In a healthcare approach such as an education programme on HFSS products, a successful programme would see HFSS revenues fall, but these would not be expected to be estimated or included in the economic evaluation. It is also important to note that the government estimate of £370,000 per year for enforcement may be high, reducing estimated cost-effectiveness here. The predictability of future direct or indirect industry effects is very difficult. For example, people may consume less HFSS products but will need to exchange those purchases for different ones and may not consume less volume or cost of foods overall (indeed, they may do more of both). Then, even if food spending did fall, it is possible that aggregate demand would change very little, maintaining total commercial revenues, though hopefully to less damaging products. Then, these changes to demand are not the only impact on industry, but, as noted, potential benefits to productivity should also be considered alongside costs. Each of these issues raises enormously challenging research questions.

Industry costs have been estimated for the government's Impact Assessment(198) and it would be feasible to refashion these figures for use in an economic analysis in a comparable

way to the government implementation costs. This comes with the caveat that those costs could not be verified, which is in common with the implementation costs, though a greater problem for industry costs as there is likely to be considerably greater unpredictability what the net system effects on industry might be.

It is important to consider what the purpose of including a set of costs or effects is. The government makes investments in healthcare or public health, and in return expects improved health, quantifying how much of that health was gained given how much investment was made in the form of QALYs. This happens in isolation of other effects on the system such as industry revenues, employment, or economic productivity (eg. related to presenteeism, absenteeism, long-term sickness), though every intervention does have these effects to a greater or lesser degree. Comparability with other evaluations is desirable, indeed NICE issues extensive guidance(211) aiming to achieve greater consistency.

At higher levels of abstraction on industry impacts, people may eat less HFSS foods and less calories, but spending on food may remain the same or even increase. Therefore, if we are interested in private sector business we might be interested in the distribution of costs and benefits across the private sector (eg. as advertising agencies may be less likely to have other benefits circulated back to them) but this is an entirely different research question requiring different methods.

A consideration involved in using the economic perspective used is that costs to the government are accumulated at the UK level while the distribution of benefits is heterogeneous across England (here estimated at the local authority level but presumably with meaningful heterogeneity at the community and individual levels). This introduces the question of how to distribute national-level costs locally. It is easy and standard to compress

all costs and benefits into one national level as this concept of a range of local cost-effectiveness for a national intervention is complex. If the intention is to reduce health inequalities, then capturing differences in effectiveness between areas (or other relevant groups) is useful. Cost-effectiveness may have additional considerations. In some instances, cost-effectiveness may be positively or negatively associated with deprivation, which could be politically more difficult. For example, if average cost effectiveness was higher in more deprived areas, but effects greater in less deprived, then the more deprived areas would be in effect subsidising the less deprived areas to increase inequalities. Additionally, it can be useful to calculate cost-effectiveness if a value such as NMB helps to communicate the local variation effectively.

Time horizon

To estimate appropriate input values given different time horizons, assumptions must be made. Benefits accrue over the lifetime, with most morbidity avoided as relative risks and baseline incidence and prevalence rise in later life, but the window of costs is the 13 years the intervention is in place and acting on the cohort of current 5-17 year-olds', with benefits accruing as a legacy of those 13 years, regardless of what then happens. The immediate payback to the government over those 13 years would be small, as only a few diseases have any effect, so if cost-effectiveness were calculated strictly within only these years or within the assumed lifetime of the regulations (25 years) then cost effectiveness may be very different.

Much of the benefit estimated in this analysis is discounted away as it accrues later in life, decades after the costs, making it difficult to conceive of a fair equivalence of costs and

benefits in such a HEE. This raises the interesting issue of why we discount health at all. There is debate if health should be discounted more, less or the same as costs.(215) It is clear why we discount costs related to the time preference for the present and an opportunity cost of an investment (that we might quantify (204,210), according to NICE or the Treasury). However, benefits are still experienced the same regardless of when, there is no opportunity cost to having better health and no opportunity cost eroding relative benefits because they happen later. The Green Book justified discounting health on the basis that people have a preference for things closer in the future,(204) though with health there is not a now or later choice that we might apply to costs; good health can be enjoyed now and in the future. Discounting costs is conceptually coherent with how costs work in the real world, discounting health is not coherent to health in the real world. The Treasury Green Book and NICE Public Health Guidance are in agreement that health should be discounted less than costs, justified explicitly in the Green Book that health is not subject to the 'wealth effect' to financial outcomes, which represents decreasing marginal utility of a given amount of spending relative to increasing wealth with economic growth in future.(204,210) This conceptual approach is open to the critique that discounting health benefits more than costs leads to a situation where an intervention becomes ever more efficient the longer it is delayed, though this has been found to not be affecting real-world decision-making. It is therefore unclear that the answer should be to discount health and cost, rather than costs alone.(215,216)

Interpretation

This intervention takes a population approach, as first described by Rose in 1985,(17) in that it targets an entire population, regardless of level of baseline risk. Adams *et al*(36) described that population approaches may require less agency and may benefit those with the greatest need more, consistent with Shafir's theory on cognitive scarcity.(217) This intervention has potential to reduce health inequalities, consistent with the UK government's 'levelling up' agenda.(31)

The overlaps in effect between quintiles are an important area for consideration – differences within quintiles were often greater than differences between quintile means. It is interesting that IMD quintile is not enough to explain the wide variation between areas' effects, with greater exploration required to fully understand the pattern. That variation in point estimates of QALY or cost benefit is greater for more deprived areas in the main analysis may be expected, even though the isolated effects of factors (shown in figures 8.8 and 8.9) showed a weaker tendency for this pattern. This may be explained by the process of the model being multiplicative at various stages. Each input has a multiplicative impact on the estimated QALYs saved, so it is the areas that consistently fall in the upper ends of the distributions that get slightly greater multiplicative effects at each stage and benefit more overall from the intervention.

IMD does not predict QALY or healthcare cost outcomes particularly well, especially for more deprived areas. This is not necessarily a problem – the IMD quintiles are only used here to aggregate results and demonstrate the impacts of the intervention on inequalities. Meanwhile, part of the purpose of modelling by area-specific data is that it can be useful to understand that some areas may benefit more or less than apparently comparable areas without being able to precisely explain why. Indeed, if we knew precisely how and why

areas would vary in effect without modelling, then the modelling would be obsolete. On the other hand, we could ask if this lack of distinction is an artefact of the statistical methods in that the data used to capture differences between areas did not explain enough of the variation. The variables used may correlate with deprivation but are not a direct predictor of deprivation score, so the full relationship between these restrictions and deprivation may be greater in the real world than captured here.

The overlaps also demonstrate that modelling the mean cost and health impacts at the IMD quintile level may miss a great deal of the real-world heterogeneity and potentially provide misleading estimates of costs or effects for a single local authority. Within this, we see that some areas in the highest deprivation quintiles achieve comparable costs and effects to many of the least deprived areas. It may be important for policymakers to better understand which areas are harder to reach and with which interventions. For example, obesity rates happen to be lower than other areas of comparable deprivation, we need not worry about obesity-related inequalities in that area. If, however, obesity rates are high, but the effect of various obesity interventions is repeatedly low, then more thought needs to be given to how to target the need in that area. It is also not obvious why the population structure of higher deprivation areas should drive effect per person (as seen in figure 8.8), once other factors are kept stable. Most of the total effect of locally-specific population is due to size, which is big but largely uninteresting as it is an arbitrary bounding issue, ie. the boundaries between authorities could be anywhere and there could be any number of them. For example, it is a historic quirk that the conurbation around Birmingham is divided into six boroughs and Manchester into 10, leaving The City of Birmingham borough approximately twice the population of The City of Manchester borough despite comparable

overall metro sizes. After accounting for size, the remnant of effect that population structure benefits more deprived areas should be considered a coincidence.

Currently, the implementation of the proposed advertising restrictions have been put on hold by the government(202) following large increases in inflation and concern that the suite of proposed obesity policies including advertising restrictions may increase the costs further, particularly for the least well off. It is unclear why these further restrictions to advertising to children should meaningfully impact shopping decisions made by their parents based on affordability, indeed this implication contradicts the government's own assertion that the evidence for the impact of HFSS advertising on adult choices is too weak to include in the impact assessment.(198) This aside, this paper does not address the impacts of the proposed advertising restrictions on individual-, group- or area-level costs of living, nor headline or groceries-specific inflation, let alone how any of these issues should be dealt with from a policy point of view. With implementation on hold, there is an opportunity to improve policy design from the sensitivity analysis (also demonstrated in Mytton *et al*(203)) showing that displacing adverts until after the watershed would not only dramatically reduce expected benefits but also reduce the benefit for the most deprived by far more than the least deprived. The intervention is being implemented with major costs to the government and industry and it is non-trivial that benefits can be greatly enhanced with a small change to its implementation – for example increasing the restrictions window to, say, 11pm.

There are also effects of the intervention on society that are more difficult to quantify or that are plainly intangible, such as the impact on "freedom of choice" from not seeing given adverts, or that investment in children's television may decline and therefore have lower

educational quality.(199,200) It was certainly not feasible here to account for all such factors – if it is even possible.

This analysis shows a larger effect than estimated by Mytton *et al*(203) as expected due to the larger coefficients used to convert minutes of exposure to calories consumed. The sensitivity analysis making the same assumptions on this coefficient found broadly comparable number of QALYs saved over 100 years, at an equivalent of 202,000 QALYs rather than the estimate of 240,000 DALYs averted in Mytton *et al*(203) NHS savings in the Government Impact assessment were estimated at £18m compared to a total of £524m here, though apart from the time horizon here being four times longer, the exact parameters and assumptions made in the impact assessment are not clear enough to know why.(198)

Strengths

This modelled analysis uses locally-estimated input parameters to provide new insight into the distributional implications of population-wide public health policy. The underlying strengths of the modelling method are discussed below in Chapter 9, but briefly would include modelling across the full range of local authority areas, which is a lower level of geography than previously available for a scenario model outside the context of the National Diabetes Prevention Model(73) and PHE CVD model(74), both of which model on specific pre-set interventions and focus on their eponymous diseases. Most input parameters were locally-estimated, providing layers of locally-specific dynamics to the modelling. What this low level of geography provides is a greater area-specific granularity

through the modelling process, for example allowing the identification of overlaps between areas of different deprivation. Assuming that an area in each IMD quintile will behave as the mean area of that quintile would obliterate a great deal of within and between-quintile variation, while potentially producing an estimate that is very inaccurate the to the area being examined. How each area will respond needs to be modelled separately to be meaningful, consistent with basic ideas of complexity leading to dimensions of local variation that cannot simply be assumed.

In this case, the analysis has estimated locally-specific effects for a proposed national-level intervention that is currently suspended. That the effects of proposed television advertising restrictions for the most deprived areas vary more than for the least deprived, and there is a great deal of overlap between IMD quintiles. This tendency was more pronounced for cost implications than health implications and although there was no relationship between NMB with level of deprivation, there was wide variation in NMB for any given level. Uncertainty was modelled via 2000 runs of Monte Carlo analysis for 10 areas, demonstrating that magnitude of uncertainty tended to be directly related to magnitude of effect, as we might expect. More work is required to understand both dynamics of variation between point estimates and uncertainty intervals.

What is important here is that this could not have been shown without locally-specific modelling. Modelling on deprivation alone provides information on how health inequalities may be impacted on average, though as the IMD is an abstraction and real people live in specific locations, it does not tell us anything useful about where will be benefitted by a given intervention. The key implication of the wide variation in local effects is that it starts

to change the dynamic to how we consider population-level public health interventions. Theory is that population-level interventions may be more like to benefit people in line with need and deprivation.(36) This is a simplifying assumption that looks like it may be more of an oversimplification than is generally accepted. Any assumption that these relationships will reliably conform at the small group level as they do on the larger aggregate level appears very flawed.

Limitations

There are important limitations to this analysis. The estimation of the intervention effect used a comparable logic model to that used in Mytton *et al*.(203) The inputs to this process were taken from other primary research on the advertising system and its effects on health. The number of adverts seen per day was taken from a paper by Adams *et al*(207) from 2012, so of writing 11 years old in a rapidly changing media environment. These estimates were broken down by social grade, though no uncertainty was available on them. The evidence linking seconds of advert exposure with change to calorie intakes was a meta-analysis of a systematic review of randomised studies examining children's feeding behaviour after watching HFSS adverts. This came with the limitation that to the effect by healthy and overweight children, only studies that did this subgroup analysis could be included, resulting in a greater average effect than in Mytton *et al* – 18.2kcal/ day for healthy weight children and 28.5kcal/ day for overweight/ obese children, rather than the 14.2kcal/ day for the whole population in Mytton *et al*.(203) This difference between healthy weight and overweight/ obese children was also only quantified by these two categories, so had to be

implemented as such in the estimation process, rather than having a continuous distribution of effect across the range of BMI (and, indeed, with other variation such as by age and sex). Additionally, we need to be wary that what happens in experimental conditions is not necessarily what happens in life. The social grade item of the Active Lives Survey was taken at the local authority level and was essential in shaping local differences in effect. This variable is a derived variable from five self-reported items, which is a system designed by the ONS to shortcut the more detailed formal system of determining social grade. The two approaches have 75% agreement, with a small net tendency for people to overestimate their social grade in the shorter version.(218) Two impacts on the modelled results are relevant here: first, this would reduce the overall estimated effect of the intervention, and second, dilute differences between areas, both of which would confer conservative bias on interpretation.

Uncertainty on the effect of advertising restrictions on energy consumption was estimated via a Monte Carlo analysis, using uncertainty on the inputs where available (shown in table 8.1), to generate uncertainty around the estimated effect. Each paper uses slightly different age ranges and the Systematic Review/ meta-analysis used for estimating the impact of advertising to consumption includes samples with different age ranges, so it has been necessary to assume that these all have an equivalence to one another and to the population modelled here. The uncertainty for the number of seconds in each advert was available as the SD(203) – representing variation, not parameter uncertainty – while the uncertainty for the effect of advertising on consumption and for social grade were given as SE. Moreover, the SE for proportions in each social grade were given by a Monte Carlo analysis, so are sensitive to the number of runs chosen to complete. This combination of

variation and parameter uncertainty is somewhat conceptually difficult in terms of describing what that estimate of uncertainty of the advertising restrictions on consumption represents, possibly best described as 'the local variation in change to energy consumption related to proposed HFSS television advertising restrictions, accounting for variation in the duration of adverts and the parameter uncertainty on the social makeup of each area and on the effect of advertising exposure on energy consumption'.

NICE recommends that direct costs in HEE include both healthcare and personal social services, while it was only possible here to account for healthcare costs. Calculating variation in social services costs was considered but it was concluded that it would not be feasible. There are a variety of reasons why it would be very challenging, related to the social care system being much less centralised than the NHS. Spending responsibility is delegated from central government to local authorities, which have very different amounts of both need and resources, services are supplied to people with very variable individual needs, by largely private providers, with highly variable local cost inputs and without collated patient data or a national cost collection process. Social care savings have been estimated alongside healthcare savings in modelling studies for NCD prevention, for example Briggs *et al*(157) estimated salt reformulation targets saving £142m to the NHS and £725m to social care over 10 years – a factor of 5.1 – though a PA programme saved the NHS £10m and social care £14m, only 1.4 times the amount.

There are also broader limitations to the approach itself. These will be explored in greater detail in the Chapter 9, though they are briefly laid out here. Modelling relies on the quality of inputs. Here, they have had to be extensively estimated at the local authority level as

described in the above chapters. Each approach has its own strengths and limitations that have already been mentioned, though they include the small assumptions such as estimating the populations of people aged 90 and over, full synthetic estimation of BMI distributions and the need to rely on IMD quintile average disease epidemiology. All of these data are good approximations of their estimand, but by their very nature are difficult to validate independently. That disease epidemiology had to be estimated at the IMD quintile level was a necessary compromise as the run time for local-specific epidemiology was estimated to take five months, but remains a limitation for the reasons set out above. The modelling process itself also has limitations in terms of making simplifying assumptions on how real-world processes can be simplified and simulated while still capturing the core features of the phenomenon to provide valid modelled estimates. Due to run time, only 10 areas' uncertainty modelled and more work is required to understand if and how uncertainty on effects varies by area beyond the apparent positive relationship between scale of effect and scale of uncertainty.

Conclusions

This analysis modelled the locally-specific effect of proposed television advertising restrictions on the health and cost consequences for the cohort of children aged 5-17 years for each lower tier and unitary local authority in England. This used a locally-parameterised version of the PRIMEtime model that accounts for local differences in population, BMI distribution, disease epidemiology and intervention effect. It found that there were likely to

be wide variation in the effects between the areas that benefitted the most and least. Effects increased somewhat in line with deprivation, but with large overlaps between levels of deprivation.

HEE indicated that the intervention was cost-saving and therefore dominant compared to the business-as-usual situation. PSA indicated it was highly unlikely that the intervention would not be cost-saving from the given perspective. Health benefit (QALYs) and cost savings were both greatest for more deprived areas, but cost saving per QALY did not vary meaningfully by deprivation. A variety of conceptual issues in the HEE indicate that interpreting health economic evidence on population health interventions can be challenging.

This model captures some real-world local heterogeneity in outcomes from the proposed advertising restrictions that would not be foreseeable without performing locally-specific modelling. This intervention is likely to benefit the most deprived the most, though with a great deal of overlap between quintiles of deprivation, so, more work is required to better understand the predictors of benefit from the intervention. There is potential that increasing the window of time that adverts are restricted could greatly enhance the health and healthcare cost benefits of the restrictions. Consistent with the government's 'levelling up' agenda, this analysis provides evidence that the proposed television advertising restrictions would reduce geographic health inequalities, while being cost-saving to the government overall. The government should consider extending the watershed to prevent two thirds of the benefit from being lost to child viewing television after 9pm.

Chapter 9: Discussion and conclusion

Thesis overview

Chapter 1 introduces the general aims and context of the thesis. Namely, the aim was to produce a local authority level version of the PRIMEtime model for BMI-related disease burden in England. This was felt to be valuable to allow the impacts of national-level policies on health and healthcare costs to be estimated locally, with implications for health inequalities and the government's "levelling up" agenda more broadly. Chapter 2 is a multi-arm review of grey and peer-reviewed literature, of local-level NCD scenario models in the UK, their capabilities and compliance with best practice recommendations. This found four models previously developed to estimate the impacts of interventions for BMI, diet or PA on health and/ or healthcare cost outcomes. Two were simpler models that used Markov cohort model structures, and two were more complicated Microsimulation models with more representative disease states and better exploration of forms of uncertainty. The Microsimulations were more successful at meeting best practice recommendations.

Chapter 3 describes how the new local authority model will function using a PMSL, requiring local-level parameters to be estimated. Chapters 4-6 lay out the estimation of local-level parameters for the use in the new local level model. Chapter 4 uses a Bayesian modelling method(101,102) to estimate consistent sets of disease epidemiology for local areas and estimating case fatality rates from incidence and prevalence rates, by using these parameters' logical interdependencies. This estimated rates at the local authority IMD

2019(219) quintile level due to runtime limitation of the Bayesian model preventing each area having rates estimated separately.

Chapter 5 estimates the distributions of BMI for adults, aged 18 and over, in each upper tier local authority area in England, by age and sex. This built on previous small area estimation methods by taking granularity of Census data and quality of BMI data from the HSE (2018)(112) These were combined using GLM, which is well placed to cope with and detect the positive skew in the BMI data. These findings supported traditional notions of BMI being higher in older and more deprived areas. Chapter 6 demonstrates the process for estimating BMI distributions for children aged 2-17 years. This used the locally-specific NCMP data, which has an almost complete sample of measured BMI for children in reception year (aged 4-5 years) and year 6 (aged 10-11 years) in England. The HSE was then used to interpolate the BMI on missing years of age along the s-shaped curve observed as children grow.

Chapter 7 describes how local-level healthcare costs were estimated. National-level average disease specific costs to the NHS have been previously estimated.(47) New patient-level cost data from NHS Digital(186) were used to quantify how much NHS hospital trusts varied from one another at the level of the clinical specialty, then this variation was used to infer each trust's annual disease-specific costs from the national average.

PRIMEtime_local is then used to estimate the potential local-specific costs of proposed increases to the restrictions on advertising HFSS foods to children in Chapter 8. This found that more deprived areas would benefit more on average, but with large overlaps between quintiles of deprivation. Healthcare cost savings rose in line with health burden spared, but did not meaningfully vary by deprivation level. An assessment of modelling best practices is included in Appendix 2c.

Strengths and limitations

Strengths and limitations of scenario modelling as an approach

First of all, all scenario modelling involves certain limitations that are important to bear in mind. George Box famously said “all models are wrong, but some are useful”(220) or as framed by Peter Scarborough, “all models are simplifications of reality, and if they weren’t then they wouldn’t be useful”. This careful balance of simplifying a phenomenon into its key components to allow it to be quantitatively described without simplifying it beyond a useful level of reflection is the core skill of modelling. Doing this well depends on the understanding not only of the phenomenon and principles of modelling science but also the understanding of what elements of both need to be adhered more or less strictly, or where simplifying assumptions may be dramatic but unlikely to change the answer to the question being asked. Sometimes there is little choice on what data is available to parameterise models. Appropriate interpretation is also nuanced to each model and scenario, while communicating exactly what the modelling means and preventing those results from being applied in inappropriate contexts is another challenge altogether.

Appropriately projecting a current scenario into the future is dependent on multiple known unknowns and unknown unknowns, particularly around the logic model of the intervention. Policy options that are seriously being considered should generally have a robust logic model and evidence to parameterise that logic model. There is a level of direct logic – the

chain of causality between an intervention and its intended effects (and unintended effects) but gaining good-quality evidence to parameterise these chains can be more challenging for some types of intervention than others. All evidence is flawed to some extent due to real world constraints, but well designed and conducted RCTs that are less likely to offer biased estimates of an intervention effect are often not feasible to conduct for population-level interventions. For population approaches, evidence is more likely to rely on natural experiments and other observational designs, where true effects are harder to know than with RCTs and the impacts of different forms of uncertainty is harder to quantify. For example, it is possible to run an RCT of a weight loss referral service, but probably not for voluntary industry responsibility agreements for sugar reduction. Instead, certain assumptions have to be made to estimate the impacts of the policy on energy intakes.(51)

Other known unknowns include future trends to risk factor prevalence, disease incidence or case fatality – these things will be different in future, but we don't know when, how much or with what subgroup differences.

There is also another more indirect level of consideration to uncertainty in the intervention: these logic models are abstractions of reality that may or may not hold in the real world, even if they make sense and there is some evidence to support that model. These effects are due to emergent complex system dynamics of how people and institutions will behave. For example, will industry respond to sugary drinks tax by reformulating or by passing on the price increase? Will consumers inadvertently consume more sugar and calories elsewhere in their diets? Will they explicitly reject reformulation,(221) or even reject state paternalism and overcompensate for the intervention by consuming more in the way of harmful products elsewhere in their diets? These complex system dynamics go even further

into unknown unknowns that are definitionally impossible to anticipate. We must be mindful and realistic that this introduces a risk of interventions' effects being quite different from those anticipated. These models are therefore specifically scenario models, not predictive models. PRIMEtime_local does not aim to accurately predict the future but provide a framework to estimate potential health and cost implications of given population-level changes to BMI under the best evidence available. Generally, there is no way to then validate the quantitative outputs – even if sufficient time and resources were available to collect data, the 'signal' of the effects of population effects are often lost in the 'noise' of other influences and random variation, with difficulty attributing causality back to specific events.

Strengths and limitations of the PRIMEtime structure

There are some important limitations and simplifying assumptions to PRIMEtime that trade-off with its strengths as an NCD scenario model.

The PMSL structure assumes people sit in one of three disease states for each disease: 'no disease' 'disease' or 'dead from disease'. This provides simplicity and computational efficiency, but comes with trade-offs. Firstly, disease severity is not accounted for. This means that interventions improving quality of life from a disease or those reducing mortality of only a subgroup cannot easily be captured, though there would potentially be ways of incorporating it if sufficient data and computational power were available.

The PMSL also assumes the risk of incident cases of different diseases is independent for a given individual, apart from T2DM with IHD or stroke. In some cases this may be fair but it may also result in the artificial increase or decrease in rates of multimorbidity (concurrent NCD diagnoses for a single individual), with commensurate effects on the distribution of utility and costs within a population. This may increase or decrease average unit healthcare costs as multimorbidity is associated with higher costs(222) and there is an additional utility score for multimorbidity(98) that is difficult to account for in PRIMEtime without knowing the overall distribution of multimorbidity. An individual with high BMI is more likely to have any or all of the modelled diseases than someone with low BMI. This therefore risks failing to account for the additional utility burden of multimorbidity and a conservative estimate in the number of QALYs saved. If multimorbidity is independently associated with higher mortality rates, failure to account for that could increase or decrease the number of life-years lived, and therefore person years at risk and then further cases, costs and QALYs. There is therefore a risk that estimates overstate the numbers of cases prevented and QALYs saved (as additional mortality is not accounted for) and either overstating or understating costs, depending on whether the effect of reduced cases or increased unit costs of disease were greater. As risks are also not randomly distributed in the population (the prevalence of most risk factors increases with deprivation), it is also possible these effects are greater for more deprived areas. This is dealt with at least in part by varying all-cause mortality with disease-specific epidemiology.

The reliance on BMI in this PRIMEtime structure needs consideration. BMI is a well-accepted proxy for adiposity but the simple calculation of weight/height^2 leaves it open to appearing elevated where people's weight is raised for other reasons such as high skeletal muscle mass, so does not necessarily predict fat mass. The location of fat on the body also makes a

difference – central and visceral adiposity is more strongly associated with unfavourable health outcomes than fat around the pelvic girdle and thighs, which BMI does not account for.(223)

Treating T2DM as a disease outcome as well as a risk factor for IHD and stroke allows this particularly important example of non-independent probabilities to be accounted for. If diabetes risk and BMI risk were both applied independently to IHD or stroke, this would lead to double counting due to BMI also having a relationship with T2DM, so an adjustment factor is also applied to prevent this double counting.(47)

How cancers are modelled here is particularly limited. Similar issues could be argued to affect other diseases such as IHD or stroke, but these situations are less notable. A highly simplified conceptual model might be that cancers occur as either a case of limited disease ("curable") or terminal disease ("incurable"). Some cases of limited disease are successfully treated and a person may return to a near-baseline health after some time. Others will progress to a terminal stage despite treatment and other cases that are initially successfully treated may recur as incurable cases years later. As it stands, all cases are treated as prevalent, aggregating all cases into a single state ("Ever had cancer X") with case fatality relating to that total pool of cases, not only the cases of terminal disease. The effects on costs are small as average costs are calculated and framed as an average lifetime cost of disease. Estimated improvement in QALYs may be upwardly biased as people are not modelled to return to a remission state after curative treatment, so their utility weight remains at the disease state lifelong. Correcting for this in a three-state model would be challenging, and introducing a remission state and separating curative and non-curative cases would be preferable. This would require good data on cancer stages at first

presentation, and remission rates after radical treatment for curative intent. There would be considerable data requirements for this that put it outside the scope of this work, but it is a priority for future development.

PRIMEtime is a useful structure for overcoming various difficulties in public health modelling. NICE and ISPOR-SMDM have published extensive guidance(80,211,224) on implementing consistent and comparable approaches to the economic evaluation of healthcare innovations, though public health has additional complications such as longer-term impacts, the identification of effect in random variation ('signal-to-noise'), and greater consideration of wider societal benefits. The CHEERS checklist (ie. Consolidated Health Economic Evaluation Reporting Standards) represents reporting standards on economic evaluations more broadly(225) and health economic models have other suggested standards than those provided by ISPOR-SMDM.(226) The PMSL is well suited to achieve consistency with the NICE reference case, which now includes a broad acceptability of approaches to economic evaluation of health interventions, advises that the time horizon is able to include the important costs and effects, economic perspective can be flexible and may include non-health benefits. These changes allow public health evaluations to be compared on a more like-for-like basis against healthcare interventions, with a spectrum of approaches being accepted.(211,224,227) PRIMEtime achieves these aims through an efficient structure that allows multiple disease endpoints (and multiple risk factors, where appropriate) and cost impacts to be accounted for along a theoretical future lifetime. It does not privilege one particular disease state, instead intending to capture as much disease burden as feasible. It is a well-respected NCD modelling approach, having been used in many NCD prevention analyses.(39,49,50,52,53,86–88) It is specifically designed for scenario analysis through its parallel baseline and counterfactual scenario calculations, allowing the

testing of these scenarios under multiple assumptions. The local model described here adds the novel ability to model granular local heterogeneity, allowing the analysis of geographic equality impacts.

Strengths and limitations to national-level inputs

Using old (2000-3) US population-based EQ-5D response data and less outdated UK preference-based values to obtain utility data to populate the model may not be generalisable to a UK population. The inherent differences in national context may impact on how much weight individuals put on their level of disability. In addition, advances in medical care may mean that medical conditions now exert less disability than they did in the past and they may do so less in the UK with a socialised health system than in the USA, where healthcare resources may not be well allocated. In addition, the responses from the MEPS data may underrepresent either the most or least advanced cases of a disease due to severe disability or undiagnosed cases. However without an equivalent UK-based catalogue of EQ-5D responses by disease, these US/UK-based utilities remain the best accepted and most widely used source of utilities for the UK.

While prospective cohort studies constitute best evidence in the absence of RCTs, their limitations remain. Causation is not directly implied, whatever the association identified in the PCS, and an apparent effect may be due in part to both causation and additional association. The core assumptions is that correlation equals causation such that interventions to reduce BMI will necessarily reduce disease risk exposure. Therefore, to help support the case for causal relationships between BMI and the modelled disease outcomes,

Systematic Review Meta-Analysis studies specifically exploring other notions of causality such as Bradford-Hill criteria were preferentially selected.(47,228)

There are also alternative equations to quantify the relationship between energy intake and body weight. Those by Hall *et al*(92,93) have the most empirical backing and are generally considered a gold-standard, though the full sets of adult equations are not implemented as they rely on population heterogeneity that is difficult to capture in a population model. However, empirical work by Hall *et al* showed this linear model performed very closely to the full dynamic models.(92)

The use of national average RRs may conservatively bias results in terms of estimating lower impacts of interventions on disease burden (at least for some areas) and reducing modelled variation between areas. For example, Relative Risks between BMI and T2DM are higher for South Asian people and other ethnic minorities, or framed conversely, BMI category cut-offs differ by ethnicity (eg. what BMI categorises people as healthy or obese).(5) As RRs tend to be higher for non-white ethnicities, and non-white people are non-uniformly distributed into more urban areas, the estimates of effects for these areas are likely to be selectively and materially biased downwards. Therefore, it would be valuable in future to incorporate these dynamics into local authority modelling, via an approach such as ethnicity-weighted average Relative Risks for each area.

Strengths and limitations of PRIMEtime_local approach

PRIMEtime_local is flexible in offering the production of disease burden and healthcare cost point estimates for 315 areas or PSA for up to 10 selected areas. Both run time and programming limitations would make running PSA for all 315 areas challenging. Running 2000 iterations of the Monte Carlo analysis would take approximately 6 days and would produce quantities of data that were potentially challenging to handle. Population composition, epidemiology, risk factor distribution and intervention effect can all be allowed to vary locally or be standardised.

PIFs are influenced by any factor that can affect RR, risk factor prevalence or risk factor exposure intensity. These can be affected by factors such as baseline trends over time. Trend in BMI is not accounted for here due to uncertainty in trajectory even at the national level,(195) and inability to meaningfully judge change over time locally. Although the purpose of this local modelling is to estimate difference between areas, it is possible that if BMI rises and falls roughly in parallel between areas, there will be no overall effect. If, however, as with tobacco we see BMI falling in the least deprived first and most, then only after a lag more deprived areas progressively following suit, the impact of trend could be important.

The PMSL has challenges to modelling subgroups(229) in that the model must either be disaggregated or the model run multiple times for different subgroups. Both of these approaches require additional parameters to be estimated (such as those estimated through Chapters 4-7), with increasing numbers of subgroups introducing increasing challenges in terms of the computational and data requirements. For example, the data requirements for the disease epidemiology for local areas requires published local-area related epidemiological data– other subgroups such as by ethnicity do not have this data so it would

require additional assumptions to model them. Even the local authority level of geographic subgroups required a great deal of new data to be estimated in ways that may not be possible for other groups (particularly non-geographic groups, such as social grade or ethnicity). For example, budget data are provided and published by area, not (explicitly) by demographic groups. PRIMEtime_local offers a route to overcoming the limitation with the PMSL approach that modelling subgroups is more difficult, indeed, this model offers unique level of local area granularity for modelling the impacts of BMI interventions.

Costs are implemented as either single incident costs or ongoing prevalent costs, but not both for the same disease. This is because the national-level cost estimates that the local estimates rely on did not separate out these acute from chronic presentations, or separate by time (such as the first year of disease from subsequent years). Some diseases such as IHD or stroke may have higher first year costs than subsequent years' costs, reflecting treatment of an acute event then subsequent care. As first-year costs are included in subsequent years by the method, the inability to separate these out means that they may slightly overestimate the lifetime cost of disease where longer life expectancy leads these subsequent costs to get counted over longer periods. This overestimation is likely to be extremely small.

It was also not felt to be feasible to include Social Care costs ('personal social services) in line with NICE guidance.(211) Estimating local variation in Social Care costs would be challenging as the system is quite unalike the NHS cost system, in that while the NHS has a centralised funding model, Social Care is funded through a combination of local authority budgets and personal contributions,(230) to privately owned and profit-making providers.

There is no cost collection to detail variation in costs between areas, so the local-specific figures estimated for healthcare costs would not be achievable for social care.

Strengths and limitations to local-level model parameter inputs

In terms of the local model inputs, the strengths and limitations of each has been discussed in Chapters 4-7, though summarised again here.

The epidemiological data are mainly limited in terms of needing to aggregate to the area IMD quintile level rather than local authority level, losing an important opportunity to integrate locally-specific data. These parameters are ultimately based on locally-collected data, though how these are synthesised by the GBD study is opaque. Further, the case fatality rate has to be synthesised by a secondary piece of Bayesian modelling, Disbayes. This is closely related to the DISMOD models(104–106) that are well respected, having been used extensively for the GBD studies and comes with the important strength of creating consistent sets of incidence, prevalence and case fatality rates, though with the disadvantage that the methods are highly esoteric, making the process somewhat a black box. The expectation might have been for monotonically-increasing rates of disease burden in line with deprivation, but this was not observed. Instead, quintile-to-quintile ordering of deprivation was lost in the epidemiology. There is presumably wide within-quintile variation, which is also lost here, along with the opportunity to explore what factors lie beneath the associations between disease burden and deprivation (though it is important to note that deprivation here was total deprivation from the IMD, not health-related deprivation alone). It is important that there may be entirely valid reasons that deprivation

does not map onto disease burden neatly, for example, less deprived groups may be able to afford less healthy behaviours such as eating at takeaways, or spend longer commuting to high-wage but further away jobs than lower income people. Meanwhile, other people in the same IMD quintile areas may follow the more classical pattern of health behaviours and health following education, income, SES (bearing in mind none of these are perfect predictors of health either). It is unfortunately not easy to anticipate what impact the loss of within-quintile variation has on modelled results. This highlights the importance of locally-measured data and that the incorporation of some subgroup epidemiology remains an important strength relative to previous local modelling work.

The local BMI estimates are limited in that they rely on old Census data and indirect relationships between BMI and predictor variables in the HSE that don't vary locally. Validating these is also challenging, though multiple approaches to validating these estimates have been implemented and have shown good concordance with other data.

The estimates of local costs come with the limitations described above in Chapter 7, briefly that they rely on previous national estimates of disease costs, rely on cost data that lack good transparency or track record in published economic evaluations, and assume that variation in disease-level costs of treatment reflects specialty-level variation in costs of treatment within a hospital Trust.

None of the local level inputs had uncertainty distributions estimated to be incorporated into PRIMEtime's PSA. It is accepted that estimating uncertainty on small area estimates can be challenging, particularly with approaches using area microdata.(138) Likewise, the approach for estimating costs did not allow for uncertainty to be captured. Previously, uncertainty was included around costs in PRIMEtime using an arbitrary 20% variation

following comparable precedent.(231) This would be technically possible for child or adult BMI distributions, though there is no precedent for how much to vary either mean on SD of BMI. Attempts to include uncertainty around disease epidemiology parameters in the PSA introduces different issues that are discussed in Chapter 4. Both the GBD data and the Bayesian modelling tool used to estimate case fatality rates both offer sources of uncertainty, though introducing it into the modelling is not confounded by a lack of good estimates, but by the computational difficulty maintaining consistency between inter-dependent parameters.

Practical and academic context of the work

Comparison with other work

The most comparable local area modelling work in England is that of the DPP RoI model and the CVD prevention tool, mentioned above in the literature review in Chapter 2.

These are microsimulations, whose local populations are estimated by weighting a nationally-representative population from the HSE 2011, scaled to size by age, sex, ethnicity (three levels of 'white', 'Asian' and 'other'). The DPP models seven diseases (IHD, diabetic retinopathy, diabetic ulcers/ amputations, breast cancer, colorectal cancer, osteoarthritis and depression) while the CVD tool adds two more to these (Congestive Heart Failure and dementia). This is fewer than the 12 states modelled by PRIMEtime_local, but with specificity around diabetes states. Baseline disease epidemiology does not vary by geography. BMI distributions vary implicitly in the weighting process from the HSE, which

achieves an equivalent effect on local risk to the BMI distributions used in PRIMEtime_local. As with PRIMEtime_local, local effects can be varied by the user. Costs are implemented as national averages, lacking the local-specific reflections of cost submissions used in PRIMEtime_local (but are more granular by disease state than those implemented here, for example breaking down first year costs and subsequent costs). Unlike PRIMEtime_local, effects on the economy are modelled, implemented via the average cost of a day of lost output due to absence and the average cost of recruitment after a death. A range of exemplar scenarios with PSA are provided for the DPP tool, but it is not possible for the CVD prevention tool and no deterministic sensitivity analyses are presented for either.

Overall, PRIMEtime_local adds local specificity on these previous local authority models by adding greater variability of disease epidemiology and costs, and allows PSA. An increased breadth of disease states are included, but at the sacrifice of granularity in some of these disease states.

Since PRIMEtime_local was conceived of, parallel work has been done elsewhere on disaggregating PMSL models for analysing equity impacts of interventions in New Zealand. The method involves iteratively rescaling all-cause mortality rates (ACMR), and incidence and case fatality rates for IHD and stroke, to represent each of five quintiles of deprivation. At each cycle, consistency is maintained between the combination of the disaggregated models with the original aggregated model. A range of test scenarios were performed to reduce mortality rates for subgroups and to test the principles of the approach, then a scenario was used to estimate the impacts of a policy option on inequalities that involved substituting 59% of New Zealand's sodium chloride for potassium chloride (which does not have detrimental impact on blood pressure or IHD/ stroke risk). The salt scenario showed

that 1.3 times more benefits in terms of Health-Adjusted Life-Years would accrue to the most deprived quintile than the least deprived, and 1.1 times for ACMR. Interestingly, the proportional benefits in QALYs seen between most and least deprived quintiles in the television advertising restrictions scenario was also 1.3 times (Chapter 9, table 9.3).

Contribution to the research field

The literature review in Chapter 2 demonstrated that NCD modelling capabilities at the local authority level in England was limited, with only two really good examples of previous work. This set the precedent for developing a new model for flexibly estimating the local effects of BMI interventions for each local authority area. This review also took a robust approach to reviewing peer-reviewed and four arms of grey literature review. The approach taken for any of these four arms could be implemented or modified as a fair approach to searching grey literature on comparable topics.

New methods are presented for estimating small-area data, each built from strengths of existing data and previous methods. Adult BMI distributions are estimated to the highest degree of granularity yet for England, by 5-year age band, sex and local authority area (Census groups where necessary). These use a combination of synthetic estimation and microdata to produce locally-specific and granular data. New datasets of child BMI and costs of disease are presented for each lower tier local authority. Both of these also involved developing new methods, though neither is as detailed and technical as that for adult BMI distributions. IMD quintile level disease epidemiology offers a level of granularity not used in previous modelling work in the UK. This data also demonstrates in clear terms the complexity of dealing with local-specific data. That the order of disease burden was not

directly in line with deprivation demonstrates that the patterning we *expect* is not necessarily what we will find, so modelling to greater degrees of granularity is valuable in its own right, where locally-collected data can be used. The comparison between the HSE-derived estimates and ALS self-reported BMI categories showed there was probably materially important under-reporting of BMI in the ALS, (Chapter 4, figure 4.10 and discussion) showing that the cost-sparing approach of self-reporting health data comes with a major trade-off for quality.

The capability to model local areas flexibly for BMI interventions is a new contribution to the field of NCD modelling. PRIMEtime_local is up-to-date and allows easy PSA, and more granular cost and epidemiological data than previous modelling, with updated and highly granular BMI distributions estimated for children and adults. The scenario analysis examining the effect of proposed television advertising restrictions helps demonstrate why this level of granularity is valuable. The traditional approach of modelling the mean and it's uncertainty intervals would miss all the variation identified, but even an approach modelling on IMD quintile alone could easily misrepresent that each local authority in a given quintile could be expected to benefit as much as any other, while the work presented here demonstrates clearly that the assumption that benefits will fall in line with deprivation cannot be taken for granted, so assumptions that benefit of population-level interventions are likely to reduce inequalities if less agency is required to access the benefits(36) is only the beginning of the story. This broad observation is likely to be true but we see huge amounts of overlap between IMD quintiles, meaning some areas in both the highest and lowest deprivation quintiles only gain benefits around the same as those at the mean of the third (central) quintile. Therefore, even if inequalities fall overall, the patterning of benefit

beneath the averages cannot be assumed. Some areas that are harder to benefit may need special public health attention.

Implications for the real world

The aim of the model was to allow the variation of a national-level policy to be understood and to allow the effects of a local policy to be estimated using relevant local parameters. The former of these is demonstrated in the scenario in Chapter 8 on television advertising restrictions of HFSS foods. The latter was not demonstrated as a separate process, but the capabilities required for this local modelling were demonstrated through Chapter 8, namely the ability to model point estimate and PSA for an intervention, with the incorporation of outputs into HEE. With local authorities responsible for Public Health budgets as well as other areas of public policy that influence health, such as planning of takeaway outlets, active transport infrastructure or restricting outdoor HFSS product advertising, there could be many potential applications to this model. The government's new interest in population approaches and in 'levelling up' should make PRIMEtime_local well-placed to influence public policy decisions at the national level. With the proposed advertising restrictions currently on hold, there is a window of opportunity to influence thinking in government and contribute towards their implementation.

More broadly there are also implications of this work for how we think about data collection. The observation that locally-measured data matters suggests that our key national surveys might need to be scaled to allow local-level estimates to be produced. The idea that locally-specific data adds value is not new,(138) and nor is the trade-off between data requirement and cost in determining sample size a new consideration. However, the

increasing acknowledgement in the 21st century of the important implications of population heterogeneity may imply that the balance of factors is shifting in favour of increasing sample sizes. That local-level data are better measured than inferred is in line with complexity theory, which takes the perspective that the influences on system behaviour are too great in number and unpredictable in their relationships than we can (currently) measure or understand, so approaches that use estimation as a substitute for measurement will always be an inferior choice. In the absence of this locally-measured data, using sophisticated small area estimation methods such as these presented remains even more important.(138)

Areas for future development

First of all, there are some issues mentioned throughout this thesis that it would be of benefit to come back to. Incorporation of disease epidemiology into the modelling at the IMD quintile level is an important limitation and priority for future work. It would be useful to better understand the drivers of local variation in disease burden, as ways of better describing the complex adaptive system that health arises in. Equally, there has not been the scope to explore in great detail the associations between the local estimates and other area-level data, such as how BMI varies geographically by different measures of economic performance, elements of the IMD, or by RUC. An exploratory analysis of associations between healthcare cost estimates and other factors was not directly relevant here (Chapter 7) but again, this could prove very useful for better exploring system determinants of Trusts' costs. The combination of better understanding the system determinants of both

costs and disease burden could help towards the long-term need to understand demand, quality and efficiency in the NHS.

Which areas benefitted more or less in the television advertising restrictions scenario also needs greater examination. As mentioned, starting to understand which areas may be harder to reach for a given intervention would be of benefit to help address mismatch between need and effect. This would also help explore the causality of variation in effect arising within PRIMEtime_local as an aid to interpretation and future use.

It is regrettable that the development of a method for estimating local variation in social care costs was not feasible here. Incorporating social care costs is recommended by NICE and they are potentially of an order of magnitude great enough to meaningfully alter the conclusions of an economic analysis: previous analyses(157) using a comparable PRIMEtime structure have found these costs to be 1.4 and 5.1 times those for healthcare. In the case of the television advertising scenario, this was not a key issue, as the intervention was found to be cost effective (indeed, cost saving) even without social care costs included, but in other situations this may not be the case.

Two future uses for PRIMEtime_local have already been identified. Colleagues at University College London and Cambridge University have requested to fund the use of PRIMEtime_local, firstly to model the health equity implications of secular rises to child obesity since 1970, and secondly to model the locally-specific effects of local authority policies to restrict the expansion of takeaway outlets on BMI-related health. The model is well-placed for these academic uses and other related research, though the aim is that it could be useful to support local- and national-level decision-making.

Modelling variation also leads to reflection on what we are trying to achieve in the field of NCD scenario modelling and public health policy research more broadly. Scenario modelling has traditionally involved measuring and reporting the mean estimate and the confidence/ credibility we can place in that central estimate. The increasing understanding of geographic and demographic heterogeneity implies this may be no longer adequate for modern public and policy expectations. That different areas may have dramatically different outcomes may introduce the question of if it is enough modelling the mean and its *uncertainty* in public health interventions, rather than present the spread of *variation* of the modelled estimates. Modelling the level of local granularity presented here is certainly challenging, though there are a variety of other ways of modelling variation. For example, it would still be of benefits to estimate intervals representing population variation, or make estimates of difference in benefit for different population groups (eg. by deprivation or ethnicity) or geographic areas (deprived coastal areas versus wealthy suburban ones) if the requirements of local authority area modelling are too onerous. Variation analyses could be done via parallel Monte Carlo analyses – one parameterised with SEM of input parameters and the other with sample SDs. This would involve some important statistical and epistemological work (for example coping with uncertainty of the spread of data), but would constitute a major advance in understanding how our public health – and possibly healthcare – interventions may affect different the spread of risk across society, to help prioritise and plan the implementation of public policies.

Conclusion

This work has arisen from the intersection of multiple factors. There is an ongoing need to address the huge burden of BMI-related disease in England, alongside devolution of Public Health budgets to local authorities and a recent political expedience of reducing regional inequalities in the UK. To help understand how locally- or nationally-implemented BMI interventions may affect local areas differently, the aim was set to develop a local authority NCD model for BMI interventions.

Estimating the local parameters required was done via different methods. BMIs were estimated by age, sex and lower tier local authority. For children this was done by interpolating the years between ages 5 and 11 years in the NCMP – an almost complete set of measured BMI data for these two ages. For adults this was done via the building of a Generalise Linear Regression model to predict BMI in the HSE 2018 then using this model to produce synthetic estimates of BMI at the individual level in the Census microdata 2011. Healthcare costs were estimated at the upper tier level by using new PLICS cost collection data, then using the local variation identified to vary previous national-level cost estimates from previous modelling work. Disease epidemiology was derived from the GBD and the missing case fatality parameters estimated via the logical interdependencies between epidemiological parameters, using the Bayesian modelling tool, Disbayes.

A scenario was used to explore the outputs of PRIMEtime_local. This estimated the health and healthcare cost implications of restricting television advertising of high fat, sugar and salt foods, for each of 315 lower tier local authorities in England. This found that areas would broadly benefit more if they were in more deprived quintiles, but that there was a lot of overlap, with some lower deprivation areas achieving more benefit than some higher deprivation areas. People are estimated to benefit by 0.0222-0.521 QALYs per person, with

£11.7-£42.8 in healthcare cost savings per person, and the average difference in benefit for health was 1.30 times higher in the most deprived quintile of areas than lowest. If adverts are delayed rather than cancelled due to the policy, the effect may be reduced by two thirds, so extending the hours of the proposed restrictions would provide a major opportunity for the benefits of the policy to be protected, rather than burdening both the public sector and industry without delivering the intended public benefit.

This work demonstrates the importance of local-specific data collection and modelling. Population diversity is increasingly acknowledged by the public and policymakers, while the development of complexity theory leads us to realise this diversity may have unpredictable impacts on the patterning of health determinants. In turn, the distribution of benefits of Public Health policies may be difficult to anticipate, and the importance of good-quality locally-collected data and modelling capabilities is ever more apparent.

References

1. Newton JN, Briggs ADM, Murray CJL, Dicker D, Foreman KJ, Wang H, et al. Changes in health in England, with analysis by English regions and areas of deprivation, 1990â€“2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet [Internet]. 2015 [cited 2023 Feb 21];386:2257–74. Available from: http://dx.doi.org/10.1016/

2. Steel N, Ford JA, Newton JN, Davis ACJ, Vos T, Naghavi M, et al. Changes in health in the countries of the UK and 150 English Local Authority areas 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet [Internet]. 2018 Nov 3 [cited 2023 Feb 22];392(10158):1647–61. Available from: http://www.thelancet.com/article/S0140673618322074/fulltext

3. Health matters: obesity and the food environment - GOV.UK [Internet]. [cited 2023 Feb 21]. Available from: https://www.gov.uk/government/publications/health-matters-obesity-and-the-food-environment/health-matters-obesity-and-the-food-environment--2

4. National Institute for Health and Care Excellence. NICE Recommendations: Obesity: identification, assessment and management. Guidance (CG189). [Internet]. NICE; 2014 [cited 2023 Feb 21]. Available from: https://www.nice.org.uk/guidance/cg189

5. Caleyachetty R, Barber TM, Mohammed NI, Cappuccio FP, Hardy R, Mathur R, et al. Ethnicity-specific BMI cutoffs for obesity based on type 2 diabetes risk in England: a population-based cohort study. Lancet Diabetes Endocrinol. 2021 Jul 1;9(7):419–26.

6. Obesity Profile, Public Health England Public Health Profiles 2021, https://fingertips.phe.org.uk/profile/national-child-measurement-programme, Crown copyright 2021.

7. National Child Measurement Programme, England 2019/20 School Year - NDRS [Internet]. [cited 2023 Feb 21]. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/national-child-measurement-programme/2019-20-school-year

8. di Cesare M, Sorić M, Bovet P, Miranda JJ, Bhutta Z, Stevens GA, et al. The epidemiological burden of obesity in childhood: a worldwide epidemic requiring urgent action. BMC Med [Internet]. 2019 Nov 25 [cited 2023 Feb 21];17(1). Available from: /pmc/articles/PMC6876113/

9. Esposito K, Giugliano D. The metabolic syndrome and inflammation: association or causation? Nutrition, Metabolism and Cardiovascular Diseases. 2004 Oct 1;14(5):228–32.

10. How can local authorities reduce obesity? Insights from NIHR research [Internet]. 2022 [cited 2023 Feb 23]. Available from: https://evidence.nihr.ac.uk/how-local-authorities-can-reduce-obesity/

11. Tackling obesity: empowering adults and children to live healthier lives - GOV.UK [Internet]. [cited 2023 Feb 21]. Available from: https://www.gov.uk/government/publications/tackling-obesity-government-strategy/tackling-obesity-empowering-adults-and-children-to-live-healthier-lives

12. Childhood Obesity: A Plan for Action, Chapter 2 [Internet]. 2018 Jun. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/718903/childhood-obesity-a-plan-for-action-chapter-2.pdf

13. Theis DRZ, White M. Is Obesity Policy in England Fit for Purpose? Analysis of Government Strategies and Policies, 1992–2020. Milbank Q [Internet]. 2021 Mar 1 [cited 2023 Feb 21];99(1):126. Available from: /pmc/articles/PMC7984668/

14. Butler AR, Astbury NM, Goddard L, Hajizadeh A, Seeber P, Crawley B, et al. Setting the top 10 priorities for obesity and weight-related research (POWeR): a stakeholder priority setting process. BMJ Open [Internet]. 2022 [cited 2023 Feb 21];12:58177. Available from: http://dx.doi.org/10.1136/bmjopen-2021-058177

15. World Obesity Federation Policy Dossiers [Internet]. World Obesity Federation. [cited 2023 Feb 21]. Available from: https://www.worldobesity.org/resources/policy-dossiers

16. Nuffield Bioethics Council. Public health: ethical issues [Internet]. 2007 [cited 2023 Feb 21]. Available from: http://www.nuffieldbioethics.org

17. Rose G. Sick individuals and sick populations. Int J Epidemiol [Internet]. 1985 [cited 2023 Feb 21];14(1):32–8. Available from: https://pubmed.ncbi.nlm.nih.gov/3872850/

18. Sally Davies. Time to Solve Childhood Obesity: An Independent Report by the Chief Medical Officer [Internet]. 2019 [cited 2023 Feb 21]. Available from: https://www.gov.uk/government/publications/time-to-solve-childhood-obesity-cmo-special-report

19. Butland B, Jebb S, Kopelman P, Mcpherson K. Tackling Obesities: Future Choices-Project Report 2 nd Edition Government Office for Science Foresight Tackling Obesities: Future Choices-Project report. [cited 2023 Feb 21]; Available from: www.foresight.gov.uk

20. Hammond RA, Cdc for. Complex Systems Modeling for Obesity Research. 2009 [cited 2023 Feb 21]; Available from: http://www.cdc.gov/pcd/issues/2009/jul/09_0017.htm.Accessed[date].

21. Cookson R, Griffin S, Norheim OF, Culyer AJ, Chalkidou K. Distributional Cost-Effectiveness Analysis Comes of Age. Value in Health [Internet]. 2021 Jan 1 [cited 2023 Feb 21];24(1):118. Available from: /pmc/articles/PMC7813213/

22. NHS England. NHS Five Year Forward View [Internet]. 2014 [cited 2022 May 18]. Available from: https://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf

23. NHS England. The NHS Long Term Plan [Internet]. 2019 [cited 2023 Feb 21]. Available from: www.longtermplan.nhs.uk

24. Wanless D. Securing our Future Health: Taking a Long-Term View Final Report [Internet]. 2002 [cited 2023 Feb 21]. Available from: https://www.yearofcare.co.uk/sites/default/files/images/Wanless.pdf

25. HM Revenue and Customs. Policy Paper: The Soft Drinks Industry Levy [Internet]. 2016 [cited 2023 Feb 21]. Available from: https://www.gov.uk/government/publications/soft-drinks-industry-levy/soft-drinks-industry-levy

26. Briggs ADM, Mytton OT, Kehlbacher A, Tiffin R, Elhussein A, Rayner M, et al. Health impact assessment of the UK soft drinks industry levy: a comparative risk assessment modelling study. Lancet Public Health [Internet]. 2017 Jan 1 [cited 2023 Feb 22];2(1):e15–22. Available from: http://www.thelancet.com/article/S2468266716300378/fulltext

27. Scarborough P, Adhikari V, Harrington RA, Elhussein A, Briggs A, Rayner M, et al. Impact of the announcement and implementation of the UK Soft Drinks Industry Levy on sugar content, price, product size and number of available soft drinks in the UK, 2015-19: A controlled interrupted time series analysis. PLoS Med [Internet]. 2020 Feb 11 [cited 2023 Feb 22];17(2). Available from: https://pubmed.ncbi.nlm.nih.gov/32045418/

28. Boris Johnson Says Brits Must Lose Weight to Fight COVID-19 | Time [Internet]. [cited 2023 Feb 21]. Available from: https://time.com/5872175/boris-johnson-weight-loss-coronavirus/

29. Health and Social Care Act 2012 [Internet]. Department of Health, HM Government; 2012. Available from: https://www.legislation.gov.uk/ukpga/2012/7/contents/enacted

30. Department of Communities and Local Government. A plain English guide to the Localism Bill - Update. HM Government [Internet]. 2011 [cited 2023 Feb 21]; Available from: www.communities.gov.uk

31. HM Government. White Paper: Levelling Up the United Kingdom. Executive Summary. 2022.

32. Local Government Association. What is local government? Local government touches the lives of everybody, every day. https://www.local.gov.uk/about/what-local-government.

33. Understand how your council works: Types of council - GOV.UK [Internet]. [cited 2023 Feb 21]. Available from: https://www.gov.uk/understand-how-your-council-works

34. Gray AM, Clarke PM, Wolstenholme JL, Wordsworth S. Applied Methods of Cost-effectiveness Analysis in Healthcare (Handbooks in Health Economic Evaluation). 2010 [cited 2023 Feb 22];328. Available from: http://www.amazon.co.uk/Cost-effectiveness-Analysis-Healthcare-Handbooks-Evaluation/dp/0199227284

35. Archard D. Public health ten years on - The Nuffield Council on Bioethics [Internet]. [cited 2023 Feb 22]. Available from: https://www.nuffieldbioethics.org/blog/public-health-ten-years

36. Adams J, Mytton O, White M, Monsivais P. Why Are Some Population Interventions for Diet and Obesity More Equitable and Effective Than Others? The Role of Individual Agency. PLoS Med. 2016 Apr;

37. Parliamentary Office for Science and Technology. POST Research evidence and policy-making: increasing demand, publication speed and public scrutiny [Internet]. [cited 2023 Feb 6]. Available from: https://post.parliament.uk/research-evidence-and-policy-making-increasing-demand-publication-speed-and-public-scrutiny/

38. Bermingham R. Understanding research evidence. Parliamentary Office for Science and Technology. [Internet]. POST Resources. Parliamentary Office for Science and Technology; 2020 Sep [cited 2023 Feb 22]. Available from: https://post.parliament.uk/understanding-research-evidence/

39. Hartmann-Boyce J, Cobiac LJ, Theodoulou A, Oke JL, Butler AR, Scarborough P, et al. Weight regain after behavioural weight management programmes and its impact on quality of life and cost effectiveness: Evidence synthesis and health economic analyses. Diabetes Obes Metab [Internet]. 2023 Feb 1 [cited 2023 Feb 22];25(2):526–35. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/dom.14895

40. Astbury NM, Piernas C, Hartmann-Boyce J, Lapworth S, Aveyard P, Jebb SA. A systematic review and meta-analysis of the effectiveness of meal replacements for weight loss. Obes Rev [Internet]. 2019 Apr 1 [cited 2023 Feb 22];20(4):569–87. Available from: https://pubmed.ncbi.nlm.nih.gov/30675990/

41. Bandy LK, Scarborough P, Harrington RA, Rayner M, Jebb SA. Reductions in sugar sales from soft drinks in the UK from 2015 to 2018. BMC Med [Internet]. 2020 Jan 13 [cited 2023 Feb 22];18(1). Available from: https://pubmed.ncbi.nlm.nih.gov/31931800/

42. Rutter H, Savona N, Glonti K, Bibby J, Cummins S, Finegood DT, et al. The need for a complex systems model of evidence for public health. The Lancet. 2017 Dec 9;390(10112):2602–4.

43. Ogilvie D, Adams J, Bauman A, Gregg EW, Panter J, Siegel KR, et al. Using natural experimental studies to guide public health action: turning the evidence-based medicine paradigm on its head. J Epidemiol Community Health (1978) [Internet]. 2020 [cited 2023 Feb 6];74(2). Available from: https://pubmed.ncbi.nlm.nih.gov/31744848/

44. Briggs ADM, Wolstenholme J, Blakely T, Scarborough P. Choosing an epidemiological model structure for the economic evaluation of non-communicable disease public health interventions. Popul Health Metr [Internet]. 2016 May 4 [cited 2023 Feb 22];14(1):1–12. Available from: https://pophealthmetrics.biomedcentral.com/articles/10.1186/s12963-016-0085-1

45. Briggs A, Scarborough P, Smith A. Modelling in public health. In: Public Health Intelligence: Issues of Measure and Method [Internet]. Springer International Publishing; 2016 [cited 2023 Feb 22]. p. 67–90. Available from: https://link.springer.com/chapter/10.1007/978-3-319-28326-5_4

46. Barendregt JJ, van Oortmarssen GJ, van Hout BA, van den Bosch JM, Bonneux L. Coping with multiple morbidity in a life table. http://dx.doi.org/101080/08898489809525445 [Internet]. 2009 [cited 2023 Feb

22];7(1):29–49. Available from: https://www.tandfonline.com/doi/abs/10.1080/08898489809525445

47. Cobiac LJ, Law C, Scarborough P. PRIMEtime: an epidemiological model for informing diet and obesity policy. medRxiv [Internet]. 2022 May 21 [cited 2023 Feb 22];2022.05.18.22275284. Available from: https://www.medrxiv.org/content/10.1101/2022.05.18.22275284v1

48. Scarborough P, Kaur A, Cobiac L, Owens P, Parlesak A, Sweeney K, et al. Eatwell Guide: modelling the dietary and cost implications of incorporating new sugar and fibre guidelines. BMJ Open [Internet]. 2016 Dec 1 [cited 2023 Feb 22];6(12):e013182. Available from: https://bmjopen.bmj.com/content/6/12/e013182

49. Aminde LN, Phung HN, Phung D, Cobiac LJ, Veerman JL. Dietary Salt Reduction, Prevalence of Hypertension and Avoidable Burden of Stroke in Vietnam: Modelling the Health and Economic Impacts. Front Public Health [Internet]. 2021 Jun 4 [cited 2023 Feb 22];9:682975. Available from: /pmc/articles/PMC8213032/

50. Aminde LN, Cobiac LJ, Phung D, Phung HN, Veerman JL. Avoidable burden of stomach cancer and potential gains in healthy life years from gradual reductions in salt consumption in Vietnam, 2019-2030: a modelling study. Public Health Nutr [Internet]. 2022 [cited 2023 Feb 22]; Available from: https://pubmed.ncbi.nlm.nih.gov/35983611/

51. Amies-Cull B, Briggs ADM, Scarborough P. Estimating the potential impact of the UK government's sugar reduction programme on child and adult health: modelling study. BMJ [Internet]. 2019 Apr 17 [cited 2023 Feb 23];365. Available from: https://www.bmj.com/content/365/bmj.l1417

52. Mizdrak A, Telfer K, Direito A, Cobiac LJ, Blakely T, Cleghorn CL, et al. Health Gain, Cost Impacts, and Cost-Effectiveness of a Mass Media Campaign to Promote Smartphone Apps for Physical Activity: Modeling Study. JMIR Mhealth Uhealth 2020;8(6):e18014 https://mhealth.jmir.org/2020/6/e18014 [Internet]. 2020 Jun 11 [cited 2023 Feb 22];8(6):e18014. Available from: https://mhealth.jmir.org/2020/6/e18014

53. van der Deen FS, Wilson N, Cleghorn CL, Kvizhinadze G, Cobiac LJ, Nghiem N, et al. Impact of five tobacco endgame strategies on future smoking prevalence, population health and health system costs: two modelling studies to inform the tobacco endgame. Tob Control [Internet]. 2018 [cited 2023 Feb 22];27:278–86. Available from: http://dx.doi.org/10.1136/tobaccocontrol-2016-053585

54. Lajoie J. Understanding the Measurement of Global Burden of Disease. National Collaborating Centre for Infectious Diseases [Internet]. 2013 [cited 2023 Feb 22]; Available from: www.ccnmi.ca

55. Pettitt D, Raza S. The Limitations of QALY: A Literature Review. J Stem Cell Res Ther. 2016;06(04).

56. von Neumann J, Morgenstern O. Theory of games and economic behavior. Theory of Games and Economic Behavior (60th Anniversary Commemorative Edition). Princeton, New Jersey: Princeton University Press; 1944.

57. Thurstone LL. The Measurement of Values. Chicago: The University of Chicago Press; 1959.

58. George Torrance by W, Thomas WH, Sackett DL, Torrance GW. A Utility Maximization Model for Evaluation of Health Care Programs. Health Serv Res [Internet]. 1972 [cited 2023 Feb 22];7(2):118. Available from: /pmc/articles/PMC1067402/?report=abstract

59. Sassi F. Calculating QALYs, comparing QALY and DALY calculations. Health Policy Plan [Internet]. 2006 Sep 1 [cited 2023 Jun 20];21(5):402–8. Available from: https://dx.doi.org/10.1093/heapol/czl018

60. Unal B, Capewell S, Critchley JA. Coronary heart disease policy models: A systematic review. BMC Public Health [Internet]. 2006 Aug 18 [cited 2023 Feb 22];6(1):1–10. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-6-213

61. Grover A, Joshi A. An Overview of Chronic Disease Models: A Systematic Literature Review. Glob J Health Sci [Internet]. 2015 [cited 2023 Feb 22];7(2). Available from: doi:10.5539/gjhs.v7n2p210

62. Levy DT, Mabry PL, Wang YC, Gortmaker S, Huang TTK, Marsh T, et al. Simulation models of obesity: a review of the literature and implications for research and policy. Obesity Reviews [Internet]. 2011 May 1 [cited 2023 Feb 22];12(5):378–94. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1467-789X.2010.00804.x

63. Li Y, Lawley MA, Siscovick DS, Zhang D, Pagán JA. Agent-Based Modeling of Chronic Diseases: A Narrative Review and Future Research Directions. Prev Chronic Dis [Internet]. 2016 [cited 2023 Feb 22];13:150561. Available from: www.cdc.gov/pcd/issues/2016/15_0561.htm

64. Cochrane Collaboration. Cochrane Handbook of Systematic Reviews of Interventions (v6.3) [Internet]. Cochrane Training. [cited 2023 Feb 22]. Available from: https://training.cochrane.org/handbook/PDF/v6.3

65. Mulrow CD. Systematic Reviews: Rationale for systematic reviews. BMJ [Internet]. 1994 Sep 3 [cited 2023 Feb 22];309(6954):597–9. Available from: https://www.bmj.com/content/309/6954/597

66. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ [Internet]. 2009 Jul 21 [cited 2023 Feb 22];339. Available from: https://www.bmj.com/content/339/bmj.b2700

67. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ [Internet]. 2015 Jan 2 [cited 2023 Feb 22];349. Available from: https://www.bmj.com/content/349/bmj.g7647

68. Lhachimi SK, Nusselder WJ, Smit HA, van Baal P, Baili P, Bennett K, et al. DYNAMO-HIA--a Dynamic Modeling tool for generic Health Impact Assessments. PLoS One [Internet]. 2012 May 10 [cited 2023 Feb 22];7(5). Available from: https://pubmed.ncbi.nlm.nih.gov/22590491/

69. Boshuizen HC, Lhachimi SK, van Baal PHM, Hoogenveen RT, Smit HA, Mackenbach JP, et al. The DYNAMO-HIA Model: An Efficient Implementation of a Risk Factor/Chronic Disease Markov Model for Use in Health Impact Assessment (HIA). Demography. 2012 Nov 1;49(4):1259–83.

70. Jaccard A. Modelling the economics of chronic disease with the EConDA tool and the UKHF microsimulation model. Eur J Public Health [Internet]. 2016 [cited 2023 Feb 22];26. Available from: https://academic.oup.com/eurpub/article/26/suppl_1/ckw171.046/2448931

71. Kypridemos C. Modelling the effectiveness and equity of primary prevention policies in England: a stochastic dynamic microsimulation for the joint prevention of non communicable diseases. University of Liverpool;

72. Briggs A. Estimating and comparing the cost-effectiveness of primary prevention policies affecting diet and physical activity in England. Vol. 14, Population Health Metrics. University of Oxford; 2017.

73. Breeze PR, Thomas C, Squires H, Brennan A, Greaves C, Diggle P, et al. SPHR Diabetes Prevention Model: Detailed Description of Model Background, Methods, Assumptions and Parameters. 2015 Jun 15 [cited 2023 Feb 22]; Available from: http://www.sheffield.ac.uk/scharr/sections/heds/discussion-papers/15-01-1.472126

74. Cardiovascular Disease Prevention Return on Investment Tool [Internet]. 2018 [cited 2023 Feb 22]. Available from: www.facebook.com/PublicHealthEngland

75. Kefford B. About British Library EThOS - search and order theses online [Internet]. British Library. The British Library; [cited 2023 Feb 22]. Available from: https://www.bl.uk/ethos-and-theses/about-ethos

76. Grey Matters: a practical tool for searching health-related grey literature | CADTH [Internet]. [cited 2023 Feb 22]. Available from: https://www.cadth.ca/grey-matters-practical-tool-searching-health-related-grey-literature

77. Haddaway NR, Collins AM, Coughlin D, Kirk S, Wray KB. The Role of Google Scholar in Evidence Reviews and Its Applicability to Grey Literature Searching. 2015 [cited 2023 Feb 22]; Available from: http://www.bangor.ac.uk/

78. Godin K, Stapleton J, Kirkpatrick SI, Hanning RM, Leatherdale ST. Applying systematic review search methods to the grey literature: a case study examining guidelines for school-based breakfast programs in Canada. 2015.

79. Adams J, Hillier-Brown FC, Moore HJ, Lake AA, Araujo-Soares V, White M, et al. Searching and synthesising 'grey literature' and 'grey information' in public health: critical reflections on three case studies. Syst Rev. 2016;5(164).

80. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--1. Value Health [Internet]. 2012 [cited 2023 Feb 22];15(6):796–803. Available from: https://pubmed.ncbi.nlm.nih.gov/22999128/

81. Matrix TKMG, National Institute for Health Excellence. Estimating Return on Investment for interventions and strategies to increase physical activity - Technical Report. 2006.

82. Public Health England. User Guide: weight management economic assessment tool (version 2) [Internet]. 2016 [cited 2023 Mar 7]. Available from: https://webarchive.nationalarchives.gov.uk/ukgwa/20170110165409/https://www.noo.org.uk/visualisation

83. 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 [Internet]. 2007 Jul 19 [cited 2023 Feb 22];335(7611):136. Available from: https://www.bmj.com/content/335/7611/136

84. NCDprime: Modelling the impact of national policies on noncommunicable disease (NCD) mortality using PRIME: a policy scenario modelling tool NCDprime [Internet]. Copenhagen; 2019 [cited 2023 Feb 23]. Available from: http://www.euro.who.int/pubrequest

85. Scarborough P, Harrington RA, Mizdrak A, Zhou LM, Doherty A. The Preventable Risk Integrated ModEl and Its Use to Estimate the Health Impact of Public Health Policy Scenarios. Scientifica (Cairo) [Internet]. 2014 [cited 2023 Feb 23];2014:1–21. Available from: https://pubmed.ncbi.nlm.nih.gov/25328757/

86. Cobiac LJ, Scarborough P. Translating the WHO 25×25 goals into a UK context: the PROMISE modelling study. BMJ Open [Internet]. 2017 Apr 1 [cited 2023 Feb 22];7(4):e012805. Available from: https://bmjopen.bmj.com/content/7/4/e012805

87. Cobiac LJ, Scarborough P. Modelling the health co-benefits of sustainable diets in the UK, France, Finland, Italy and Sweden. Eur J Clin Nutr [Internet]. 2019 Apr 1 [cited 2023 Feb 22];73(4):624–33. Available from: https://pubmed.ncbi.nlm.nih.gov/30755710/

88. Briggs ADM, Cobiac LJ, Wolstenholme J, Scarborough P. PRIMEtime CE: A multistate life table model for estimating the cost-effectiveness of interventions affecting diet and physical activity. BMC Health Serv Res [Internet]. 2019 Jul 16 [cited 2023 Feb 23];19(1):1–19. Available from: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-019-4237-4

89. Blakely T, Moss R, Collins J, Mizdrak A, Singh A, Carvalho N, et al. Proportional multistate lifetable modelling of preventive interventions: concepts, code and worked examples. Int J Epidemiol [Internet]. 2020 Oct 1 [cited 2023 Feb 23];49(5):1624–36. Available from: https://academic.oup.com/ije/article/49/5/1624/5920732

90. Barendregt JJ, Veerman JL. Categorical versus continuous risk factors and the calculation of potential impact fractions. J Epidemiol Community Health (1978) [Internet]. 2010 Mar 1 [cited 2023 Feb 23];64(3):209–12. Available from: https://jech.bmj.com/content/64/3/209

91. Cobiac LJ, Scarborough P. Translating the WHO 25×25 goals into a UK context: the PROMISE modelling study (supplementary material). 2017 [cited 2023 Mar 6]; Available from:

https://bmjopen.bmj.com/content/bmjopen/7/4/e012805/DC2/embed/inline-supplementary-material-2.pdf?download=true

92. Hall KD, Sacks G, Chandramohan D, Chow CC, Wang YC, Gortmaker SL, et al. Quantification of the effect of energy imbalance on bodyweight. The Lancet [Internet]. 2011 Aug 27 [cited 2023 Feb 22];378(9793):826–37. Available from: http://www.thelancet.com/article/S014067361160812X/fulltext

93. Hall KD, Butte NF, Swinburn BA, Chow CC. Dynamics of childhood growth and obesity: development and validation of a quantitative mathematical model. Lancet Diabetes Endocrinol [Internet]. 2013 Oct 1 [cited 2023 Feb 22];1(2):97–105. Available from: http://www.thelancet.com/article/S2213858713700512/fulltext

94. Ezzati M, Lopez AD, Rodgers A, Murray CJL. Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors (Volume 1). 2004.

95. Renehan AG, Soerjomataram I, Tyson M, Egger M, Zwahlen M, Coebergh JW, et al. Incident cancer burden attributable to excess body mass index in 30 European countries. Int J Cancer [Internet]. 2010 Feb 1 [cited 2023 Feb 22];126(3):692–702. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/ijc.24803

96. Abbafati C, Abbas KM, Abbasi-Kangevari M, Abd-Allah F, Abdelalim A, Abdollahi M, et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet [Internet]. 2020 Oct 17 [cited 2023 Feb 22];396(10258):1223–49. Available from: http://www.thelancet.com/article/S0140673620307522/fulltext

97. Global Health Data Exchange (GHDx) [Internet]. Institute for Health Metrics and Evaluation; [cited 2023 Feb 22]. Available from: https://ghdx.healthdata.org/

98. Sullivan PW, Slejko JF, Sculpher MJ, Ghushchyan V. Catalogue of EQ-5D Scores for the United Kingdom. http://dx.doi.org/101177/0272989X11401031 [Internet]. 2011 Mar 21 [cited 2023 Feb 22];31(6):800–4. Available from: https://journals.sagepub.com/doi/10.1177/0272989X11401031

99. Sullivan PW, Ghushchyan V. Preference-Based EQ-5D Index Scores for Chronic Conditions in the United States. [cited 2023 Feb 22]; Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2634296/pdf/nihms-81406.pdf

100. Dolan P. Modeling Valuations for EuroQol Health States. Med Care [Internet]. 1997 [cited 2023 Feb 23];35(11):1095–108. Available from: https://www.jstor.org/stable/3767472

101. Jackson C. Bayesian estimation of chronic disease epidemiology from incomplete data: the disbayes package • disbayes [Internet]. GitHub. [cited 2023 Feb 22]. Available from: https://chjackson.github.io/disbayes/articles/disbayes.html

102. Jackson C, Zapata-Diomedi B, Woodcock J. Bayesian multistate modelling of incomplete chronic disease burden data. ArXiv [Internet]. 2021 Nov 28 [cited 2023 Feb 22];arXiv:2111.14100. Available from: https://ui.adsabs.harvard.edu/abs/2021arXiv211114100J/abstract

103. Thatcher R, Kannisto V, Andreev K. The Survivor Ratio Method for Estimating Numbers at High Ages. Demogr Res [Internet]. 2002 Jan 4 [cited 2023 Feb 22];6:1–18. Available from: https://www.demographic-research.org/volumes/vol6/1/default.htm

104. Flaxman AD, Vos T, Murray CJL. An integrative metaregression framework for descriptive epidemiology [Internet]. Seattle: University of Washington Press; 2015 [cited 2023 Feb 22]. 217 p. Available from: https://uwapress.uw.edu/book/9780295991849/an-integrative-metaregression-framework-for-descriptive-epidemiology

105. Flaxman AD, Vos T, Murray CJL. An Integrative Metaregression Framework for Descriptive Epidemiology Contents. [cited 2023 Feb 22]; Available from: https://pypi.org/project/dismod-mr

106. Barendregt JJ, van Oortmarssen GJ, Vos T, Murray CJL. A generic model for the assessment of disease epidemiology: The computational basis of DisMod II. Popul Health Metr [Internet]. 2003 Apr 14 [cited 2023 Feb 22];1(1):1–8. Available from: https://pophealthmetrics.biomedcentral.com/articles/10.1186/1478-7954-1-4

107. Kypridemos C, Guzman-Castillo M, Hyseni L, Hickey GL, Bandosz P, Buchan I, et al. Estimated reductions in cardiovascular and gastric cancer disease burden through salt policies in England: an IMPACTNCD microsimulation study. BMJ Open [Internet]. 2017 Jan 1 [cited 2023 Feb 22];7(1):e013791. Available from: https://bmjopen.bmj.com/content/7/1/e013791

108. Lv JC, Zhang LX. Prevalence and Disease Burden of Chronic Kidney Disease. In: Advances in Experimental Medicine and Biology [Internet]. New York: Springer New York LLC; 2019 [cited 2023 Feb 22]. p. 3–15. Available from: https://link.springer.com/chapter/10.1007/978-981-13-8871-2_1

109. Population estimates for the UK, England and Wales, Scotland and Northern Ireland [Internet]. Office for National Statistics. 2019 [cited 2023 Feb 23]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2018

110. L Murray CJ. The Global Burden of Disease Study at 30 years. Nat Med [Internet]. 2022 [cited 2023 Mar 6];28:2019–2016. Available from: https://doi.org/10.1038/s41591-022-01990-1

111. Steel N, Ford JA, Newton JN, Davis ACJ, Vos T, Naghavi M, et al. Changes in health in the countries of the UK and 150 English Local Authority areas 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016 (Supplementary Appendix). The Lancet [Internet]. 2018 Nov 3 [cited 2023 Feb 22];392(10158):1647–61. Available from: http://www.thelancet.com/article/S0140673618322074/fulltext

112. NatCen Social Research, University College London, Department of Epidemiology and Public Health. (2023). Health Survey for England, 2019. UK Data Service. SN: 8860, DOI: 10.5255/UKDA-SN-8860-1 [Internet]. [cited 2023 Feb 22]. Available from: https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=8860

113. University of Cambridge, MRC Epidemiology Unit, NatCen Social Research. (2023). National Diet and Nutrition Survey Years 1-11, 2008-2019. 19th Edition. UK Data

Service. SN: 6533, DOI: 10.5255/UKDA-SN-6533-19 [Internet]. [cited 2023 Feb 22]. Available from: https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=6533

114. Sport England. (2020). Active Lives Survey, 2017-2018. UK Data Service. SN: 8651, http://doi.org/10.5255/UKDA-SN-8651-1.

115. Rahman A, Harding A. Small Area Estimation, section 2.2.3. In: Rahman A, Harding A, editors. Small area estimation and microsimulation modelling [Internet]. 2019 [cited 2023 Feb 22]. p. 495. Available from: https://www.routledge.com/Small-Area-Estimation-and-Microsimulation-Modeling/Rahman-Harding/p/book/9780367261269

116. Bajekal M, Scholes S. Synthetic Estimation of Healthy Lifestyles Indicators: Stage 1 Report IMPACTsec-analysis of drivers of differentail falls in CVD mortality for socioeconomic groups in England View project [Internet]. 2004. Available from: https://www.researchgate.net/publication/255631127

117. Marshall A. Small area estimation using ESDS government surveys – An introductory guide. Economic and Social Data Service. 2010.

118. Rahman A. Estimating small area health-related characteristics of populations: A methodological review. Geospat Health. 2017;12(1).

119. Rahman A, Harding A. Some Health Related Issues in Australia and Methodologies for Estimating Small Area Health Related Characteristics; Working Paper no 15 [Internet]. Canberra; 2010 Dec [cited 2023 Feb 22]. Available from: www.natsem.canberra.edu.au

120. Office for National Statistics. (2020). 2011 Census: Aggregate Data. UK Data Service. SN: 7427, DOI: 10.5257/census/aggregate-2011-2 [Internet]. [cited 2023 Feb 22]. Available from: https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=7427

121. Macintyre S, Ellaway A, Cummins S. Place effects on health: how can we conceptualise, operationalise and measure them? Soc Sci Med. 2002 Jul 1;55(1):125–39.

122. Twigg L, Moon G, Jones K. Predicting small-area health-related behaviour: a comparison of smoking and drinking indicators. Soc Sci Med. 2000 Apr 1;50(7–8):1109–20.

123. Charlton J. Use of the Census Samples of Anonymised Records (SARs) and survey data in combination to obtain estimates at local authority level. Vol. 30, Environment and Planning A. 1998.

124. 2011 Census Microdata Individual Safeguarded Samples – User Guide [Internet]. 2016 [cited 2023 Feb 22]. Available from: file:///C:/Users/benam/Downloads/safeguardedmicrodatauserguideregandgrouplav10_tcm77-390387.pdf

125. Office for National Statistics. (2023). 2011 Census Microdata Individual Safeguarded Sample (Local Authority): England and Wales. UK Data Service. SN: 7682, DOI:

10.5255/UKDA-SN-7682-1 [Internet]. [cited 2023 Feb 22]. Available from: https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=7682

126. Moon G, Quarendon G, Barnard S, Twigg L, Blyth B. Fat nation: Deciphering the distinctive geographies of obesity in England. Soc Sci Med. 2007 Jul 1;65(1):20–31.

127. Whitworth A, Aitken G, Mclennan D, Taylor J. Evaluations and improvements in small area estimation methodologies, National Centre for Research Methods Methodological Review paper. 2013.

128. Edwards KL, Clarke GP, Thomas J, Forman D. Internal and External Validation of Spatial Microsimulation Models: Small Area Estimates of Adult Obesity. Appl Spat Anal Policy. 2011 Nov;4(4):281–300.

129. Edwards KL, Clarke GP. The design and validation of a spatial microsimulation model of obesogenic environments for children in Leeds, UK: SimObesity. Soc Sci Med. 2009 Oct;69(7):1127–34.

130. Census geographies [Internet]. Office for National Statistics. 2022 [cited 2023 Feb 23]. Available from: https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeographies

131. NatCen Social Research, NHS Digital. Health Survey for England 2018 - Methods Report [Internet]. 2019 Dec [cited 2023 Feb 23]. Available from: http://healthsurvey.hscic.gov.uk/media/81661/HSE18-Methods-rep.pdf

132. NatCen Social Research, NHS Digital. Health Survey for England 2018 - User Guide [Internet]. 2018 [cited 2023 Feb 23]. Available from: http://doc.ukdataservice.ac.uk/doc/8649/mrdoc/pdf/8649_hse_2018_user_guide.pdf

133. National Diet and Nutrition Survey: Results from Years 5 and 6 (combined) Rolling Programme (2014/2015-2015/2016) [Internet]. London; 2018 [cited 2023 Feb 23]. Available from: www.facebook.com/PublicHealthEngland

134. National Diet and Nutrition Survey: results from Years 5 and 6 (combined) - Appendices and Tables [Internet]. GOV.UK. 2018 [cited 2023 Feb 23]. Available from: https://www.gov.uk/government/statistics/ndns-results-from-years-5-and-6-combined

135. Active Lives Survey Y2 2016-17 Data - User Guide [Internet]. 2018 [cited 2023 Feb 23]. Available from: http://www.ipsos-mori.com/terms.http://www.ipsos-mori.com/terms.

136. Sport England. Active Lives Adult Survey 2016/17 - Understanding Behaviour [Internet]. 2019 Feb [cited 2023 Feb 23]. Available from: https://www.active-together.org/researchandevidence/active-lives-adult-survey-understanding-behaviours-2019

137. 2011 Rural-Urban Classification [Internet]. Office for National Statistics. 2016 [cited 2023 Feb 23]. Available from: https://www.ons.gov.uk/methodology/geography/geographicalproducts/ruralurbanclassifications/2011ruralurbanclassification

138. Rahman A, Harding A. Small area estimation and microsimulation modelling [Internet]. New York: Chapman & Hall; 2019 [cited 2023 Feb 22]. 495 p. Available from: https://www.routledge.com/Small-Area-Estimation-and-Microsimulation-Modeling/Rahman-Harding/p/book/9780367261269

139. Deb P, Norton EC. Modeling Health Care Expenditures and Use. Annu Rev Public Health [Internet]. 2018 Apr 2 [cited 2023 Feb 23];39:489–505. Available from: https://www.annualreviews.org/doi/abs/10.1146/annurev-publhealth-040617-013517

140. Mohammad N, McGivern L. Comparing Gamma and Log-Normal GLMs in R Using Simulation and Real Data Set [Internet]. ERIC Institute for Education Sciences. 2020 Aug [cited 2023 Feb 23]. Available from: https://files.eric.ed.gov/fulltext/ED607605.pdf

141. Mansournia MA, Nazemipour M, Naimi AI, Collins GS, Campbell MJ. Reflection on modern methods: Demystifying robust standard errors for epidemiologists. Int J Epidemiol. 2021 Feb 1;50(1):346–51.

142. Estimates of the very old, including centenarians [Internet]. Office for National Statistics. 2019 [cited 2023 Feb 23]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/methodologies/estimatesoftheveryoldincludingcentenariansukqmi

143. Scarborough P, Allender S, Rayner M, Goldacre M. Validation of model-based estimates (synthetic estimates) of the prevalence of risk factors for coronary heart disease for wards in England. Health Place. 2009;15(2):596–605.

144. Breusch TS, Pagan AR. A Simple Test for Heteroscedasticity and Random Coefficient Variation. Econometrica. 1979 Sep;47(5):1287.

145. Cook RD, Weisberg S. Diagnostics for heteroscedasticity in regression. Biometrika [Internet]. 1983 Apr 1 [cited 2023 Feb 23];70(1):1–10. Available from: https://academic.oup.com/biomet/article/70/1/1/240227

146. English Indices of Deprivation 2019, Department for Communities and Local Government, Crown Copyright 2019, https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019.

147. Chooi YC, Ding C, Magkos F. The epidemiology of obesity. Metabolism [Internet]. 2019 Mar 1 [cited 2023 Mar 7];92:6–10. Available from: doi:10.1016/j.metabol.2018.09.005

148. Pickering K, Scholes S, Bajekal M. Synthetic estimation of healthy lifestyles indicators: Stage 3 report. 2005.

149. Scholes S, Bajekal M. Synthetic estimation of healthy lifestyles indicators: User guide [Internet]. Available from: https://www.researchgate.net/publication/265658543

150. NatCen Social Research, NHS Digital. Health Survey for England 2021 - Official Statistics [Internet]. [cited 2023 Jun 20]. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england

151. National Child Measurement Programme, England, 2021/22 school year [Internet]. 2022 Nov [cited 2023 Feb 23]. Available from: https://digital.nhs.uk/data-and-

information/publications/statistical/national-child-measurement-programme/2021-22-school-year

152. National Child Measurement Programme 2014-15 [Internet]. NHS Digital. 2015 [cited 2023 Feb 23]. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/national-child-measurement-programme/2014-15-school-year

153. UK-WHO growth charts - 2-18 years [Internet]. World Health Organisation, Royal College of Paediatrics and Child Health. [cited 2023 Feb 23]. Available from: https://www.rcpch.ac.uk/resources/uk-who-growth-charts-2-18-years

154. 2011 Residential-based area classifications - Methodology and variables [Internet]. Office for National Statistics. 2018 [cited 2023 Feb 23]. Available from: https://www.ons.gov.uk/methodology/geography/geographicalproducts/areaclassifications/2011areaclassifications/methodologyandvariables

155. Fact Finder tool on private education [Internet]. Private Education Policy Forum. 2022 [cited 2023 Feb 16]. Available from: https://www.pepf.co.uk/fact-finder/facts-and-figures/

156. Briggs ADM, Scarborough P, Wolstenholme J. Estimating comparable English healthcare costs for multiple diseases and unrelated future costs for use in health and public health economic modelling. PLoS One [Internet]. 2018 [cited 2023 Feb 23];13(5). Available from: https://doi.org/10.1371/journal.pone.0197257

157. Briggs ADM, Wolstenholme J, Scarborough P. Estimating the cost-effectiveness of salt reformulation and increasing access to leisure centres in England, with PRIMEtime CE model validation using the AdViSHE tool. BMC Health Serv Res [Internet]. 2019 Jul 16 [cited 2023 Feb 23];19(1):1–14. Available from: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-019-4292-x

158. NHS England Our annual Report 2018/19 [Internet]. London; 2018 [cited 2023 Feb 23]. Available from: https://www.england.nhs.uk/wp-content/uploads/2019/07/Annual-Report-Full-201819.pdf

159. The Hospital Episode Statistics (HES) database [Internet]. NHS Digital. 2022 [cited 2023 Feb 23]. Available from: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics

160. Kearns B. The relevance of future unrelated health costs in economic evaluation in NICE appraisals: Report by the decision support unit [Internet]. University of Sheffield; 2020. Available from: www.nicedsu.org.uk

161. Perry-Duxbury M, Asaria M, Lomas J, Van Baal P. Health Policy Analysis Cured Today, Ill Tomorrow: A Method for Including Future Unrelated Medical Costs in Economic Evaluation in England and Wales [Internet]. Available from: www.elsevier.com/locate/jval

162. de Vries LM, van Baal PHM, Brouwer WBF. Future Costs in Cost-Effectiveness Analyses: Past, Present, Future. Pharmacoeconomics [Internet]. 2019 Feb 13 [cited 2023 Feb 23];37(2):119–30. Available from: https://link.springer.com/article/10.1007/s40273-018-0749-8

163. Dawson D, Street A. The York Series on the NHS White Paper - A Research Agenda: Discussion Paper 161 Reference Costs and the Pursuit of Efficiency in the 'New' NHS. 1998 Oct.

164. Jacobs R, Dawson D. Variation in unit costs of hospitals in the English National Health Service. J Health Serv Res Policy [Internet]. 2003 Oct [cited 2022 May 18];8(4):202–8. Available from: DOI: 10.1258/135581903322403263

165. Department of Health. The new NHS : modern, dependable. [Internet]. Stationery Office; 1997 [cited 2023 Mar 7]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/266003/newnhs.pdf

166. Department of Health. Reference Costs 2015-16 [Internet]. 2016 [cited 2023 Mar 7]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/577083/Reference_Costs_2015-16.pdf

167. Department of Health. Reference costs guidance 2015-16 [Internet]. 2016 [cited 2023 Mar 7]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/497127/Reference_costs_guidance_2015-16.pdf

168. Audit Yorkshire. Costing in the NHS: Revisited. https://www.audityorkshire.nhs.uk/news/costing-in-the-nhs/. 2019.

169. Healthcare Financial Management Association. The Newbies Guide for Reference Costing [Internet]. 2016 [cited 2022 Jan 25]. Available from: https://www.hfma.org.uk/docs/default-source/branches/West-Midlands-Branch/west-midlands/information/reference-costinge913f1cd1ab7692cb427ff0000b8cb05.pdf

170. NHS England, NHS Improvement. National Cost Collection 2019-20 Report [Internet]. 2021 Jun [cited 2022 Mar 28]. Available from: https://www.england.nhs.uk/wp-content/uploads/2021/06/National-Cost-Collection-2019-20-Report-FINAL.pdf

171. Department of Health. Reference Costs - Review of Uses by the Department of Health and National Partners. 2010.

172. Department of Health. Reference Costs - Review of Uses by NHS Bodies. 2010.

173. Department of Health. A simple guide to Payment by Results [Internet]. 2015 [cited 2022 Jan 27]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/213150/PbR-Simple-Guide-FINAL.pdf

174. Monitor. Approved Costing Guidance 2016 [Internet]. 2016 [cited 2023 Mar 7]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/498792/ACG_february_2016_final_with_cover.pdf

175. Healthcare Financial Management Association. Acute health clinical costing standards 2016/17 [Internet]. 2016 [cited 2023 Mar 7]. Available from:

https://www.hfma.org.uk/online-learning/bitesize-courses/detail/acute-health-clinical-costing-standards-2016-17

176. Ferguson B. NHS database of reference costs is severely flawed. BMJ [Internet]. 2001 [cited 2022 Jan 10];323(14 July):106. Available from: doi.org/10.1136/bmj.323.7304.106

177. NHS Improvement. Costing transformation programme - Patient-level costing: case for change [Internet]. 2016 [cited 2023 Mar 7]. Available from: https://www.hfma.org.uk/docs/default-source/our-networks/healthcare-costing-for-value-institute/approved-costing-guidance/ctp_plics_case_for_change_v1

178. Monitor. About Monitor [Internet]. 2014 Jul [cited 2023 Mar 7]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/499229/Operational_productivity_A.pdf

179. Lord Carter of Coles. Operational productivity and performance in English NHS acute hospitals: Unwarranted variations – An independent report for the Department of Health by Lord Carter of Coles. 2016 Feb.

180. Capita. Improving the quality of costing in the NHS: Payment by Results data assurance framework [Internet]. 2014 [cited 2023 Mar 7]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/319355/PbR_DAF_costing_briefing_June_2014.pdf

181. Deloitte. Reference Cost Data Quality A Final Report for Monitor [Internet]. 2014 [cited 2023 Mar 7]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/317571/Supporting_document_B_-_Deloitte_Reference_Cost_Data_Quality_for_publication8e14.pdf

182. Patient Level Information & Costing Systems (PLICS) & Reference Costs Best Practice Guide [Internet]. 2011 [cited 2022 Jan 10]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/215122/dh_132972.pdf

183. CQC. The state of health care and adult social care in England 2014/15 [Internet]. 2014 [cited 2023 Mar 7]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/467997/20151013_CQC_SoC_Summary_WEB.pdf

184. Monitor. Costing transformation programme. https://www.gov.uk/guidance/costing-transformation-programme. 2016.

185. NHS Digital. Patient Level Information and Costing System (PLICS) Data Collections: Costing Transformation Programme. https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/patient-level-information-and-costing-system-plics-data-collections. 2022.

186. NHS England. Patient Level Information and Costing System (PLICS) Data Collections. https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/patient-level-information-and-costing-system-plics-data-collections#plics-data-collections-already-undertaken. 2022.

187. Gray J. Costing Tour: An integrated approach to costing and patient care. Healthcare Finance [Internet]. 2016 Sep [cited 2023 Mar 7]; Available from: https://www.hfma.org.uk/news/healthcare-finance/feature/costing-tour

188. NHS England. Approved Costing Guidance 2022 – Standards: Integrated information requirements and costing processes, version 1 [Internet]. 2022 [cited 2023 Mar 7]. Available from: https://www.england.nhs.uk/publication/approved-costing-guidance-2022-integrated-prescribed-guidance-and-tools/

189. Technical Guide to Allocation Formulae and Pace of Change For 2016-17 to 2020-21: revenue allocations to Clinical Commissioning Groups and commissioning areas [Internet]. [cited 2021 Dec 6]. Available from: https://www.england.nhs.uk/publication/technical-guide-to-allocation-formulae-and-pace-of-change-for-2016-17-to-2020-21-revenue-allocations-to-clinical-commissioning-groups-and-commissioning-areas/

190. Consultation on 2021/22 National Tariff Payment System: A guide to the market forces factor [Internet]. 2021. Available from: www.england.nhs.uk/publication/2021-22-tariff-

191. NHS England. Equality and Health Inequalities Analysis revenue allocations to Clinical Commissioning Groups [Internet]. 2019 [cited 2023 Mar 7]. Available from: https://www.england.nhs.uk/wp-content/uploads/2019/01/04-pb-31-01-2019-ccg-allocations-ehi-analysis-form.pdf

192. Meakin S, Abbott S, Choi H, Funk S, Gruson H. NHS Trust Level COVID-19 Data Aggregated to a Range of Spatial Scales [Internet]. covid19.nhs.data. 2020 [cited 2023 Mar 9]. Available from: https://epiforecasts.io/covid19.nhs.data/index.html DOI: 10.5281/zenodo.4447465

193. CCG allocations 2019/20 to 2023/24 - all funding streams [Internet]. NHS England. 2019 [cited 2023 Feb 23]. Available from: https://www.england.nhs.uk/publication/ccg-allocations-2019-20-to-2023-24-all-funding-streams-spreadsheet/

194. Case studies and PLICS support [Internet]. Healthcare Management Financial Association. [cited 2023 Feb 23]. Available from: https://www.hfma.org.uk/our-networks/healthcare-costing-for-value-institute/case-studies-and-plics-support

195. Cobiac LJ, Scarborough P. Modelling future trajectories of obesity and body mass index in England. PLoS One [Internet]. 2021 Jun 1 [cited 2023 Feb 23];16(6):e0252072. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0252072

196. Government food strategy [Internet]. 2022 Jun [cited 2023 Feb 23]. Available from: https://www.gov.uk/government/publications/government-food-strategy/government-food-strategy

197. Nutrient Profiling Technical Guidance [Internet]. London; 2011 Jan [cited 2023 Feb 23]. Available from: http://www.dh.gov.uk/publications

198. Department of Health and Social Care. Impact Assessment (IA): Introducing a 2100-0530 watershed on TV and online restriction for paid advertising of food and drink

that are High in Fat, Salt and Sugar (HFSS) products [Internet]. 2021 [cited 2023 Feb 23]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/996232/impact-assessment-hfss-advertising.pdf

199. Introducing further advertising restrictions on TV and online for products high in fat, salt or sugar: consultation on secondary legislation [Internet]. HM Government. 2022 [cited 2023 Feb 23]. Available from: https://www.gov.uk/government/consultations/introducing-further-advertising-restrictions-on-tv-and-online-for-products-high-in-fat-salt-or-sugar-secondary-legislation/introducing-further-advertising-restrictions-on-tv-and-online-for-products-high-in-fat-salt-or-sugar-consultation-on-secondary-legislation

200. Consultation Outcome: Introducing further advertising restrictions on TV and online for products high in fat, salt and sugar: government response [Internet]. 2021 [cited 2023 Feb 23]. Available from: https://www.gov.uk/government/consultations/further-advertising-restrictions-for-products-high-in-fat-salt-and-sugar/outcome/introducing-further-advertising-restrictions-on-tv-and-online-for-products-high-in-fat-salt-and-sugar-government-response

201. Department for Digital CM and SD of H and SC. Equality assessment: introducing further advertising restrictions on TV and online for products high in fat, salt and sugar [Internet]. 2021 [cited 2023 Mar 7]. Available from: https://www.gov.uk/government/consultations/further-advertising-restrictions-for-products-high-in-fat-salt-and-sugar/outcome/equality-assessment-introducing-further-advertising-restrictions-on-tv-and-online-for-products-high-in-fat-salt-and-sugar

202. Government delays restrictions on multibuy deals and advertising on TV and online [Internet]. HM Government. 2022 [cited 2023 Feb 23]. Available from: https://www.gov.uk/government/news/government-delays-restrictions-on-multibuy-deals-and-advertising-on-tv-and-online

203. Mytton OT, Boyland E, Adams J, Collins B, O'Connell M, Russell SJ, et al. The potential health impact of restricting less-healthy food and beverage advertising on UK television between 05.30 and 21.00 hours: A modelling study. PLoS Med. 2020 Oct 13;17(10):e1003212.

204. HM Treasury. The Green Book: Central Government Guidance on Appraisal and Evaluation 2022 [Internet]. 2022 [cited 2023 Feb 23]. Available from: https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-governent/the-green-book-2020

205. Yau A, Berger N, Law C, Cornelsen L, Greener R, Adams J, et al. Changes in household food and drink purchases following restrictions on the advertisement of high fat, salt, and sugar products across the Transport for London network: A controlled interrupted time series analysis. PLoS Med [Internet]. 2022 Feb 1 [cited 2023 Jun 20];19(2):e1003915. Available from: https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1003915

206. Adams J, Ganiti E, White M. Socio-economic differences in outdoor food advertising in a city in Northern England. Public Health Nutr [Internet]. 2011 Jun [cited 2023 Feb 23];14(6):945–50. Available from: https://www.cambridge.org/core/journals/public-health-nutrition/article/socioeconomic-differences-in-outdoor-food-advertising-in-a-city-in-northern-england/192383F225D92A592658F950382D7E30

207. Adams J, Tyrrell R, Adamson AJ, White M. Socio-economic differences in exposure to television food advertisements in the UK: a cross-sectional study of advertisements broadcast in one television region. Public Health Nutr [Internet]. 2012 Mar [cited 2023 Feb 23];15(3):487–94. Available from: https://www.cambridge.org/core/journals/public-health-nutrition/article/socioeconomic-differences-in-exposure-to-television-food-advertisements-in-the-uk-a-crosssectional-study-of-advertisements-broadcast-in-one-television-region/3C93147DC9F16C32F04664DE03FC1D24

208. Russell SJ, Croker H, Viner RM. The effect of screen advertising on children's dietary intake: A systematic review and meta-analysis. Obes Rev [Internet]. 2019 Apr 1 [cited 2023 Feb 23];20(4):554–68. Available from: https://pubmed.ncbi.nlm.nih.gov/30576057/

209. Rose D, Pevalin DJ. The National Statistics Socio-economic Classification: Origins, Development and Use. Palgrave MacMillan; 2005. 1–120 p.

210. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual (PMG20) (last updated 18 January 2022) [Internet]. NICE; 2014 [cited 2023 Feb 23]. Available from: https://www.nice.org.uk/process/pmg20/chapter/introduction

211. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual (PMG36) Process and methods [Internet]. 2022 [cited 2023 Feb 23]. Available from: www.nice.org.uk/process/pmg36

212. Harris JL, Bargh JA, Brownell KD. Priming Effects of Television Food Advertising on Eating Behavior. Health Psychol [Internet]. 2009 Jul [cited 2023 Feb 23];28(4):404. Available from: /pmc/articles/PMC2743554/

213. Zimmerman FJ, Shimoga S v. The effects of food advertising and cognitive load on food choices. BMC Public Health [Internet]. 2014 [cited 2023 Feb 23];14(342). Available from: http://www.biomedcentral.com/1471-2458/14/342

214. Koordeman R, Anschutz DJ, van Baaren RB, Engels RCME. Exposure to soda commercials affects sugar-sweetened soda consumption in young women. An observational experimental study. Appetite. 2010 Jun 1;54(3):619–22.

215. Discounting: Task and finish group report; CHTE methods review [Internet]. 2020 Jul [cited 2023 Feb 23]. Available from: https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.nice.org.uk%2FMedia%2FDefault%2FAbout%2Fwhat-we-do%2Four-programmes%2Fnice-guidance%2Fchte-methods-consultation%2FDiscounting-task-and-finish-group-report.docx&wdOrigin=BROWSELINK

216. Gravelle H, Smith CHE D. Discounting for Health Effects in Cost Benefit and Cost Effectiveness Analysis; CHE Technical Paper Series 20 [Internet]. 2000 Oct [cited 2023 Feb 23]. Available from: https://www.york.ac.uk/che/pdf/tp20.pdf

217. Shah AK, Mullainathan S, Shafir E. Some consequences of having too little. Science (1979) [Internet]. 2012 Nov 2 [cited 2023 Feb 23];338(6107):682–5. Available from: https://www.science.org/doi/10.1126/science.1222426

218. Gebrekal T. Active Lives Survey 2020-2021 Year 6 (mid-year) - technical note [Internet]. 2021 Oct [cited 2023 Feb 23]. Available from: http://www.ipsos-mori.com/terms.

219. The English Indices of Deprivation 2019 [Internet]. 2019 Sep [cited 2023 Feb 23]. Available from: https://www.gov.uk/government/publications/english-indices-of-deprivation-2019-technical-report

220. Box G, Draper N. Empirical model-building and response surfaces [Internet]. John Wiley & Sons.; 1987 [cited 2023 Feb 23]. Available from: https://psycnet.apa.org/record/1987-97236-000

221. Preskey N. Irn-Bru's full sugar version returns for good . The Independent [Internet]. 2021 Mar 22 [cited 2023 Feb 23]; Available from: https://www.independent.co.uk/life-style/food-and-drink/irn-bru-full-sugar-1901-b1820667.html

222. Soley-Bori M, Ashworth M, Bisquera A, Dodhia H, Lynch R, Wang Y, et al. Impact of multimorbidity on healthcare costs and utilisation: a systematic review of the UK literature. British Journal of General Practice [Internet]. 2021 Jan 1 [cited 2023 Feb 23];71(702):e39–46. Available from: https://bjgp.org/content/71/702/e39

223. Nuttall FQ. Body Mass Index Obesity, BMI, and Health: A Critical Review. Nutrition Research. 2015 May;50(3).

224. National Institute for Health and Care Excellence. Methods for the development of NICE public health guidance (third edition) Process and methods Methods for the development of NICE public health guidance (third edition) (PMG4) [Internet]. 2012 Sep [cited 2023 Feb 23]. Available from: www.nice.org.uk/process/pmg4

225. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Statement: Updated Reporting Guidance for Health Economic Evaluations. Value in Health [Internet]. 2022 Jan 1 [cited 2023 Mar 6];25(1):3–9. Available from: http://www.valueinhealthjournal.com/article/S1098301521031466/fulltext

226. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good Practice Guidelines for Decision-Analytic Modelling in Health Technology Assessment A Review and Consolidation of Quality Assessment. Pharmacoeconomics. 2006;24(4):355–71.

227. Guide to the methods of technology appraisal 2013 Process and methods (PMG9:) [Internet]. 2013 Apr [cited 2023 Feb 23]. Available from: www.nice.org.uk/process/pmg9

228. Bradford Hill A. The Environment and Disease: Association or Causation? Proceedings of the Royal Society of Medicine: Occupational Medicine [Internet]. 1965 Jan 14 [cited 2023 Feb 23]; Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1898525/pdf/procrsmed00196-0010.pdf

229. Andersen P, Mizdrak A, Wilson N, Davies A, Bablani L, Blakely T. Disaggregating proportional multistate lifetables by population heterogeneity to estimate intervention impacts on inequalities. Popul Health Metr [Internet]. 2022 Dec 1 [cited 2023 Feb 23];20(1):1–17. Available from: https://pophealthmetrics.biomedcentral.com/articles/10.1186/s12963-022-00282-7

230. Foster D. Adult social care funding (England); Commons Library Research Briefing 7903 [Internet]. 2023 Jan [cited 2023 Feb 23]. Available from: https://researchbriefings.files.parliament.uk/documents/CBP-7903/CBP-7903.pdf

231. Blakely T, Cobiac LJ, Cleghorn CL, Pearson AL, van der Deen FS, Kvizhinadze G, et al. Health, Health Inequality, and Cost Impacts of Annual Increases in Tobacco Tax: Multistate Life Table Modeling in New Zealand. PLoS Med [Internet]. 2015 Jul 1 [cited 2023 Feb 23];12(7):e1001856. Available from: https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001856

232. Moreira PVL, Baraldi LG, Moubarac JC, Monteiro CA, Newton A, Capewell S, et al. Comparing Different Policy Scenarios to Reduce the Consumption of Ultra-Processed Foods in UK: Impact on Cardiovascular Disease Mortality Using a Modelling Approach. PLoS One [Internet]. 2015 Feb 13 [cited 2023 Feb 23];10(2):e0118353. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0118353

233. Bajekal M, Scholes S, Love H, Hawkins N, O'Flaherty M, Raine R, et al. Analysing Recent Socioeconomic Trends in Coronary Heart Disease Mortality in England, 2000–2007: A Population Modelling Study. PLoS Med [Internet]. 2012 [cited 2023 Feb 23];9(6):e1001237. Available from: https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001237

234. Scarborough P, Nnoaham KE, Clarke D, Capewell S, Rayner M. Modelling the impact of a healthy diet on cardiovascular disease and cancer mortality. J Epidemiol Community Health [Internet]. 2012 May 1 [cited 2023 Feb 23];66(5):420–6. Available from: https://jech.bmj.com/content/66/5/420

235. Unal B, Critchley JA, Capewell S. Modelling the decline in coronary heart disease deaths in England and Wales, 1981-2000: comparing contributions from primary prevention and secondary prevention. BMJ [Internet]. 2005 Sep 15 [cited 2023 Feb 23];331(7517):614. Available from: https://www.bmj.com/content/331/7517/614

236. Barton P, Andronis L, Briggs A, McPherson K, Capewell S. Effectiveness and cost effectiveness of cardiovascular disease prevention in whole populations: modelling study. BMJ [Internet]. 2011 Jul 28 [cited 2023 Feb 23];343(7819). Available from: https://www.bmj.com/content/343/bmj.d4044

237. Chamnan P, Simmons RK, Khaw KT, Wareham NJ, Griffin SJ. Estimating the population impact of screening strategies for identifying and treating people at high risk of

cardiovascular disease: modelling study. BMJ [Internet]. 2010 Apr 25 [cited 2023 Feb 23];340(7754):1016. Available from: https://www.bmj.com/content/340/bmj.c1693

238. Kypridemos C, Allen K, Hickey GL, Guzman-Castillo M, Bandosz P, Buchan I, et al. Cardiovascular screening to reduce the burden from cardiovascular disease: microsimulation study to quantify policy options. BMJ [Internet]. 2016 Jun 8 [cited 2023 Feb 23];353. Available from: https://www.bmj.com/content/353/bmj.i2793

239. Briggs ADM, Kehlbacher A, Tiffin R, Garnett T, Rayner M, Scarborough P. Assessing the impact on chronic disease of incorporating the societal cost of greenhouse gases into the price of food: An econometric and comparative risk assessment modelling study. BMJ Open [Internet]. 2013 Oct 1 [cited 2023 Feb 23];3(10):e003543. Available from: https://bmjopen.bmj.com/content/3/10/e003543

240. Aston LM, Smith JN, Powles JW. Impact of a reduced red and processed meat dietary pattern on disease risks and greenhouse gas emissions in the UK: a modelling study. BMJ Open [Internet]. 2012 Jan 1 [cited 2023 Feb 23];2(5):e001072. Available from: https://bmjopen.bmj.com/content/2/5/e001072

241. Gillies CL, Lambert PC, Abrams KR, Sutton AJ, Cooper NJ, Hsu RT, et al. Different strategies for screening and prevention of type 2 diabetes in adults: cost effectiveness analysis. BMJ [Internet]. 2008 May 22 [cited 2023 Feb 23];336(7654):1180–5. Available from: https://www.bmj.com/content/336/7654/1180

242. Marshall T, Rouse A. Resource implications and health benefits of primary prevention strategies for cardiovascular disease in people aged 30 to 74: mathematical modelling study. BMJ [Internet]. 2002 Jul 27 [cited 2023 Feb 23];325(7357):197. Available from: https://www.bmj.com/content/325/7357/197

243. Breeze PR, Thomas C, Squires H, Brennan A, Greaves C, Diggle PJ, et al. The impact of Type 2 diabetes prevention programmes based on risk-identification and lifestyle intervention intensity strategies: a cost-effectiveness analysis. Diabetic Medicine [Internet]. 2017 May 1 [cited 2023 Feb 23];34(5):632–40. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/dme.13314

244. Bós AM, Howard B V., Beresford SAA, Urban N, Tinker LF, Waters H, et al. Cost-Effectiveness Analysis of a Low-Fat Diet in the Prevention of Breast and Ovarian Cancer. J Am Diet Assoc [Internet]. 2011 Jan 1 [cited 2023 Feb 23];111(1):56–66. Available from: http://www.jandonline.org/article/S0002822310016470/fulltext

245. Breeze PR, Thomas C, Squires H, Brennan A, Greaves C, Diggle P, et al. SPHR Diabetes Prevention Model: Detailed Description of Model Background, Methods, Assumptions and Parameters. 2015 Jun 15 [cited 2023 Feb 23]; Available from: http://www.sheffield.ac.uk/scharr/sections/heds/discussion-papers/15-01-1.472126

246. Frew EJ, Bhatti M, Win K, Sitch A, Lyon A, Pallan M, et al. Cost-effectiveness of a community-based physical activity programme for adults (Be Active) in the UK: an economic analysis within a natural experiment. Br J Sports Med [Internet]. 2014 Feb 1 [cited 2023 Feb 23];48(3):207–12. Available from: https://bjsm.bmj.com/content/48/3/207

247. Gulliford MC, Bhattarai N, Charlton J, Rudisill C. Cost-effectiveness of a universal strategy of brief dietary intervention for primary prevention in primary care: Population-based cohort study and Markov model. Cost Effectiveness and Resource Allocation [Internet]. 2014 Feb 2 [cited 2023 Feb 23];12(1):1–9. Available from: https://resource-allocation.biomedcentral.com/articles/10.1186/1478-7547-12-4

248. Roberts S, Craig D, Adler A, McPherson K, Greenhalgh T. Economic evaluation of type 2 diabetes prevention programmes: Markov model of low- and high-intensity lifestyle programmes and metformin in participants with different categories of intermediate hyperglycaemia. BMC Med [Internet]. 2018 Jan 30 [cited 2023 Feb 23];16(1):1–12. Available from: https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-017-0984-4

249. Lewsey JD, Lawson KD, Ford I, Fox KAA, Ritchie LD, Tunstall-Pedoe H, et al. A cardiovascular disease policy model that predicts life expectancy taking into account socioeconomic deprivation. Heart [Internet]. 2015 [cited 2023 Feb 23];101:201–8. Available from: http://dx.doi.org/10.1136/heartjnl-2014-305637

250. Lawson KD, Lewsey JD, Ford I, Fox K, Ritchie LD, Tunstall-Pedoe H, et al. A cardiovascular disease policy model: part 2-preparing for economic evaluation and to assess health inequalities. Open Heart [Internet]. 2016 [cited 2023 Feb 23];3:140. Available from: http://openheart.bmj.com/

251. Clark SD. Modelling the Impacts of Demographic Ageing on the Demand for Health Care Services. University of Leeds; 2015.

252. Mytton OT. A Population Perspective on Physical Activity and Health. University of Cambridge; 2016.

253. Margaris P, University of Manchester. The Macroeconomic Effects of Obesity: An Application to the UK. A Job Market Paper. [Internet]. 2017 Sep [cited 2023 Mar 8]. Available from: https://www.academia.edu/34378078/The_Macroeconomic_Effects_of_Obesity_An_Application_to_the_UK

254. Stanciole AE. Essays on the Regulation of Health Care Provision and the Economics of Chronic Diseases. University of York; 2007.

255. Yao G. The economic evaluations of interventions for heart diseases. University of Birmingham; 2008.

256. Qi H. Computing Resources Sensitive Parallelization of Neural Networks for Large Scale Diabetes Data Modelling, Diagnosis and Prediction. Brunel, University of London; 2011.

257. O'Flaherty M. A better understanding of recent coronary heart disease mortality trends and determinants. University of Liverpool; 2011.

258. Douglas James Noble by. Predicting the epidemic: A study of diabetes risk profiling in a multi-ethnic inner city population. Queen Mary, University of London; 2012.

259. Lopez-Ortega M. Modelling health and frailty and its relationship to the use of health services and the supply of informal care. [London]: London School of Hygiene and Tropical Medicine; 2009.

260. Cost-effective commissioning of colorectal cancer care: An assessment of the cost-effectiveness of improving early diagnosis [Internet]. 2016 Oct [cited 2023 Feb 23]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/612370/cost-effectiveness-early-diagnosis-colorectal-cancer.pdf

261. Sassi F, Cecchini M, Lauer J, Chisholm D. Improving lifestyles, tackling obesity: the health and economic impact of prevention strategies. OECD Health Working Papers [Internet]. 2009 Nov 20 [cited 2023 Feb 23];(48):1-9,12-106. Available from: http://search.proquest.com/docview/189840328?accountid=11077

262. Department of Health and Social Care. Technical Consultation Document: Department of Health and Social Care (DHSC) Calorie Model [Internet]. 2018 [cited 2023 Mar 1]. Available from: https://www.gov.uk/government/publications/department-of-health-and-social-care-dhsc-calorie-model

263. Graf S, Cecchini M. Diet, physical activity and sedentary behaviours : Analysis of trends, inequalities and clustering in selected oecd countries. OECD Health Working Papers. OECD iLibrary [Internet]. 2017 Dec [cited 2023 Feb 23]. Available from: https://www.oecd-ilibrary.org/social-issues-migration-health/diet-physical-activity-and-sedentary-behaviours_54464f80-en

264. Singh GM, Danaei G, Farzadfar F, Stevens GA, Woodward M, Wormser D, et al. The Age-Specific Quantitative Effects of Metabolic Risk Factors on Cardiovascular Diseases and Diabetes: A Pooled Analysis. PLoS One [Internet]. 2013 Jul 30 [cited 2023 Feb 23];8(7):e65174. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0065174

265. Aune D, Sen A, Schlesinger S, Norat T, Janszky I, Romundstad P, et al. Body mass index, abdominal fatness, fat mass and the risk of atrial fibrillation: a systematic review and dose–response meta-analysis of prospective studies. Eur J Epidemiol [Internet]. 2017 Mar 1 [cited 2023 Feb 23];32(3):181–92. Available from: https://link.springer.com/article/10.1007/s10654-017-0232-4

266. World Cancer Research Fund. Body fatness and weight gain and the risk of cancer; WCRF Continuous Update Project [Internet]. 2018 [cited 2023 Feb 23]. Available from: https://www.wcrf.org/wp-content/uploads/2021/01/Body-fatness-and-weight-gain_0.pdf

267. Kyrgiou M, Kalliala I, Markozannes G, Gunter MJ, Paraskevaidis E, Gabra H, et al. Adiposity and cancer at major anatomical sites: umbrella review of the literature. The BMJ [Internet]. 2017 [cited 2023 Feb 23];356. Available from: /pmc/articles/PMC5421437/

268. Azizpour Y, Delpisheh A, Montazeri Z, Sayehmiri K, Darabi B. Effect of childhood BMI on asthma: a systematic review and meta-analysis of case-control studies. [cited 2023 Feb 23]; Available from: https://doi.org/10.1186/s12887-018-1093-z

269. Shiri R, Karppinen J, Leino-Arjas P, Solovieva S, Viikari-Juntura E. The Association Between Obesity and Low Back Pain: A Meta-Analysis. Am J Epidemiol [Internet]. 2010 Jan 15 [cited 2023 Feb 23];171(2):135–54. Available from: https://academic.oup.com/aje/article/171/2/135/130619

270. Silverwood V, Blagojevic-Bucknall M, Jinks C, Jordan JL, Protheroe J, Jordan KP. Current evidence on risk factors for knee osteoarthritis in older adults: A systematic review and meta-analysis. Osteoarthritis Cartilage. 2015 Apr 1;23(4):507–15.

271. Jiang L, Tian W, Wang Y, Rong J, Bao C, Liu Y, et al. Body mass index and susceptibility to knee osteoarthritis: A systematic review and meta-analysis. Joint Bone Spine. 2012 May 1;79(3):291–7.

272. Peters SAE, Huxley RR, Woodward M. Diabetes as a risk factor for stroke in women compared with men: A systematic review and meta-analysis of 64 cohorts, including 775 385 individuals and 12 539 strokes. The Lancet [Internet]. 2014 Jun 7 [cited 2023 Feb 23];383(9933):1973–80. Available from: http://www.thelancet.com/article/S0140673614600404/fulltext

273. The National Statistics Socio-economic classification (NS-SEC) [Internet]. Office for National Statistics. [cited 2023 Feb 23]. Available from: https://www.ons.gov.uk/methodology/classificationsandstandards/otherclassifications/thenationalstatisticssocioeconomicclassificationnssecrebasedonsoc2010

274. Castillo JJ, Reagan JL, Ingham RR, Furman M, Dalia S, Merhi B, et al. Obesity but not overweight increases the incidence and mortality of leukemia in adults: A meta-analysis of prospective cohort studies. Leuk Res. 2012 Jul 1;36(7):868–75.

275. Larsson SC, Wolk A. Body mass index and risk of non-Hodgkin's and Hodgkin's lymphoma: A meta-analysis of prospective studies. Eur J Cancer [Internet]. 2011 Nov 1 [cited 2023 Feb 23];47(16):2422–30. Available from: http://www.ejcancer.com/article/S0959804911004333/fulltext

276. Hidayat K, Li HJ, Shi BM. Anthropometric factors and non-Hodgkin's lymphoma risk: systematic review and meta-analysis of prospective studies. Crit Rev Oncol Hematol. 2018 Sep 1;129:113–23.

277. Amies-Cull B, Wolstenholme J, Cobiac L, Scarborough P. Estimating BMI distributions by age and sex for local authorities in England: a small area estimation study. BMJ Open [Internet]. 2022 Jun 1 [cited 2023 Feb 23];12(6):e060892. Available from: https://bmjopen.bmj.com/content/12/6/e060892

Appendix 1: Chapter appendices

Appendix 1a: Chapter 2 appendices

Appendix table 1a.1: Summary characteristics of models for full assessment. (PMSL – Proportional Multistate Lifetable Model, microsim – microsimulation, PA – physical activity, BMI – Body Mass Index, T2DM – Type-2 Diabetes)

Title	First author/ organisation	Year	Scenario modelling	Model structure	Weight, diet or PA risk	England subnational areas	Age range (years)	Time represented	All criteria met	Reference number
Comparing different policy scenarios to reduce the consumption of ultra-processed food in UK: impact on cardiovascular disease mortality using a modelling approach										
	Moriera	2015	Yes	microsim	processed foods	No	25+	Yes	No	(232)
Analysing different socioeconomic trends in coronary heart disease mortality in England, 2000-2007: a population modelling study										
	Bajekal	2012	Yes	cohort	BMI	No	25+	Yes	No	(233)
Modelling the impact of a healthy diet on cardiovascular disease and cancer mortality										

	Scarborough	2012	Yes	static sample	various micro/macronutrients	No	all	No	No	(234)
Modelling the decline in coronary heart disease deaths in England and Wales, 1981-2000: comparing contributions from primary prevention and secondary prevention										
	Unal	2005	Yes	cohort	-	No	25-84	Yes	No	(235)
Effectiveness and cost effectiveness of cardiovascular disease preventions in whole populations: modelling study										
	Barton	2011	Yes	cohort	-	No	40-90	Yes	No	(236)
Estimating the population impact of screening strategies for identifying and treating people at high risk of cardiovascular disease: modelling study										
	Chamnan	2010	Yes	static sample	-	No	40-74	No	No	(237)
Cardiovascular screening to reduce the burden from cardiovascular disease: microsimulation study to quantify policy options										
	Kypridemos	2016	Yes	microsim	BMI	No	40-74	Yes	No	(238)
Assessing the impact on chronic disease of incorporating the societal cost of greenhouse gases into the price of food: an econometric and comparative risk assessment modelling study										
	Briggs	2013	Yes	econometric and comparative risk assessment	food product changes	No	18+	No	No	(239)

Impact of a reduced red and processed meat dietary pattern on disease risks and greenhouse gas emissions in the UK: a modelling study										
	Aston	2012	Yes	static sample	red and processed meat	No	19-64	No	No	(240)
Different strategies for screening and prevention of type-2 diabetes in adults: cost effectiveness analysis										
	Gillies	2008	Yes	decision tree/ Markov model	-	No	45+	Yes	No	(241)
Resource implications and health benefits of primary prevention strategies for cardiovascular disease in people aged 30-74: mathematical modelling study										
	Marshall	2002	Yes	static sample	-	No	30-74	No	No	(242)
The impact of type-2 diabetes prevention programmes based on risk-identification and lifestyle intervention strategies cost-effectiveness analysis										
	Breeze	2017	Yes	microsim	BMI	No	16+	Yes	No	(243)
Cost-effectiveness of a low-fat diet in the prevention of breast and ovarian cancer										
	Bos	2019	Yes	cohort	Fat consumption	No	50+	Yes	No	(244)
SPHR Diabetes Prevention Model: detailed description of model background, methods, assumptions and parameters										
	Breeze	2015	Yes	microsim	BMI	No	16+	Yes	No	(245)

Cardiovascular Disease Prevention Return on Investment Tool Final Report										
	PHE	2018	Yes	microsim	BMI	Yes	all	Yes	Yes	(74)
Cost effectiveness of a community-based physical activity programme for adults (Be Active) in the UK: an economic analysis within a natural experiment										
	Frew	2012	Yes	Markov	PA	No	16-70	Yes	No	(246)
Cost-effectiveness of a universal strategy of brief intervention for primary prevention in primary care: population-based cohort study and Markov model										
	Gulliford	2014	Yes	Markov	fruit and vegetables	No	30-100	Yes	No	(247)
Economic evaluation of type-2 diabetes prevention programmes: Markov model of low- and high-intensity lifestyle programmes and metformin in patients with different categories of intermediate hyperglycaemia										
	Roberts	2018	Yes	Markov	Diet/ PA (T2DM)	No	40-100	Yes	No	(248)
A cardiovascular disease policy model that predicts life expectancy taking into account socioeconomic deprivation										
	Lewsey	2014	Yes	Markov/ life table	-	No	25+	Yes	No	(249)
A cardiovascular disease policy model: part 2 - preparing for economic evaluation and to assess health inequalities										
	Lawson	2016	Yes	Markov/ life table	-	No	25+	Yes	No	(250)
Estimating and comparing the cost-effectiveness of primary prevention policies affecting diet and physical activity in England										

	Briggs	2017	Yes	PMSL	BMI/ PA/ various dietary	No	17+	Yes	No	(72)
Modelling the effectiveness and equity of primary prevention policies in England										
	Kypridemos	2016	Yes	microsim	BMI, fruit and vegetables, salt, PA	No	0+	Yes	No	(71)
Modelling the impacts of demographic aging on demand for health care services										
	Clark	2015	Yes	microsim	-	Yes	50+	No	No	(251)
A population perspective on physical activity and health										
	Mytton	2016	Yes	microsim (PA) and PMSL (health)	PA	No	16+	No	No	(252)
The macroeconomic effects of obesity										
	Margaris	2017	No	-	-	-	-	-	No	(253)
Essays on the regulation of healthcare provision and economics of chronic disease										
	Stanciole	2007	No	-	-	-	-	-	No	(254)
Economic evaluations of interventions for heart diseases										
	Yao	2008	Yes	Markov	-	No	N/A	Yes	No	(255)

Computing resources sensitive parallelization of neural networks for large scale diabetes data modelling, diagnosis and prediction										
	Qi	2011	Yes	neural networks	-	No	N/A	Yes	No	(256)
A better understanding of recent coronary heart disease mortality trends and determinants										
	O'Flaherty	2011	Yes	microsim	BMI, fruit and vegetables, salt, PA	No	35+	Yes	No	(257)
Predicting the epidemic: A study of diabetes risk profiling in a multi-ethnic inner-city population										
	Noble	2012	No	-	-	-	-	-	No	(258)
Modelling health and frailty and its relationship to the use of health services and the supply of informal care										
	Lopez-Ortega	2009	No	-	-	-	-	-	No	(259)
Cost-effective commissioning of colorectal cancer care										
	PHE	2016	No	-	-	-	-	-	No	(260)
User guide: weight management economic assessment tool version 2										
	PHE	2016	Yes	Markov	BMI	Yes	18+	Yes	Yes	(82)
Improving lifestyles, tackling obesity: the health and economic impact of prevention strategies										

	OECD	2009	Yes	microsim	Fruit and vegetables, fat intake, PA, BMI	No	0-100	Yes	No	(261)
Estimating Return on Investment for interventions and strategies to increase physical activity										
	NICE	2014	Yes	Markov	PA	Yes	5 to 64	Yes	Yes	(81)
Technical Consultation Document, Department of Health and Social Care (DHSC) Calorie Model										
	DHSC	2018	Yes	Markov	BMI, energy change	No	4 to 79	Yes	No	(262)
Diet, physical activity and sedentary behaviours: Analysis of trends, inequalities and clustering in selected countries										
	OECD	2017	Yes	-	-	-	-	-	No	(263)

Appendix 1a.2: Access and presentation of results of included models

SPHR DPP RoI Tool

The tool is accessed through an online portal with input fields for location (whole of England or for a specific local authority or CCG), the target number of patients referred to the intervention each year, the percentage uptake of those referred, the planned intervention cost per patient attending (with a default provided excluding costs of recruitment), intervention effects (from meta-analysis of comparable pragmatic lifestyle diabetes interventions on weight loss, HbA1c and other physiological markers, or 25% higher or lower than the default(73)) and a time to return to baseline weight (default is 5 years). Users provide an email address and submit the input details, then the analysis is run on the University of Sheffield's servers and results are emailed to the user. A test scenario took 55 minutes to run (November 2019).

Results are presented in terms of numbers of diabetes cases, CVD events and diabetic microvascular cases averted over time horizons of 1-5 years, 10, 15 and 20 years. Cost savings can be broken down by area of the health and care service they are attributable to (primary care, secondary care, prescriptions, lab test and social care savings) or by the disease state averted (diabetes, diagnostics, statins and antihypertensives, retinopathy, neuropathy, cancer, osteoarthritis, depression), over 1-20 years. Cost per QALY and NMB are presented and local area results are compared with equivalent results for the intervention implemented across the whole of England.

PHE CVD Prevention RoI Tool

The CVD tool is accessed via an online portal. The interface has two options. If users select "What happens when I improve detection or management of key CVD risk factors?" they can set their local proportion of each CVD risk factor's cases that are detected and the proportion of each risk factor's detected cases that are well-managed. Then they set the target proportions and submit the scenario. If they select "What happens when I improve the usage of the key interventions for people at risk of CVD?" then they can change the proportion of eligible patients that receive each of the 16 included interventions for managing risk. When the scenario is submitted, the analysis is performed on University of Sheffield servers and a link to results is emailed back.

Results are presented in terms of changes to numbers of clinical events, mortality, Life Years, costs and health benefit (QALYs and NMB) with options of subgroup analyses by high-risk condition group, socioeconomic status (by IMD quintile), ethnicity and age group (<40, 40-74, >=74). Time horizons of 1-20 years are available. Discount rates for costs and QALYs can be varied. The outputs' raw data can be downloaded though it is poorly labelled and there is no uncertainty provided.

NICE Physical Activity RoI Tool

The tool has been retired from use, but was previously available as an Excel spreadsheet providing instantaneous results. Local Authority or CCG can be selected, then percent of adults and children recruited to interventions. Pre-programmed interventions for adults are one-on-one advice on PA and on transport, pedometer use, mass-media campaigns,

community-based walking programmes, cycling or walking programmes, urban infrastructure, and workplace information, workplace walking programmes and workplace multi-component interventions. Child interventions are group-based health information, and group-based walking programmes. Results are output in terms of avoidable disease burden expressed as QALYs, Net Present Value (taking a quasi-societal perspective incorporating local economy, health social care and transport costs), ICER and Benefit-Cost ratio. Results can display with comparisons of these outcome measures between baseline and alternative scenario packages of interventions.

PHE Weight Management Economic Assessment Tool

Despite being retired from use, the tool and some of its documentation remain accessible from the UK National Archives. It exists in the form of an Excel spreadsheet providing instantaneous results. This model does not specifically model an area or subgroup, but instead relies on users to input their population/ participants' gender balance, mean age and mean starting BMI. Intervention variables are the time to full recruitment, dropout rate, mean reduction in BMI, time taken to achieve this weight loss, the time that maximum weight loss is maintained, local authority costs, NHS costs and discount rates for costs and health benefits.

Results are presented as projected BMI, survival rates, QALYs and deaths, with and without the intervention. Health impact is presented by disease (diabetes, CVD, stroke, colorectal cancer and breast cancer) in terms of cases, mortality rates, case-years and costs. Other outputs are costs (by NHS and social care), QALYs and economic benefit (employment), with or without discounting.

Appendix 1b: Chapter 3 appendices

Appendix table 1b: Sources for Relative Risks used in the PRIMEtime_local model

Relative Risks	Reference
Disease	
BMI-Ischaemic heart disease	Singh GM, Danaei G, Farzadfar F, et al. The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. PLoS ONE 2013; 8: e65174(264)
BMI-Stroke	Singh GM, Danaei G, Farzadfar F, et al. The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. PLoS ONE 2013; 8: e65174 (264)
BMI-Hypertensive heart disease	Singh GM, Danaei G, Farzadfar F, et al. The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. PLoS ONE 2013; 8: e65174 (264)
BMI-Type-2 Diabetes	Singh GM, Danaei G, Farzadfar F, et al. The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. PLoS ONE 2013; 8: e65174 (264)
BMI-Atrial fibrillation and flutter	Aune D, Sen A, Schlesinger S, et al. Body mass index, abdominal fatness, fat mass and the risk of atrial fibrillation: a systematic review and dose–response meta-analysis of prospective studies. Eur J Epidemiol 2017;32(3):181-92. (265)
BMI-Breast cancer	World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project Expert Report 2018. Body fatness and weight gain and the risk of cancer. Available at dietandcancerreport.org (266)
BMI-Oesophageal cancer	Kyrgiou et al, Adiposity and cancer at major anatomical sites: umbrella review of the literature, BMJ 2017;356:j477 http://dx.doi.org/10.1136/bmj.j477 (267)
BMI-Colorectal cancer	World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project Expert Report 2018. Body fatness and weight gain and the risk of cancer. Available at dietandcancerreport.org (266)
BMI-Asthma	Azizpour et al, Effect of childhood BMI on asthma: a systematic review and meta-analysis of case-control studies,

	BMC Pediatr, 2018 Apr 26;18(1):143. doi: 10.1186/s12887-018-1093-z. (268)
BMI-Low back pain	Shiri et al, The Association Between Obesity and Low Back Pain: A Meta-Analysis, Am J Epidemiol 2010;171:135–154 DOI: 10.1093/aje/kwp356 (269)
BMI-Osteoarthritis knee	Silverwood et al, Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis, Osteoarthritis and Cartilage 23 (2015) 507-515, http://dx.doi.org/10.1016/j.joca.2014.11.019 (270)
BMI-Osteoarthritis hip	Jiang et al, Body mass index and susceptibility to knee osteoarthritis: A systematic review and meta-analysis, Joint Bone Spine 79 (2012) 291–297, doi:10.1016/j.jbspin.2011.05.015 (271)
T2DM-Ischemic Heart Disease	Peters et al, Diabetes as risk factor for incident coronary heart disease in women compared with men: a systematic review and meta-analysis of 64 cohorts including 858,507 individuals and 28,203 coronary events Diabetologia 2014 57:1542-1551
T2DM-Stroke	Peters et al 2014 Lancet; Diabetes as a risk factor for stroke in women compared with men: a systematic review and meta-analysis of 64 cohorts, including 775,385 individuals and 12,539 strokes 383: 1973-80 (272)
Utilities	
Diseases	Sullivan PW, Slejko JF, Sculpher MJ, Ghushchyan V. Catalogue of EQ-5D scores for the United Kingdom. Med Decis Making. 2011;31(6):800-4 (98)
Age, sex	Sullivan PW, Slejko JF, Sculpher MJ, Ghushchyan V. Catalogue of EQ-5D scores for the United Kingdom. Med Decis Making. 2011;31(6):800-4 (98)
Populations	
0-89 years	Population estimates for the UK, England and Wales, Scotland and Northern Ireland [Internet]. Office for National Statistics. 2019 [cited 2023 Feb 23]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2018
90-100 years	Synthesised – Chapter 3 and Appendix 2b
Anthropometrics	
Adult BMI	Synthesised – Chapter 5
Adult heights	NatCen Social Research, University College London, Department of Epidemiology and Public Health. (2023). Health

	Survey for England, 2019. UK Data Service. SN: 8860, DOI: 10.5255/UKDA-SN-8860-1 [Internet]. [cited 2023 Feb 22]. Available from: https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=8860 (112)
Children's BMI	Synthesised – Chapter 6
Children's heights	National Child Measurement Programme, England 2019/20 School Year - NDRS [Internet]. [cited 2023 Feb 21]. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/national-child-measurement-programme/2019-20-school-year (7) National Child Measurement Programme, England, 2021/22 school year [Internet]. 2022 Nov [cited 2023 Feb 23]. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/national-child-measurement-programme/2021-22-school-year (151)
Adult energy change-body weight change	Hall KD, Sacks G, Chandramohan D, Chow CC, Wang YC, Gortmaker SL, et al, Quantification of the effect of energy imbalance on bodyweight. The Lancet. 2011 Aug 27 [cited 2023 Feb 22];378(9793):826–37. Available from: http://www.thelancet.com/article/S014067361160812X/fulltext (92)
Child energy change-body weight change	Hall KD, Butte NF, Swinburn BA, Chow CC. Dynamics of childhood growth and obesity: development and validation of a quantitative mathematical model. Lancet Diabetes Endocrinol [Internet]. 2013 Oct 1 [cited 2023 Feb 22];1(2):97–105. Available from: http://www.thelancet.com/article/S2213858713700512/fulltext (93)
Epidemiology	
Incidence, prevalence and case fatality – Ischaemic Heart Disease, Stroke, Type-2 Diabetes, Hypertensive Heart Disease, Atrial Fibrillation and Flutter, and breast, colorectal and oesophageal cancers	Synthesised –Chapter 4

Incidence, prevalence and mortality – asthma	Global Health Data Exchange (GHDx) [Internet]. Institute for Health Metrics and Evaluation; [cited 2023 Feb 22]. Available from: https://ghdx.healthdata.org/ (97)
Incidence, prevalence – osteoarthritis of the knee, osteoarthritis of the hip, low back pain.	Global Health Data Exchange (GHDx) [Internet]. Institute for Health Metrics and Evaluation; [cited 2023 Feb 22]. Available from: https://ghdx.healthdata.org/ (97)
All-cause mortality	Global Health Data Exchange (GHDx) [Internet]. Institute for Health Metrics and Evaluation; [cited 2023 Feb 22]. Available from: https://ghdx.healthdata.org/ (97)
Costs	
Annual costs of disease	Synthesised –Chapter 7

Appendix 1c: Chapter 4 appendices

Appendix 1c.1: Diseases using previously-synthesised data

Hypertensive heart disease

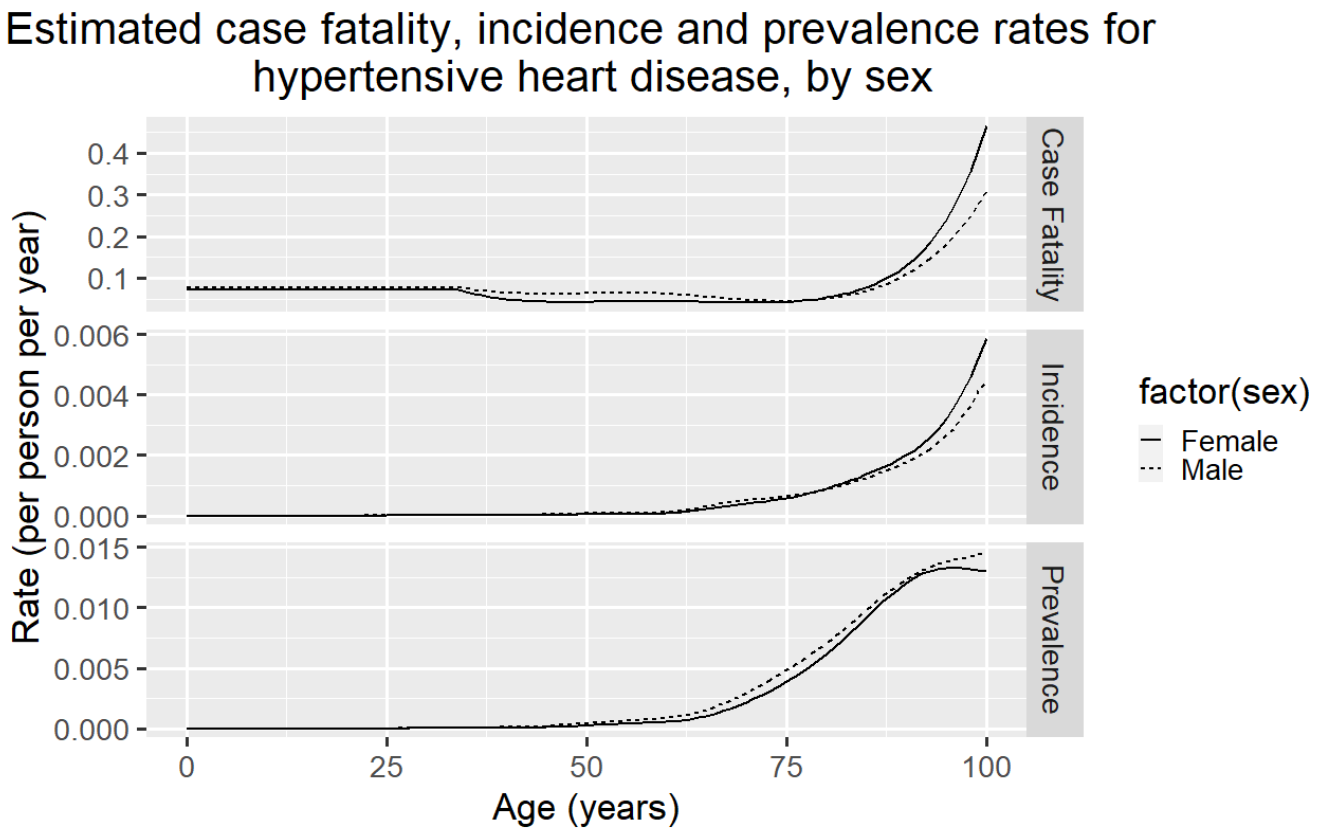


Figure 1c.1.1: England-level incidence prevalence and case fatality rates for females and males, taken from previous work on PRIMEtime(47)

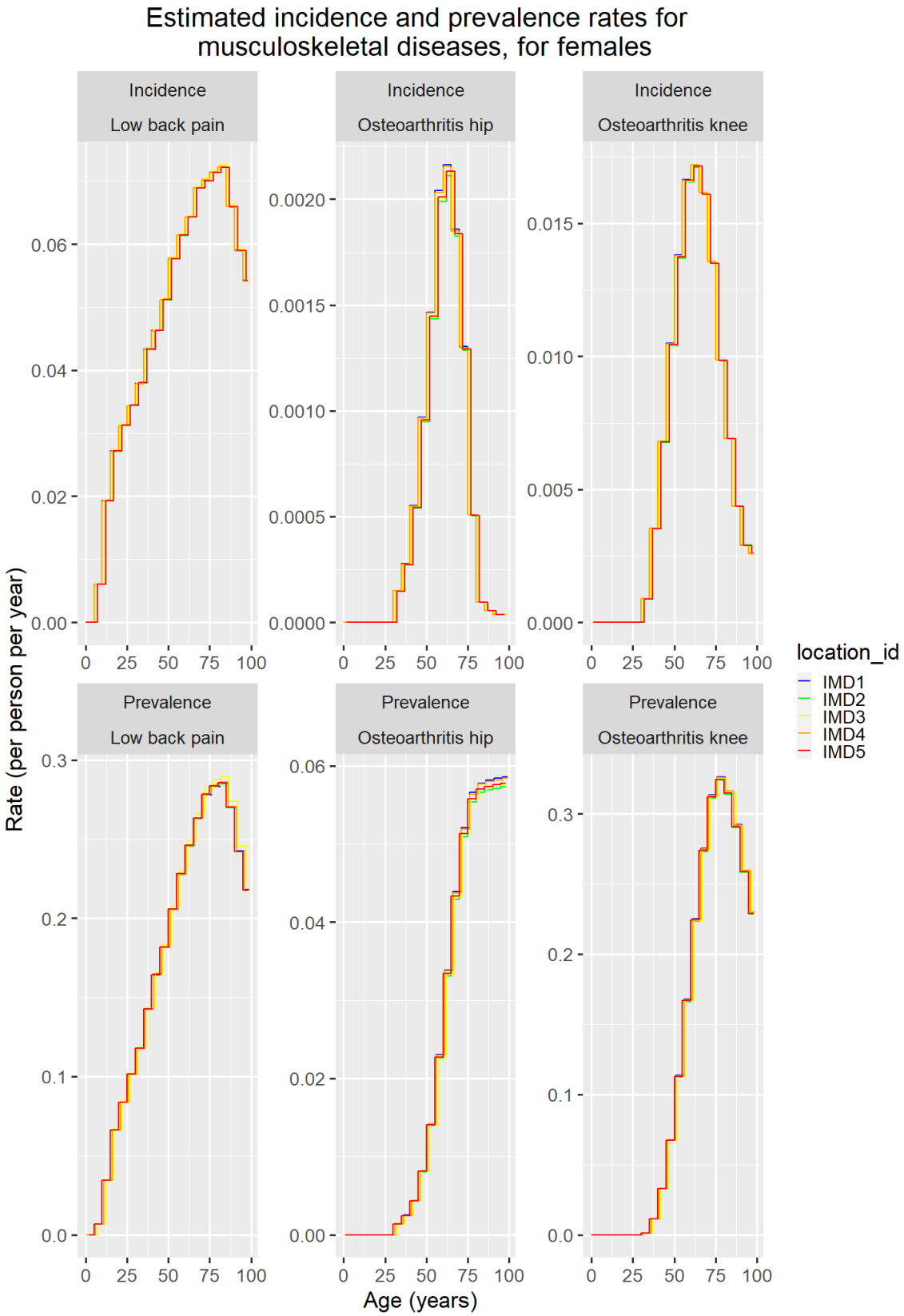


Figure 1c.1.2: Female incidence and prevalence rates of musculoskeletal diseases from GBD (published in 5-year age bands, which lead to the stepped appearance of the curves)(97)

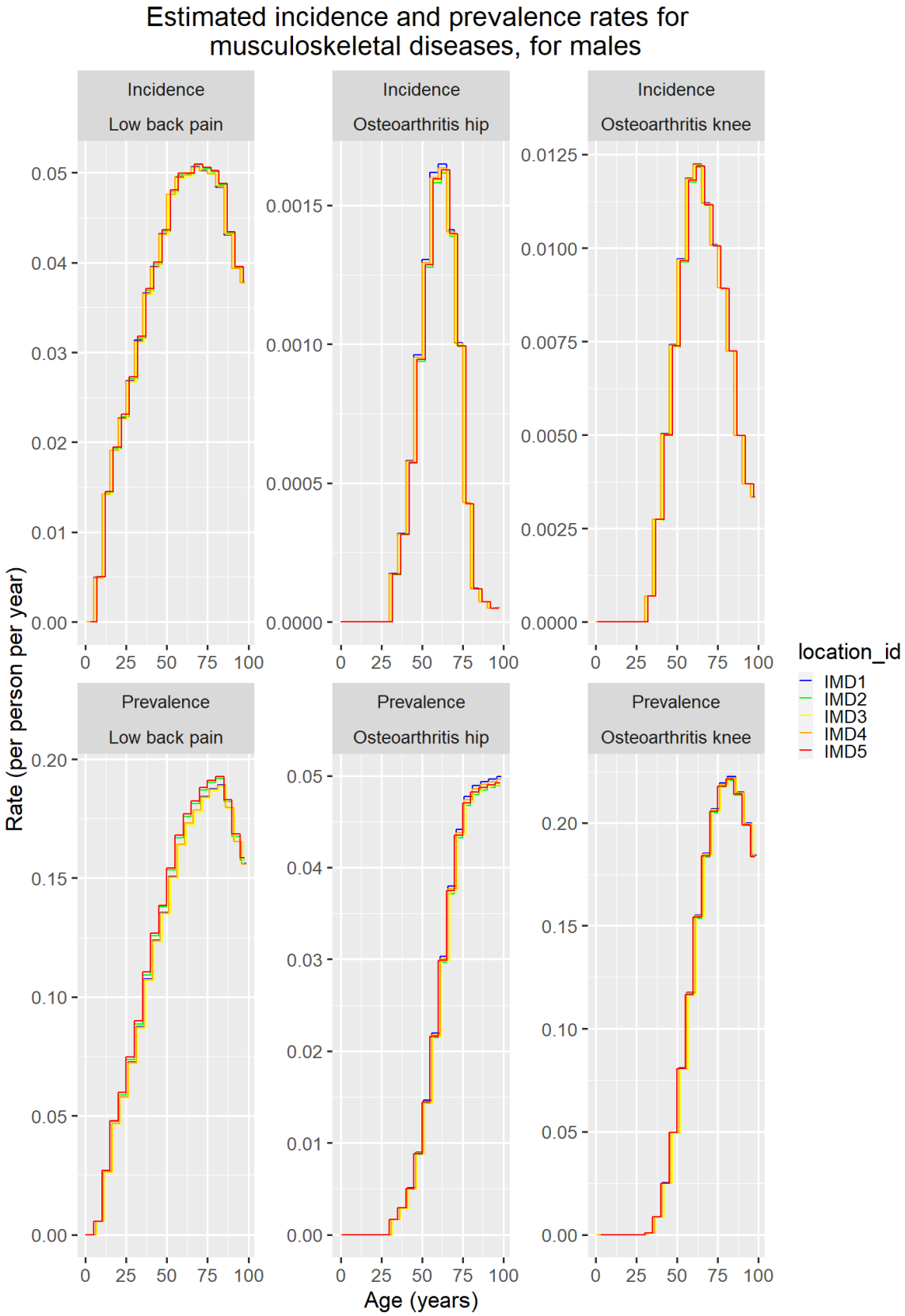


Figure 1c.1.3: Male incidence and prevalence rates of musculoskeletal diseases from the GBD (published in 5-year age bands, which lead to the stepped appearance of the curves)(97)

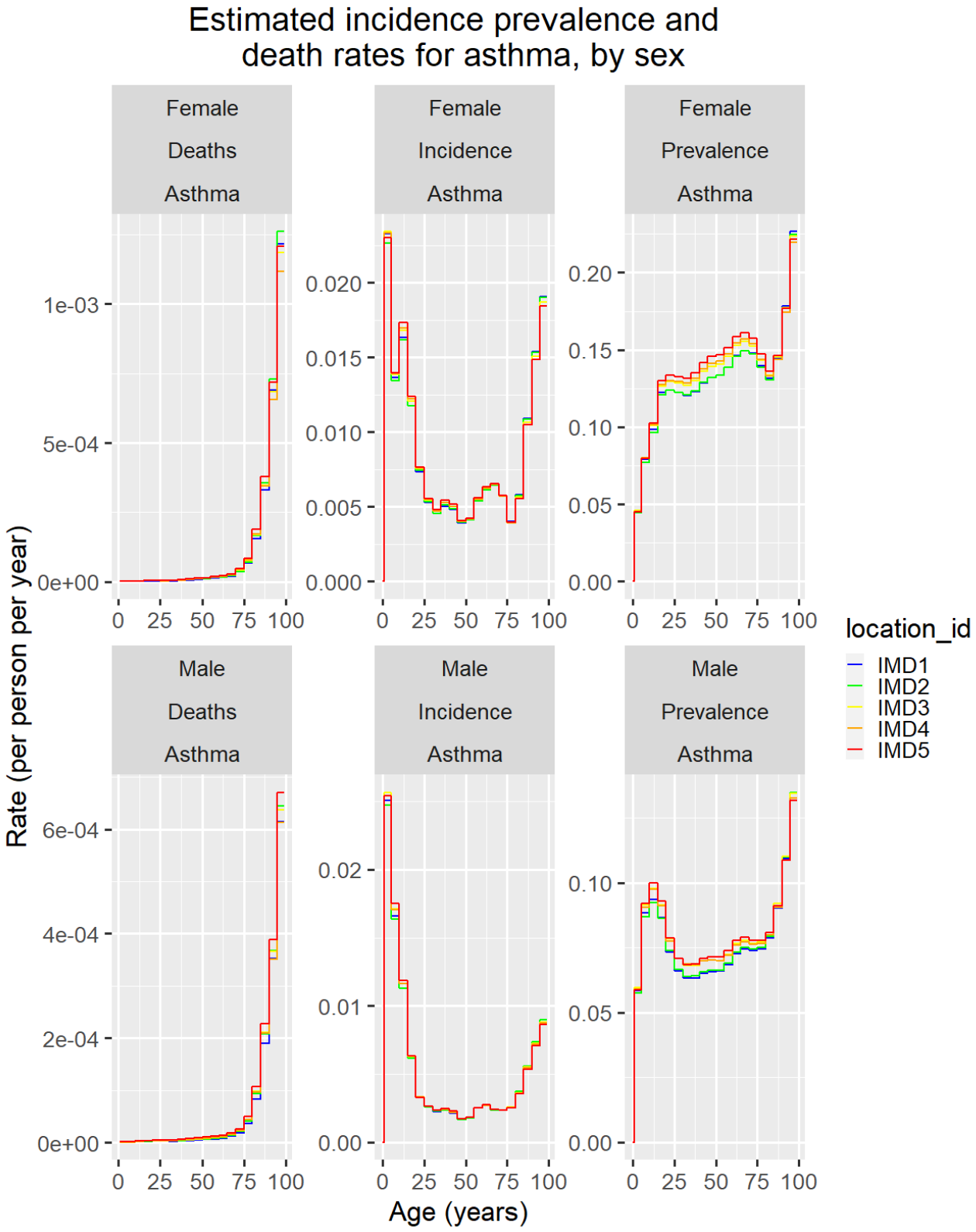


Figure 1c.1.4: Incidence, prevalence and mortality rates of asthma, for females and males, from GBD (published in 5-year age bands, which lead to the stepped appearance of the curves)(97)

Appendix 1c.2: Population-weighted average epidemiology (local-specific populations)

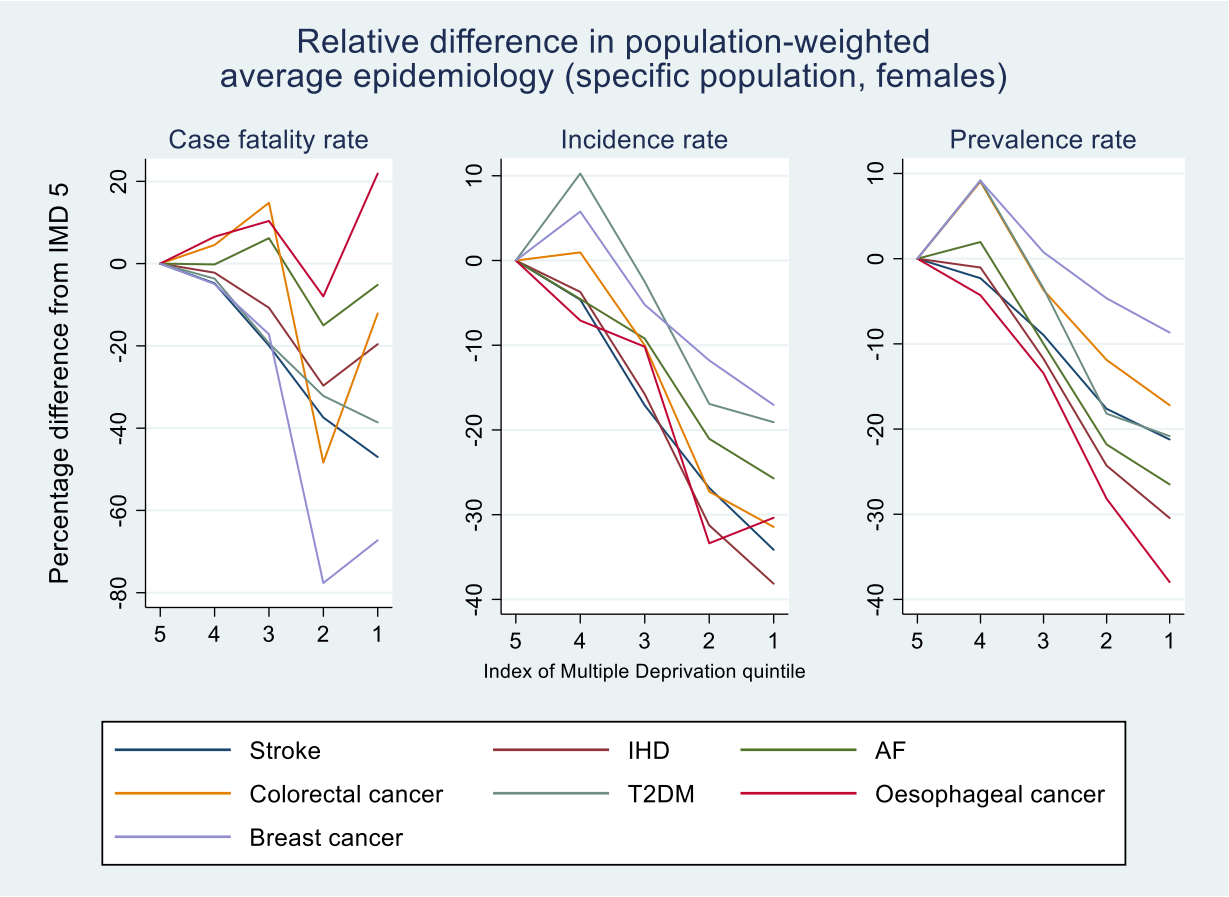


Figure 1c.2.1: Local-specific population-weighted average incidence, prevalence and case fatality rates of specified diseases for females, by Index of Multiple Deprivation quintile.

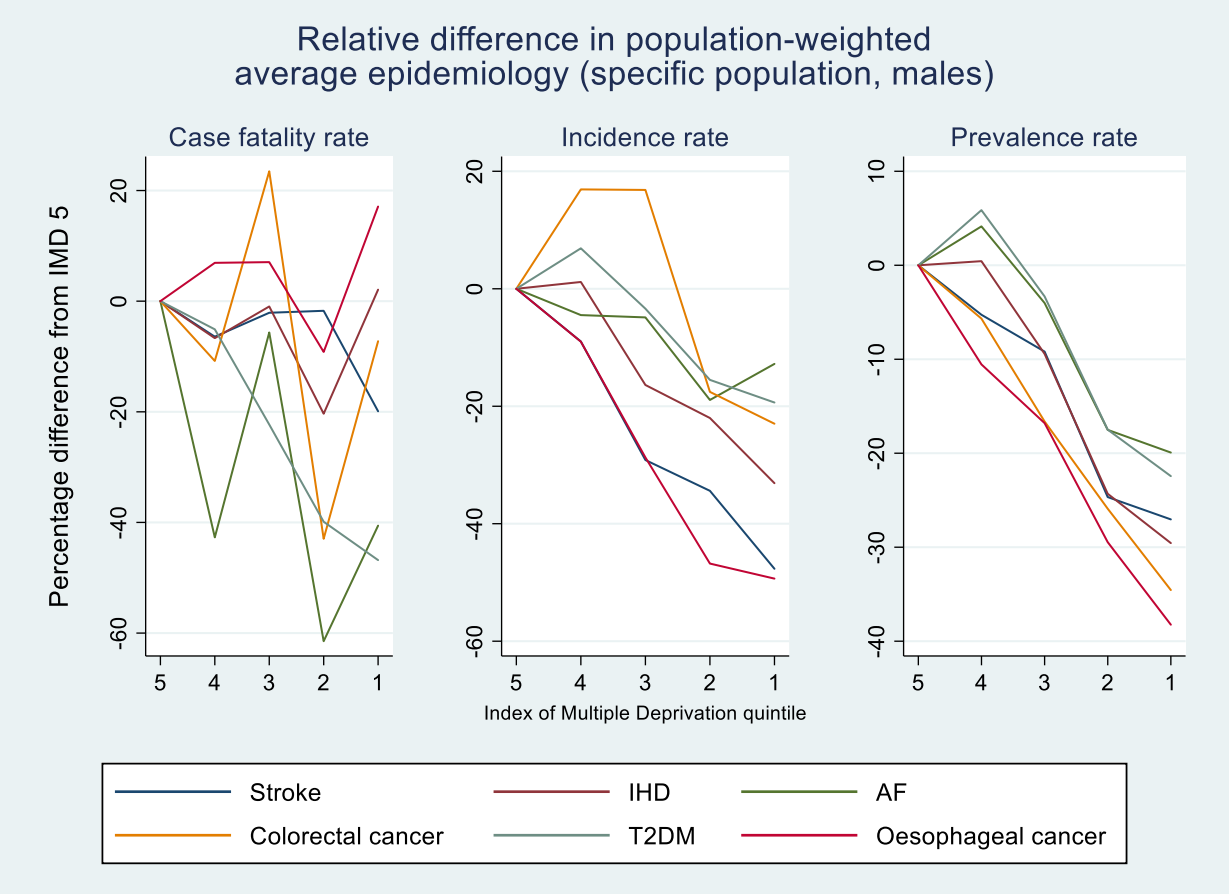


Figure A2.2: Local-specific population-weighted average incidence, prevalence and case fatality rates of specified diseases for males, by Index of Multiple Deprivation quintile.

Appendix 1c.3: Disbayes diseases by IMD quintile

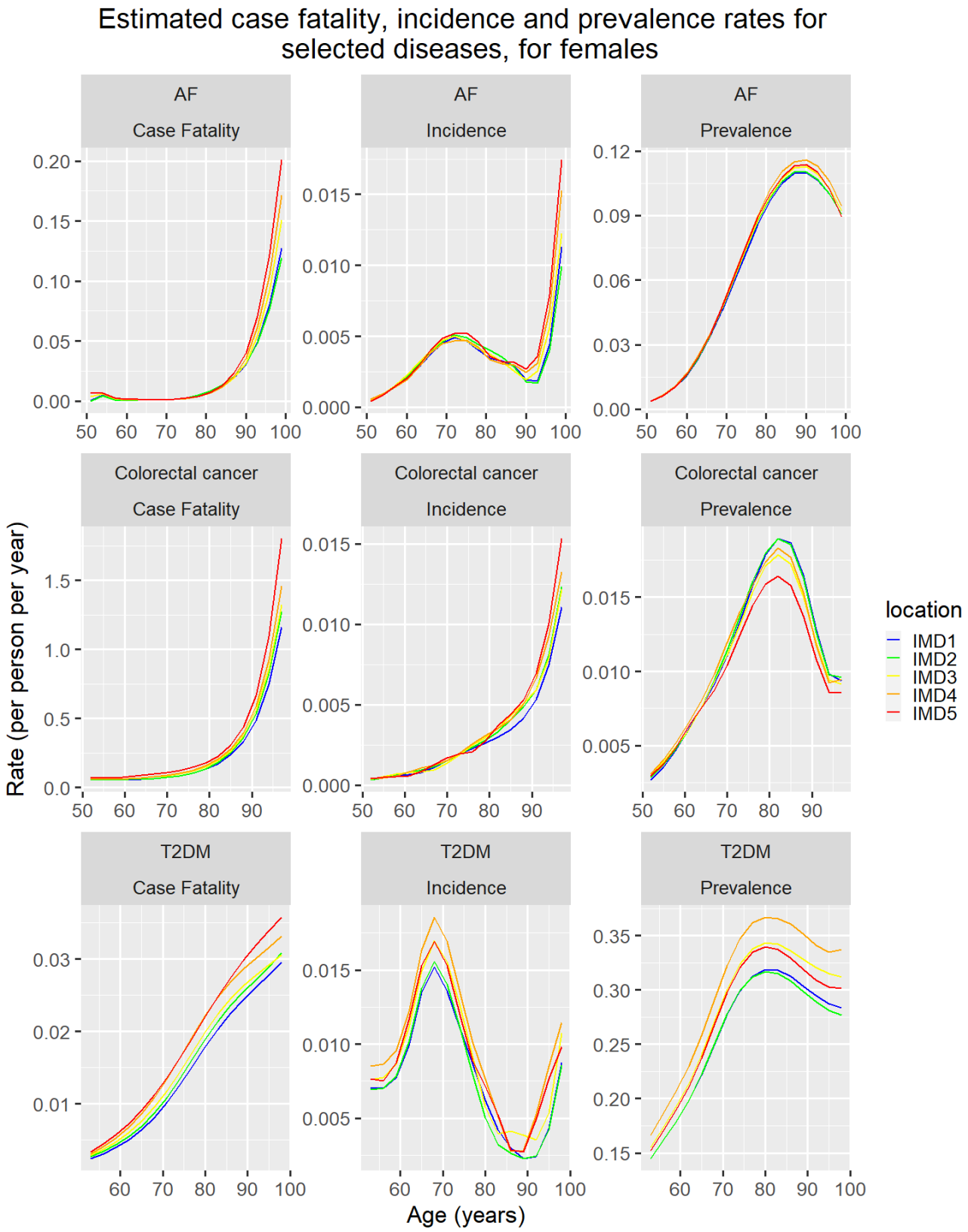


Figure 1c.3.1: Incidence, prevalence and case fatality rates of specified diseases, for females, by Index of Multiple Deprivation quintile (IMD 1 (least deprived) to 5 (most deprived)).

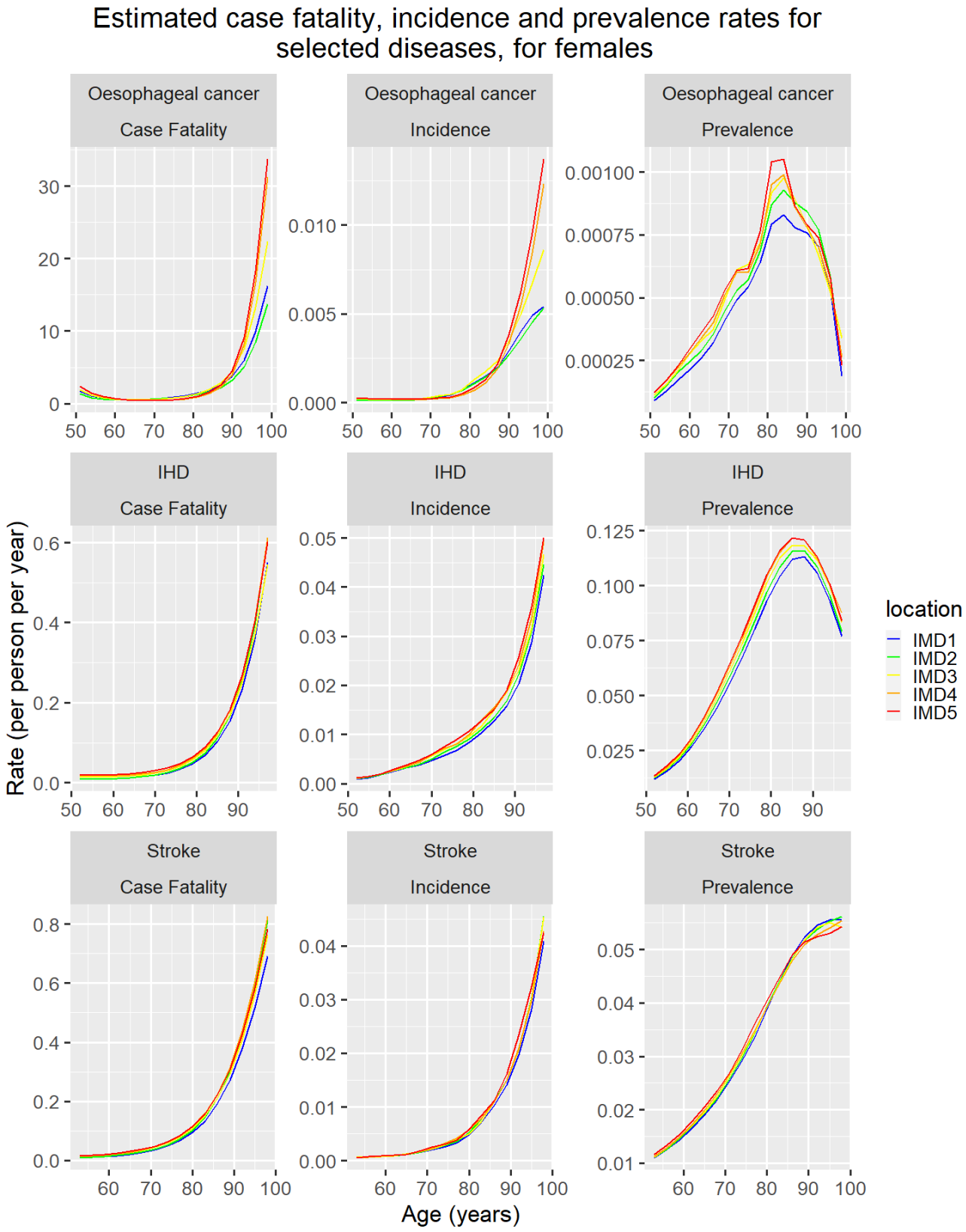


Figure 1c.3.2: Incidence, prevalence and case fatality rates of specified diseases, for females, by Index of Multiple Deprivation quintile (IMD 1 (least deprived) to 5 (most deprived)).

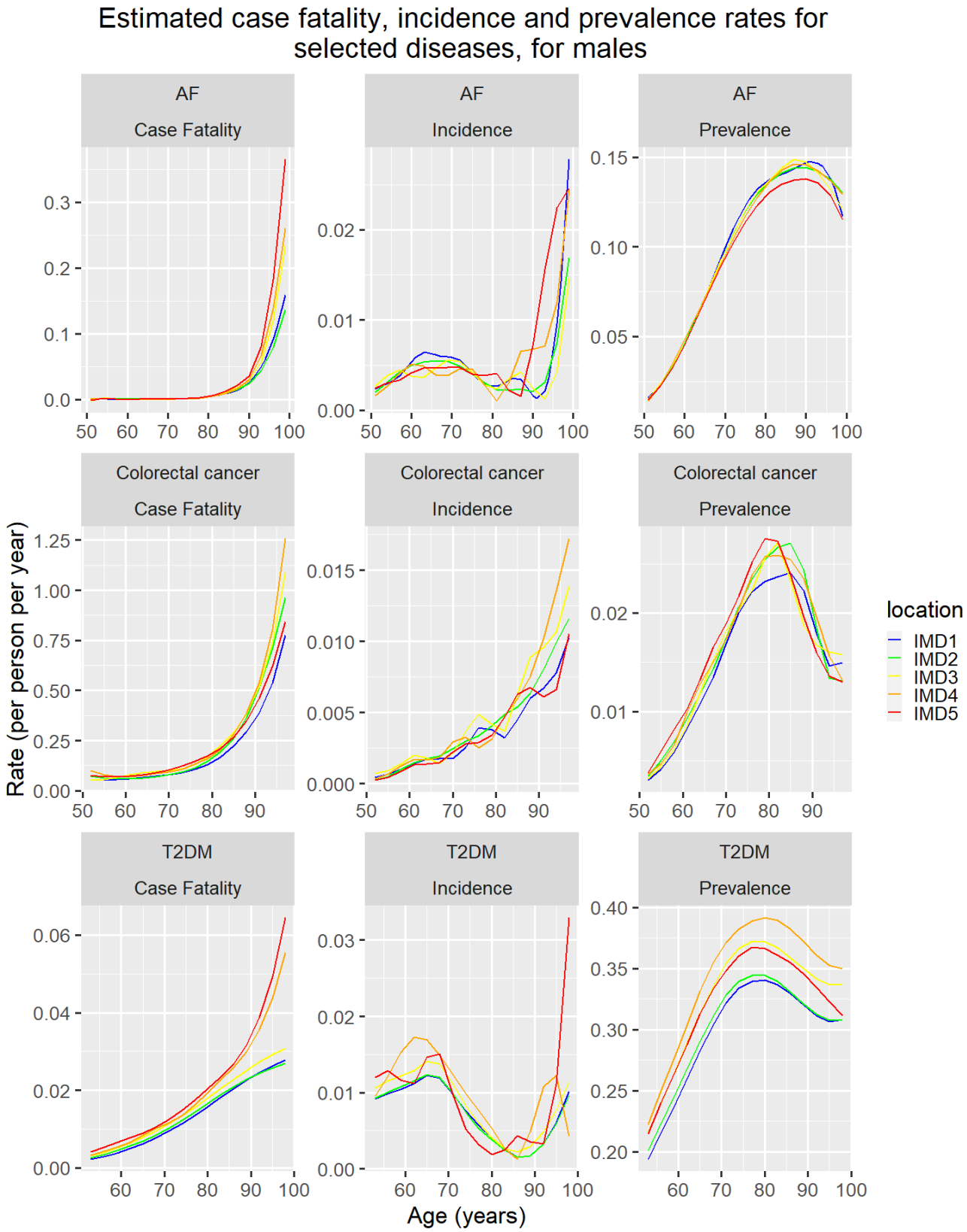


Figure 1c.3.3: Incidence, prevalence and case fatality rates of specified diseases, for males, by Index of Multiple Deprivation quintile (IMD 1 (least deprived) to 5 (most deprived)).

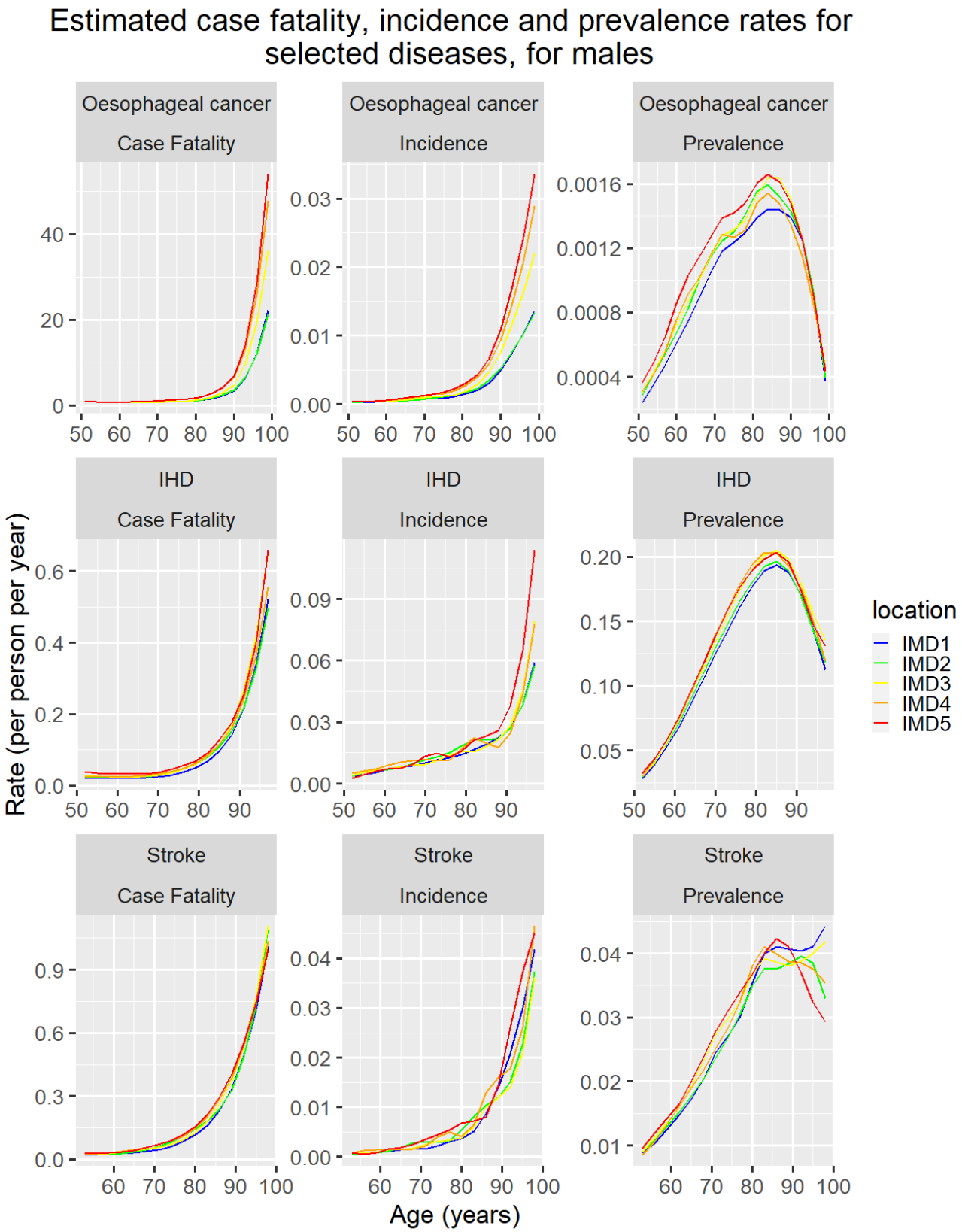


Figure 1c.3.4: Incidence, prevalence and case fatality rates of specified diseases, for males, by Index of Multiple Deprivation quintile (IMD 1 (least deprived) to 5 (most deprived)).

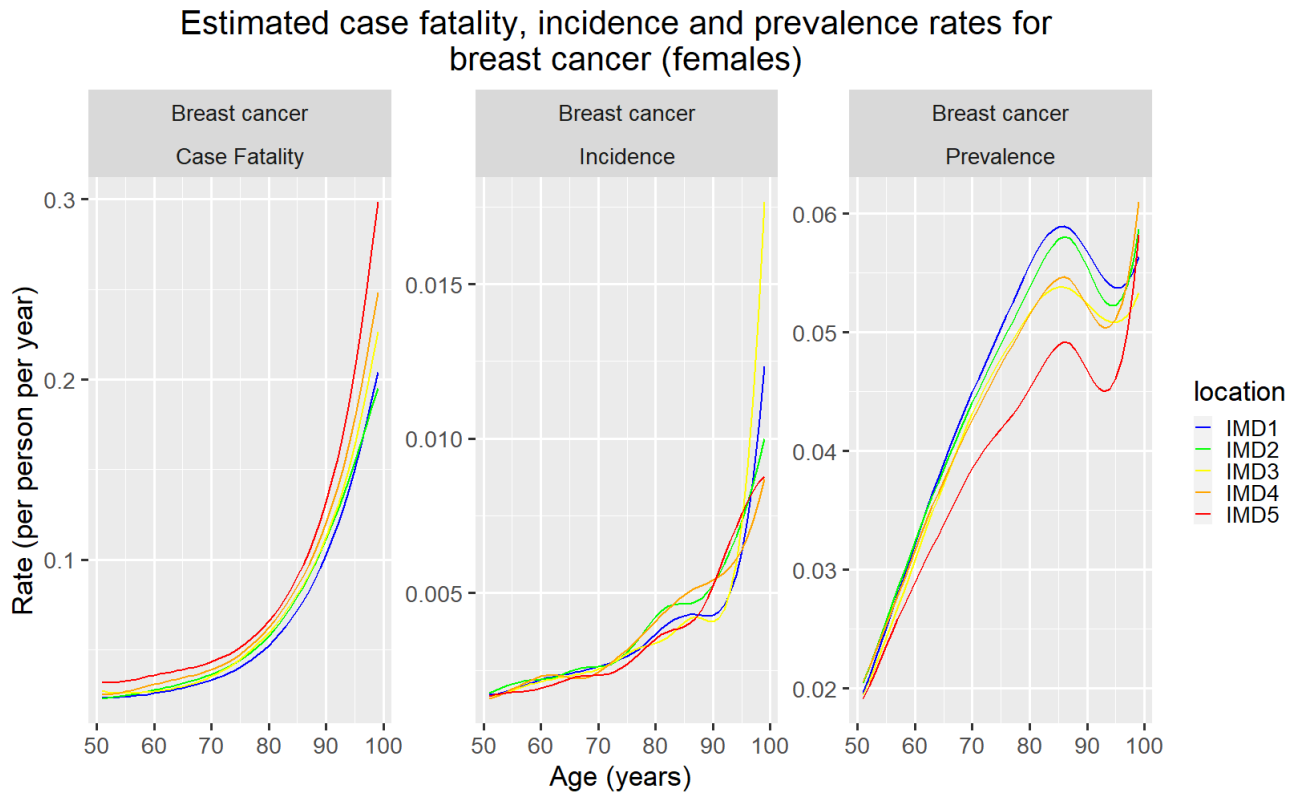


Figure 1c.3.5: Incidence, prevalence and case fatality rates of breast cancer (females), by Index of Multiple Deprivation quintile (IMD 1 (least deprived) to 5 (most deprived)).

Appendix 1c.4: Examination of outliers

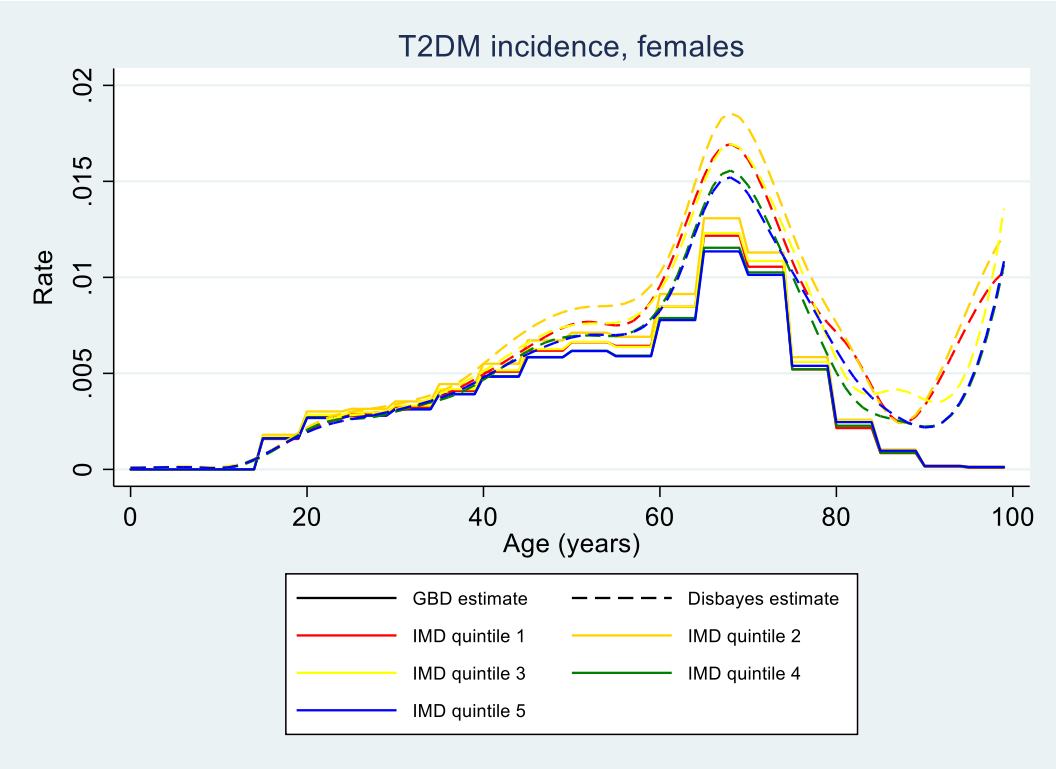


Figure 1c.4.1: Comparison of GBD and Disbayes estimates for Type-2 Diabetes incidence in females

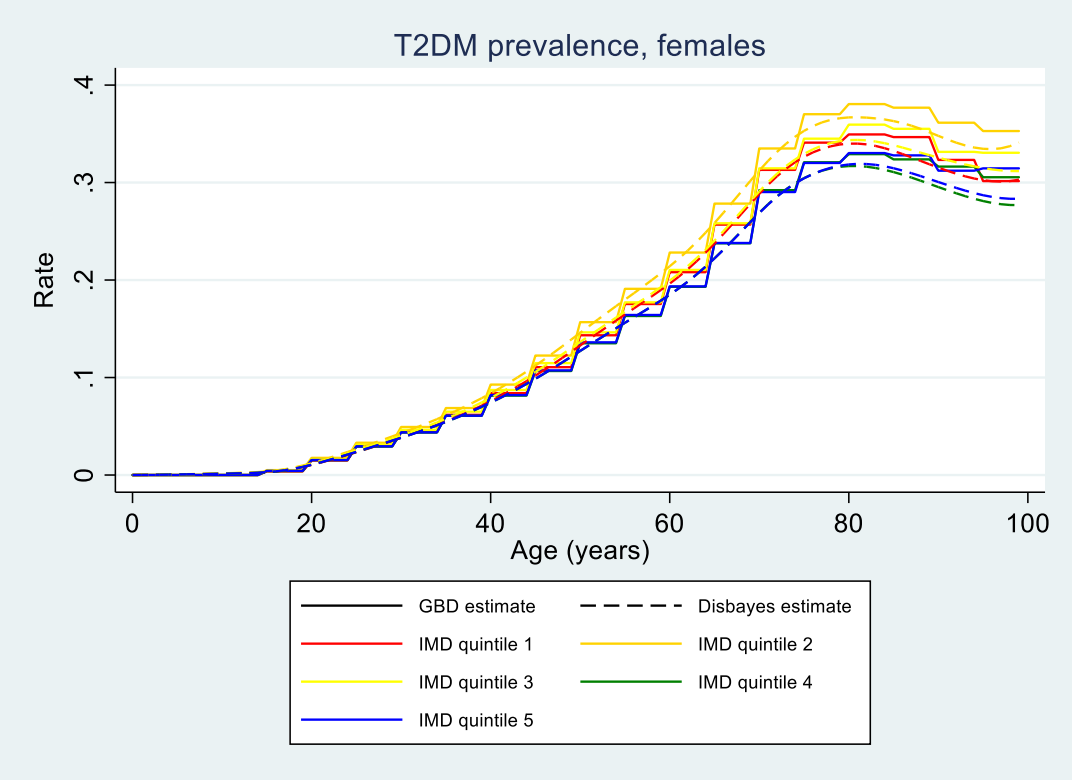


Figure 1c.4.2: Comparison of GBD and Disbayes estimates for Type-2 Diabetes prevalence in females

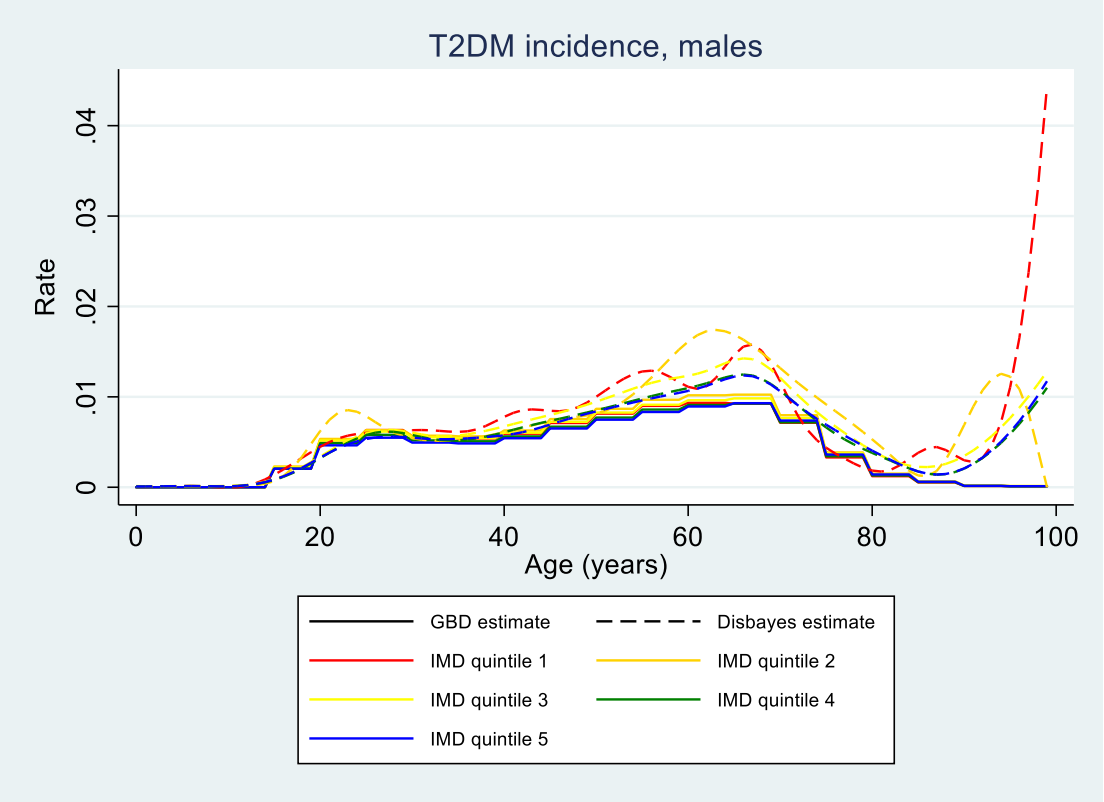


Figure 1c.4.3: Comparison of GBD and Disbayes estimates for Type-2 Diabetes incidence in males

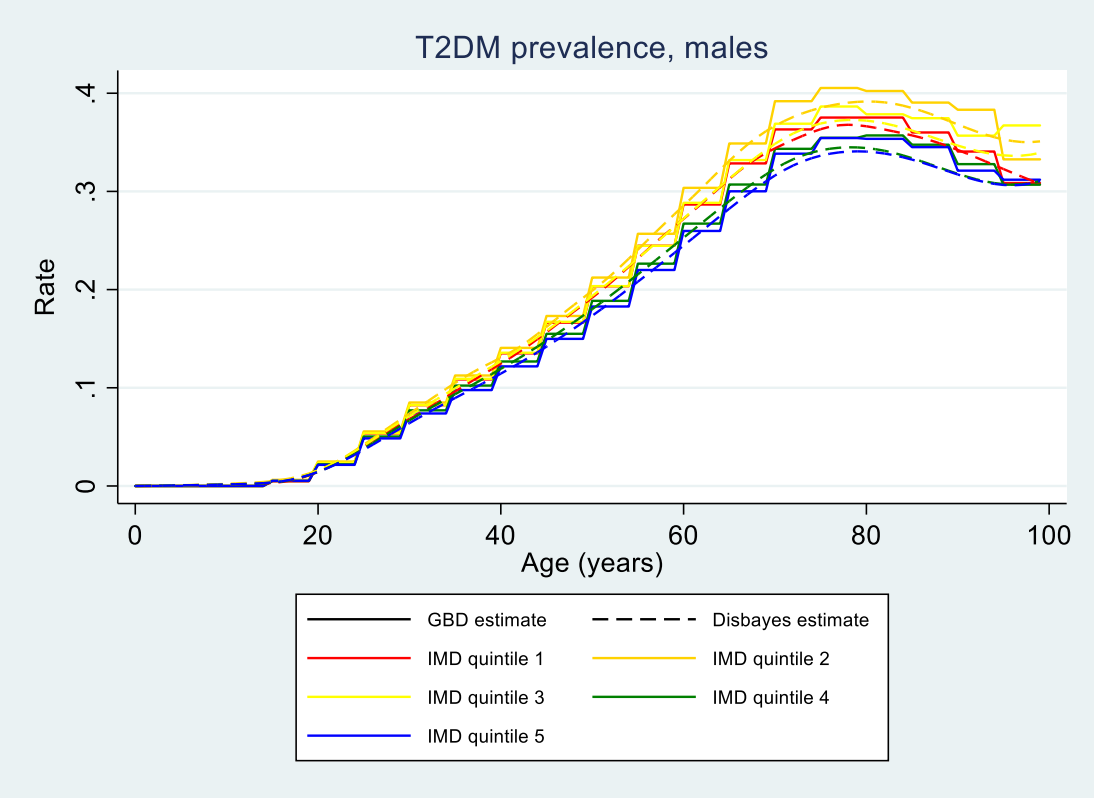


Figure 1c.4.4: Comparison of GBD and Disbayes estimates for Type-2 Diabetes prevalence in males

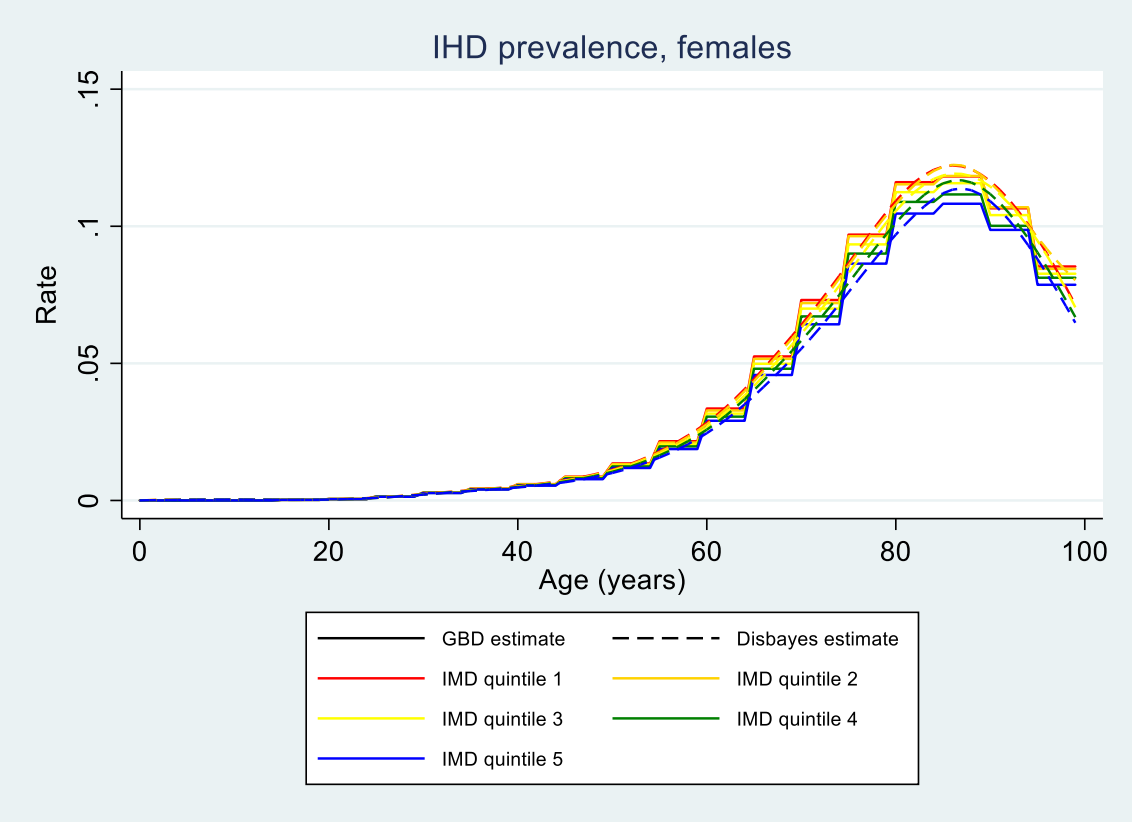


Figure 1c.4.5: Comparison of GBD and Disbayes estimates for Ischaemic Heart Disease prevalence in females.

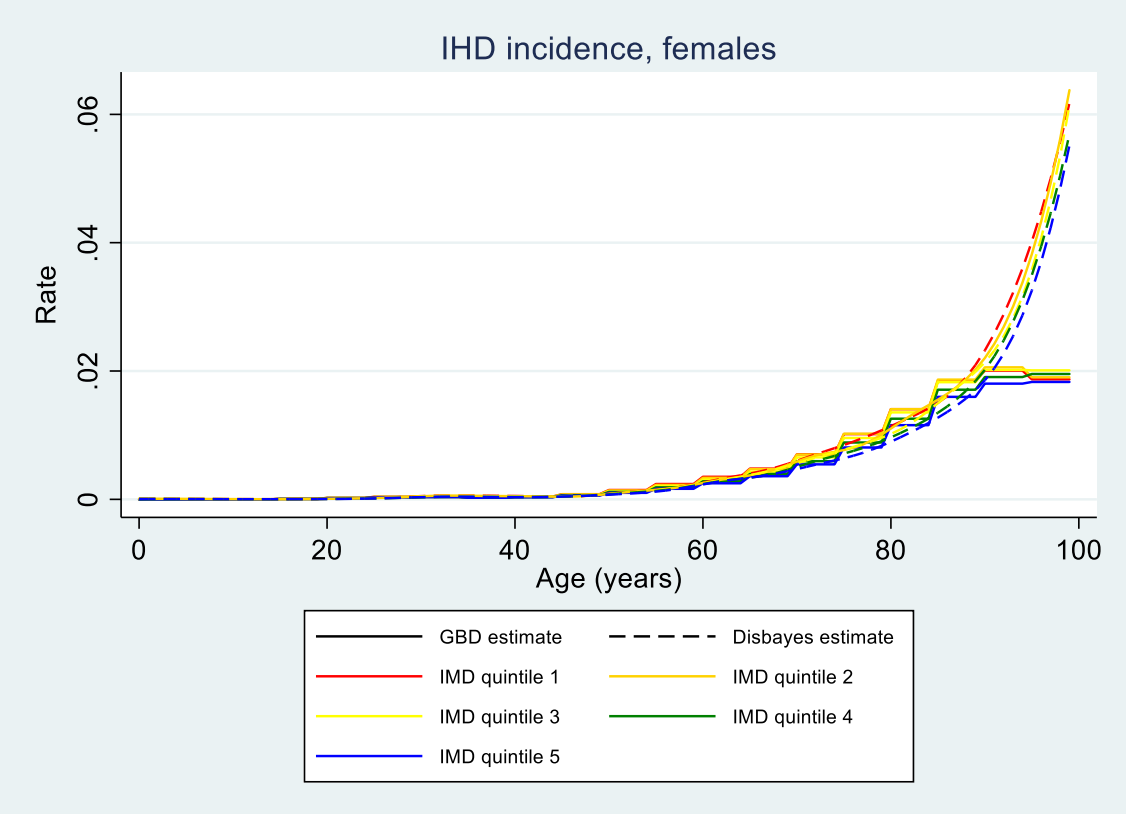


Figure 1c.4.6: Comparison of GBD and Disbayes estimates for Ischaemic Heart Disease incidence in females.

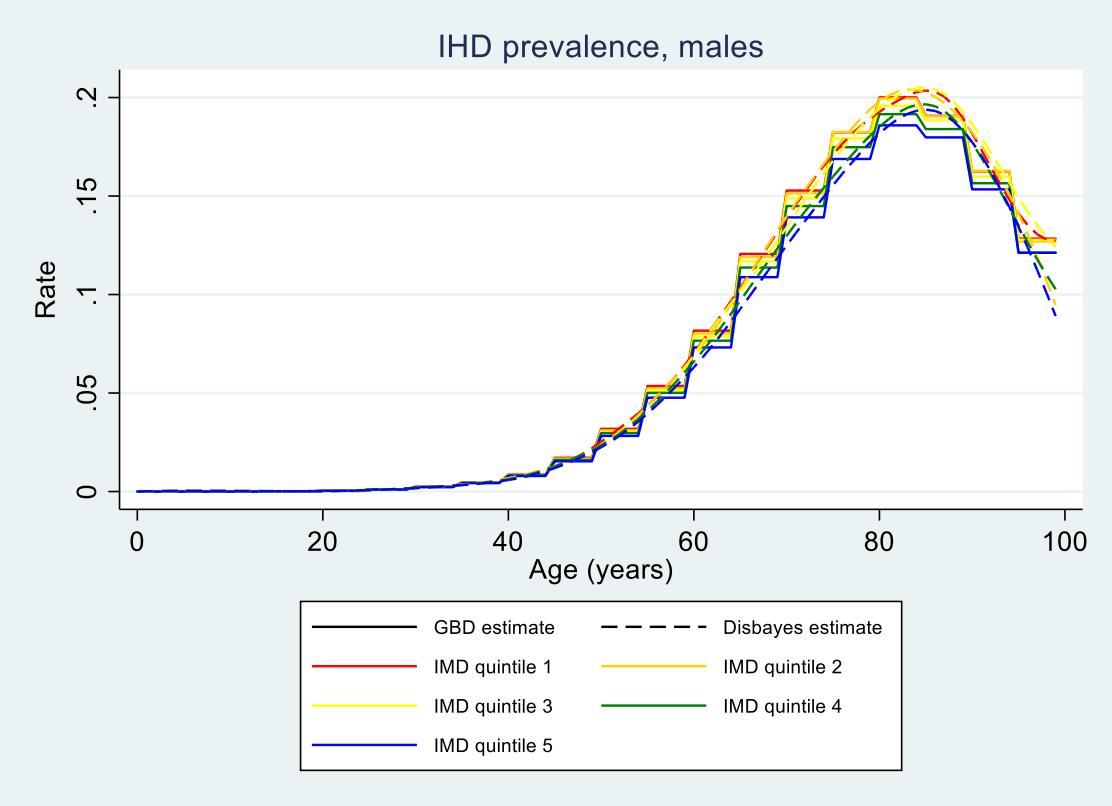


Figure 1c.4.5: Comparison of GBD and Disbayes estimates for Ischaemic Heart Disease prevalence in males.

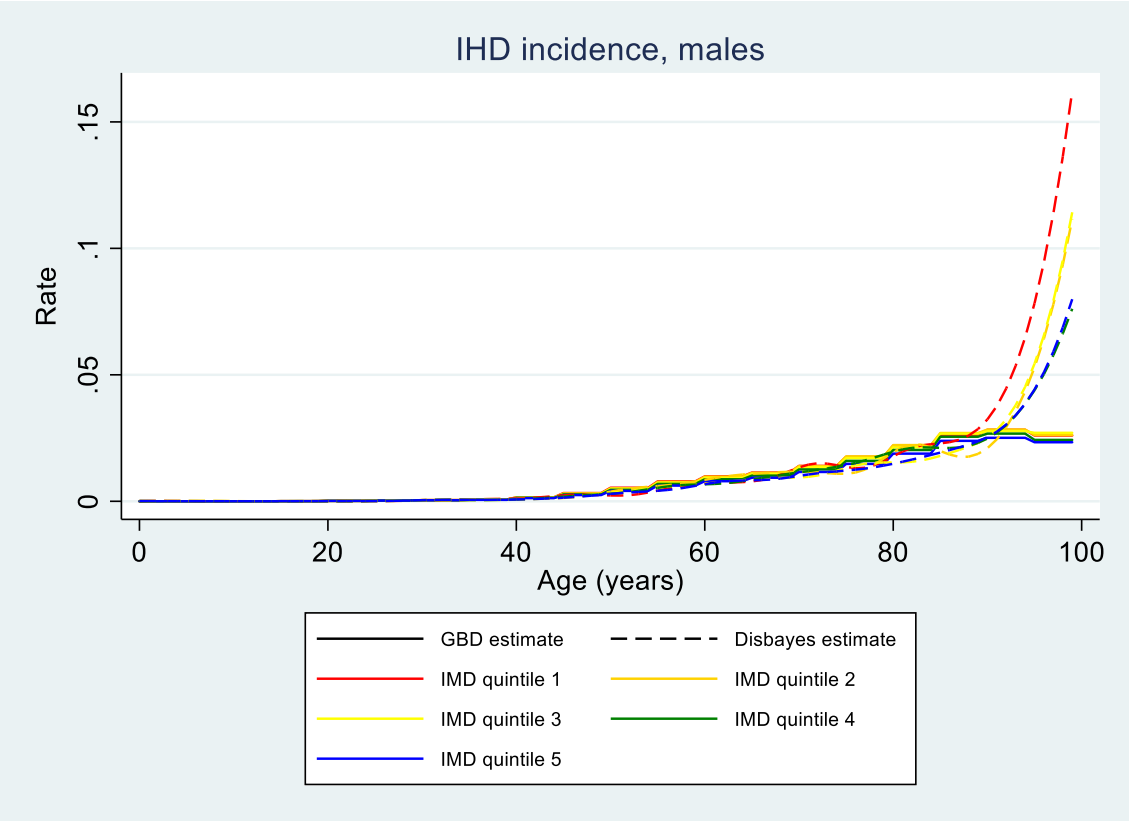


Figure 1c.4.6: Comparison of GBD and Disbayes estimates for Ischaemic Heart Disease incidence in males.

Appendix 1d: Chapter 5 appendices

Appendix 1d.1: Variable management

Appendix table 1d.1: Guide to recoding of variables in the Census microdata and HSE 2017 to make equivalent variables used in the final model.

Characteristic	Census 2011 microdata				
	Variable name	Census value	Census label	New value	New Label
Age	**AGEC**	1	0-4 year olds	(Missing)	
		2	5-9 year olds	(Missing)	
		3	10-15 year olds	(Missing)	
		4	16-18 year olds	17.5	16-18 year olds
		5	19-24 year olds	22	19-24 year olds
		6	25-29 year olds	27.5	25-29 year olds
		7	30-34 year olds	32.5	30-34 year olds
		8	35-39 year olds	37.5	35-39 year olds
		9	40-44 year olds	42.5	40-44 year olds
		10	45- 49 year olds	47.5	45- 49 year olds
		11	50- 54 year olds	52.5	50- 54 year olds
		12	55-59 year olds	57.5	55-59 year olds
		13	60-64 year olds	62.5	60-64 year olds
		14	65-69 year olds	67.5	65-69 year olds
		15	70-74 year olds	72.5	70-74 year olds

		16	75-79 year olds	77.5	75-79 year olds
		17	80-84 year olds	82.5	80-84 year olds
		18	85-89 year olds	87.5	85-89 year olds
		19	90+	F=93.0 M=92.4(1)	90+ year olds
Sex	**SEX**	1	male	0	male
		2	female	1	female
Ethnicity	**AGGETHPUK113**	1	White	1	White
		2	Mixed/multiple ethnic groups	4	Mixed/multiple
		3	Asian/Asian British: Indian	3	Asian
		4	Asian/Asian British: Pakistani	3	Asian
		5	Asian/Asian British: Bangladeshi	3	Asian
		6	Asian/Asian British: Chinese	3	Asian
		7	Asian/Asian British: Other Asian	3	Asian
		8	Black/African/Caribbean/Black British: African	2	Black
		9	Black/African/Caribbean/Black British: Black Caribbean	2	Black
		10	Black/African/Caribbean/Black British: Other Black	2	Black
		11	Other ethnic group	5	Any other group
Unemployed	**ECOPUK1**	1	Econ. Active (excl. Students), Employee, Part-time	0	Not unemployed
		2	Econ. Active (excl. Students), Employee, Full-time	0	Not unemployed
		3	Econ. Active (excl. Students), Self-employed with employees, Part-time	0	Not unemployed
		4	Econ. Active (excl. Students), Self-employed with employees, Full-time	0	Not unemployed
		5	Econ. Active (excl. Students), Self-employed no employees, Part-time	0	Not unemployed

		6	Econ. Active (excl. Students), Self-employed no employees, Full-time	0	Not unemployed
		7	Econ. Active (excl. Students), Unemployed(2)	1	Unemployed
		8	Economically Active Full-time Students, In Employment	0	Not unemployed
		9	Economically Active Full-time Students, Unemployed(2)	1	Unemployed
		10	Economically Inactive, Retired	0	Not unemployed
		11	Economically Inactive, Student	0	Not unemployed
		12	Economically Inactive, Looking after home/family	0	Not unemployed
		13	Economically Inactive, Permanently sick/disabled	0	Not unemployed
		14	Economically Inactive, Other	0	Not unemployed
Student	**ECOPUK1**	1	Econ. Active (excl. Students), Employee, Part-time	0	Not a student
		2	Econ. Active (excl. Students), Employee, Full-time	0	Not a student
		3	Econ. Active (excl. Students), Self-employed with employees, Part-time	0	Not a student
		4	Econ. Active (excl. Students), Self-employed with employees, Full-time	0	Not a student
		5	Econ. Active (excl. Students), Self-employed no employees, Part-time	0	Not a student
		6	Econ. Active (excl. Students), Self-employed no employees, Full-time	0	Not a student
		7	Econ. Active (excl. Students), Unemployed(2)	0	Not a student
		8	Economically Active Full-time Students, In Employment	1	Student
		9	Economically Active Full-time Students, Unemployed(2)	1	Student
		10	Economically Inactive, Retired	0	Not a student
		11	Economically Inactive, Student	1	Student

		12	Economically Inactive, Looking after home/family	0	Not a student
		13	Economically Inactive, Permanently sick/disabled	0	Not a student
		14	Economically Inactive, Other	0	Not a student
Has a degree or equivalent	**hlqupuk11**	10	No academic or professional qualifications	0	
		11	Level 1 (0-4 GCSE, O level, or equivalents)	0	
		12	Level 2 (5+ GCSE, O level, 1 A level, or equivalents)	0	
		13	Apprenticeship	0	
		14	Level 3 (2+ A levels, or equivalents)	0	
		15	Level 4+ (degree, postgrad, professional quals)	1	Has a degree or equivalent
		16	Other (vocational/foreign/outside UK quals)	0	
		-9	Not applicable	(Missing)	
Self-reported general health	**HEALTH**	1	Very good	1	Very good
		2	Good	2	Good
		3	Fair	3	Fair
		4	Bad	4	Bad
		5	Very bad	5	Very bad
Deprivation	**IMDQUINTE**	1	Least deprived	1	Least deprived
		2	Second least	2	Second least
		3	Intermediate	3	Intermediate
		4	Second most	4	Second most
		5	Most deprived	5	Most deprived
		-8	unknown	(Missing)	
	HSE 2017				
	Variable name	**HSE value**	**HSE label**	**New value**	**New Label**
Age	**Age16g10**	1	0-1 year olds	(Missing)	

		2	2-4 year olds	(Missing)	
		3	5-7 year olds	(Missing)	
		4	8-10 year olds	(Missing)	
		5	11-12 year olds	(Missing)	
		6	13-15 year olds	(Missing)	
		7	16-19 year olds	18	16-19 year olds
		8	20-24 year olds	22.5	20-24 year olds
		9	25-29 year olds	27.5	25-29 year olds
		10	30-34 year olds	32.5	30-34 year olds
		11	35-39 year olds	37.5	35-39 year olds
		12	40-44 year olds	42.5	40-44 year olds
		13	45- 49 year olds	47.5	45- 49 year olds
		14	50- 54 year olds	52.5	50- 54 year olds
		15	55-59 year olds	57.5	55-59 year olds
		16	60-64 year olds	62.5	60-64 year olds
		17	65-69 year olds	67.5	65-69 year olds
		18	70-74 year olds	72.5	70-74 year olds
		19	75-79 year olds	77.5	75-79 year olds
		20	80-84 year olds	82.5	80-84 year olds
		21	85-89 year olds	87.5	85-89 year olds
		22	90+	F=93.0, M=92.4(1)	90+ year olds
Sex	**Sex**	1	Male	0	male
		2	Female	1	female
		3	Refused	(Missing)	
		4	Don’t know	(Missing)	
		5	Not applicable	(Missing)	
Ethnicity	**Origin2**	1	White	1	White

		2	Black	2	Black
		3	Asian	3	Asian
		4	Mixed/multiple ethnic background	4	Mixed/multiple
		5	Any other ethnic group	5	Any other group
		6	Refusal	(Missing)	
		7	Don't know	(Missing)	
Unemployed	**Activb2**	1	Going to school or college full-time	0	Not unemployed
		2	In paid employment or self-employed	0	Not unemployed
		3	Doing unpaid work for family business	0	Not unemployed
		4	Waiting to take up paid work already obtained	0	Not unemployed
		5	Looking for paid work/ Govt. training scheme	1	Unemployed
		6	Intending to look for work, but temporary sickness	0	Not unemployed
		7	Permanently unable to work, long-term sickness	0	Not unemployed
		8	Retired from paid work	0	Not unemployed
		9	Looking after home or family	0	Not unemployed
		95	Doing something else (specify)	0	Not unemployed
		-9	No answer/refused	(Missing)	
		-8	Don't know	(Missing)	
		-1	Not applicable	(Missing)	
Student	**Activb2**	1	Going to school or college full-time	1	Student
		2	In paid employment or self-employed	0	Not a student
		3	Doing unpaid work for family business	0	Not a student
		4	Waiting to take up paid work already obtained	0	Not a student
		5	Looking for paid work/ Govt. training scheme	0	Not a student
		6	Intending to look for work, but temporary sickness	0	Not a student

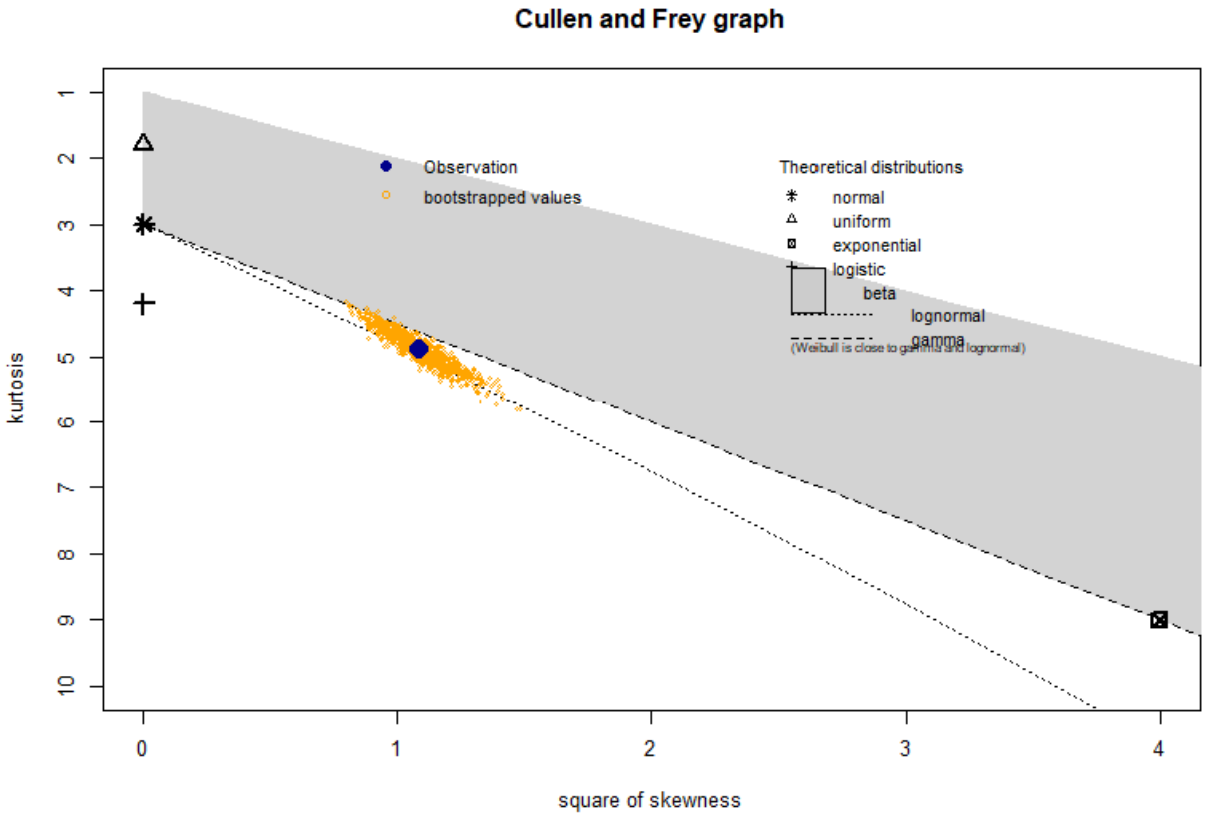
		7	Permanently unable to work, long-term sickness	0	Not a student
		8	Retired from paid work	0	Not a student
		9	Looking after home or family	0	Not a student
		95	Doing something else (specify)	0	Not a student
		-9	No answer/refused	(Missing)	
		-8	Don't know	(Missing)	
		-1	Not applicable	(Missing)	
Has a degree or equivalent	**topqual3**	1	NVQ4/NVQ5/Degree or equiv	1	Has a degree or equivalent
		2	Higher ed below degree	0	
		3	NVQ3/GCE A Level equiv	0	
		4	NVQ2/GCE O Level equiv	0	
		5	NVQ1/CSE other grade equiv	0	
		6	Foreign/other	0	
		7	No qualification	0	
		-9	Refused	(Missing)	
		-8	Don't know	(Missing)	
		-1	Not applicable	(Missing)	
Self-reported general health	**GenHelf**	1	Very good	1	Very good
		2	Good	2	Good
		3	Fair	3	Fair
		4	Bad	4	Bad
		5	Very bad	5	Very bad
		-9	Refused	(Missing)	
		-8	Don't know	(Missing)	
		-1	Not applicable	(Missing)	
Deprivation	**(Merged from external dataset(3))**	1	Least deprived – 0.48 - 8.37	1	Least deprived
		2	Second least – 8.37 - 13.92	2	Second least
		3	Intermediate – 13.92- 21.43	3	Intermediate
		4	Second most – 21.43 - 33.88	4	Second most
		5	Most deprived – 33.88 - 92.60	5	Most deprived

(1) Average age for over-90s in most recent ONS mid-year estimates of the England population is 93.0 years for females and 92.4 years for males.

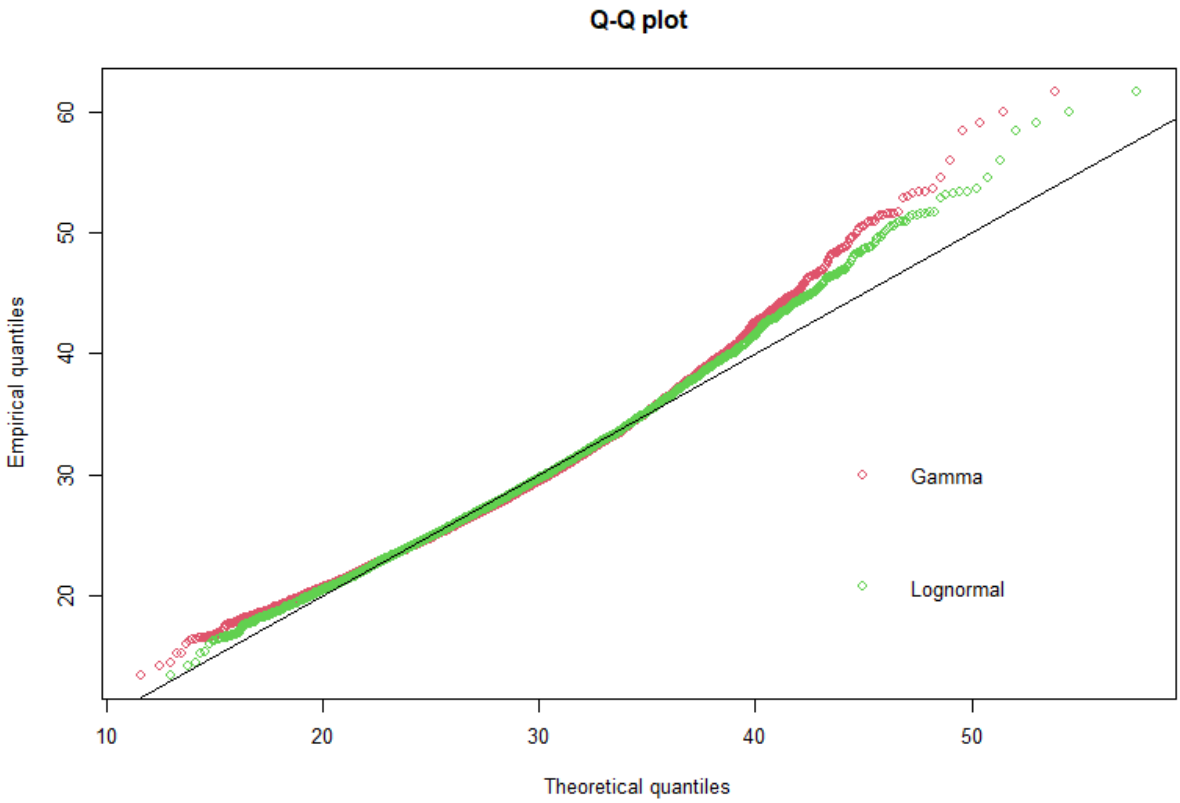
(2) **International Labour Organisation definition: seeking work and ready to start in 2 weeks, and Waiting to start a job already obtained and available to start within 2 weeks.**(273)

(3) **English Index of Multiple Deprivation**(219)

Appendix 1d.2: Distribution fitting for GLM model

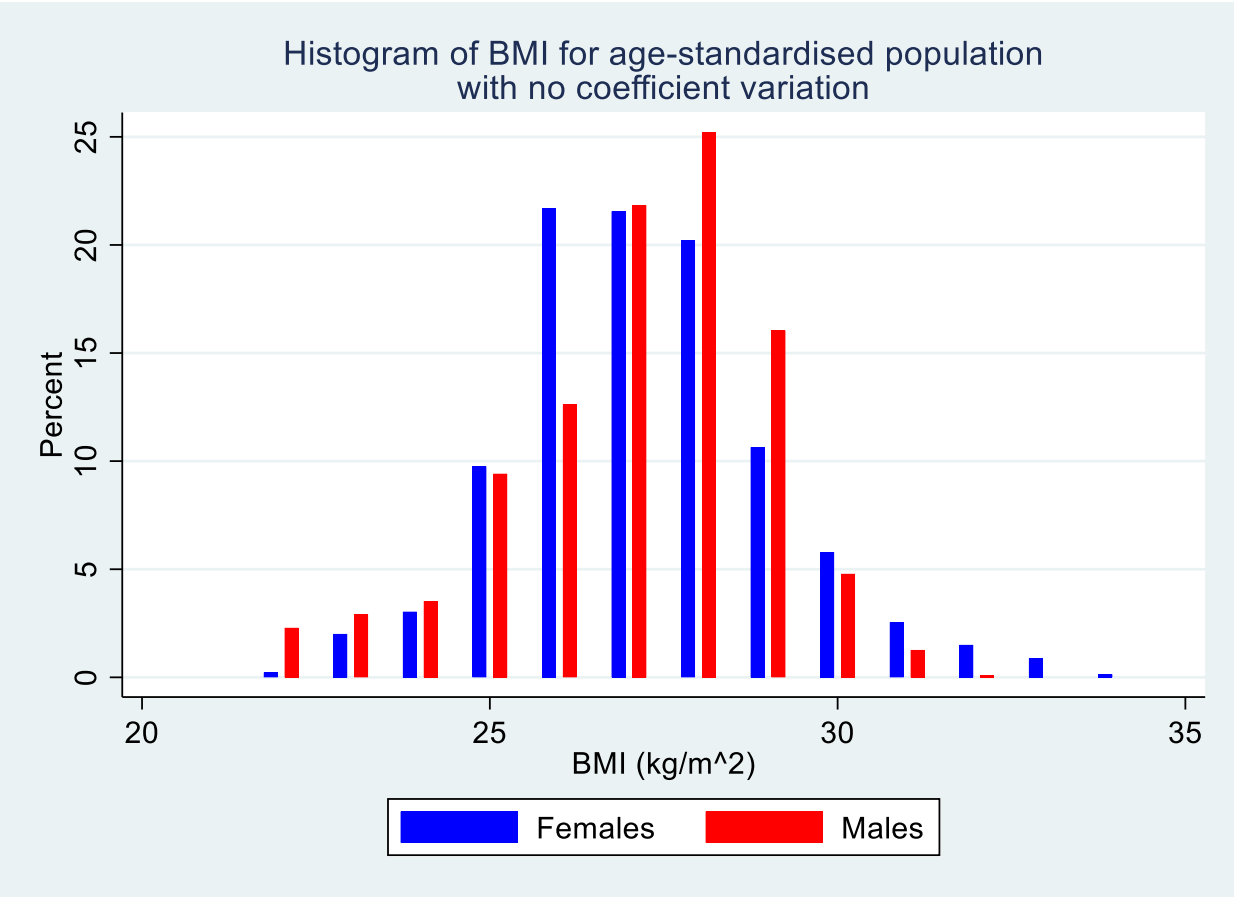


Appendix figure 1d.2a: Cullen-Frey plot for the identification of variance structure of the BMI variable.



Appendix figure 1d.2b: QQ plot comparing of the BMI variables against theoretical gamma and lognormal distributions.

Appendix 1d.3: BMI synthetic estimates without coefficient variation



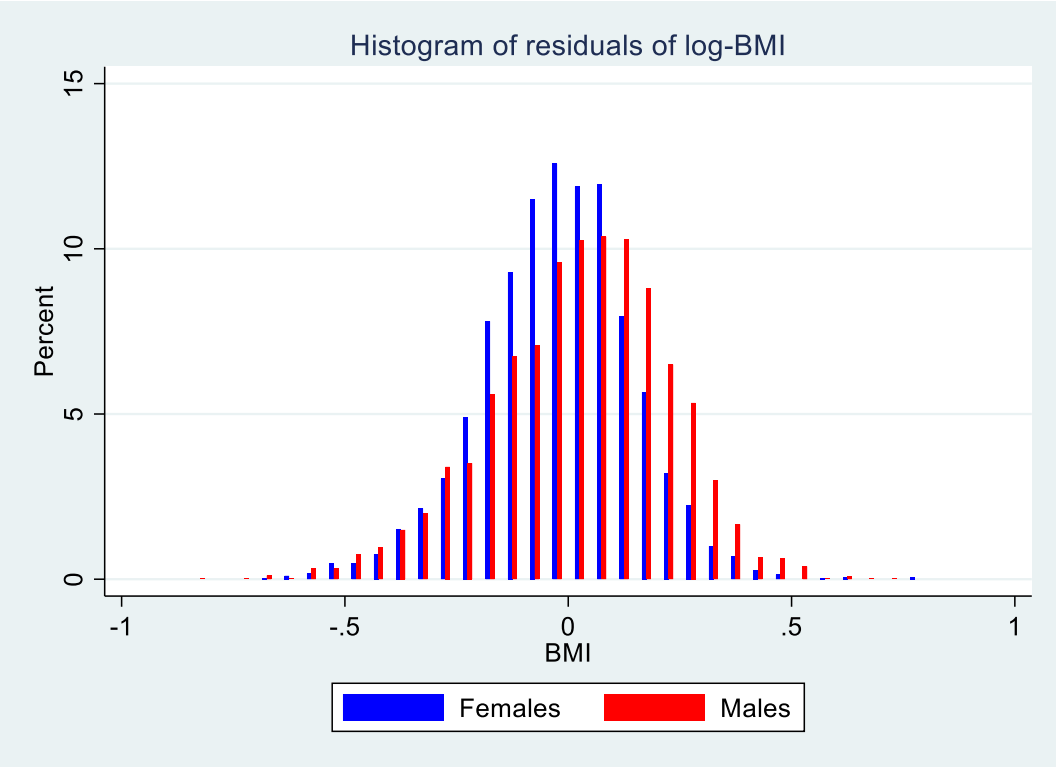
Appendix figure 1d.3: The distribution of estimated BMI generated without including variation in the model coefficients, by males and females.

Appendix 1d.4: Missing data

Appendix table 1d.4: Missing data in for the BMI variable in HSE

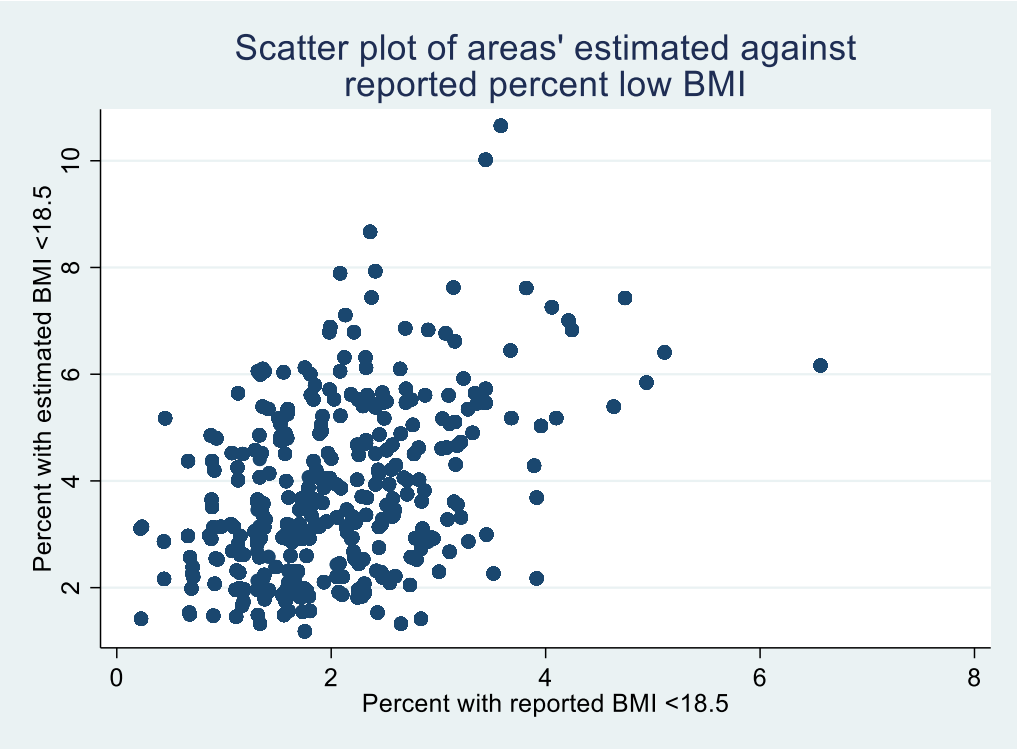
Breakdown variable	Total	Of which percent missing
Age		
18	281	17.8
22.5	366	19.4
27.5	487	17.5
32.5	581	18.1
37.5	647	14.2
42.5	684	15.8
47.5	668	15.3
52.5	686	17.9
57.5	733	16.2
62.5	634	17.5
67.5	614	14.3
72.5	625	20.0
77.5	444	22.3
82.5	298	30.5
87.5	177	36.2
90+	72	45.9
Sex		
Female	4461	18.7
Male	3536	17.8
IMD quintile		
1 (lowest deprivation)	1653	15.5
2	1707	18.3
3	1554	19.0
4	1557	19.1
5	1526	20.1

Appendix 1d.5: Histogram of residuals

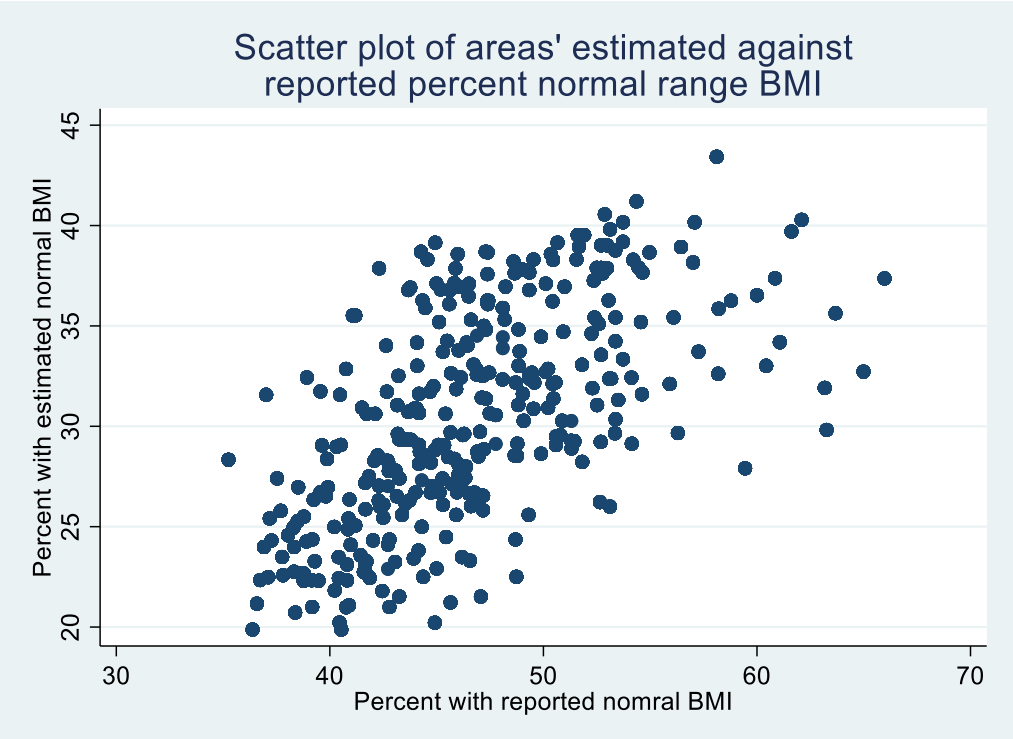


Appendix figure 1d.5: Histograms of residuals of between measured and estimated log-BMI to examine for normality.

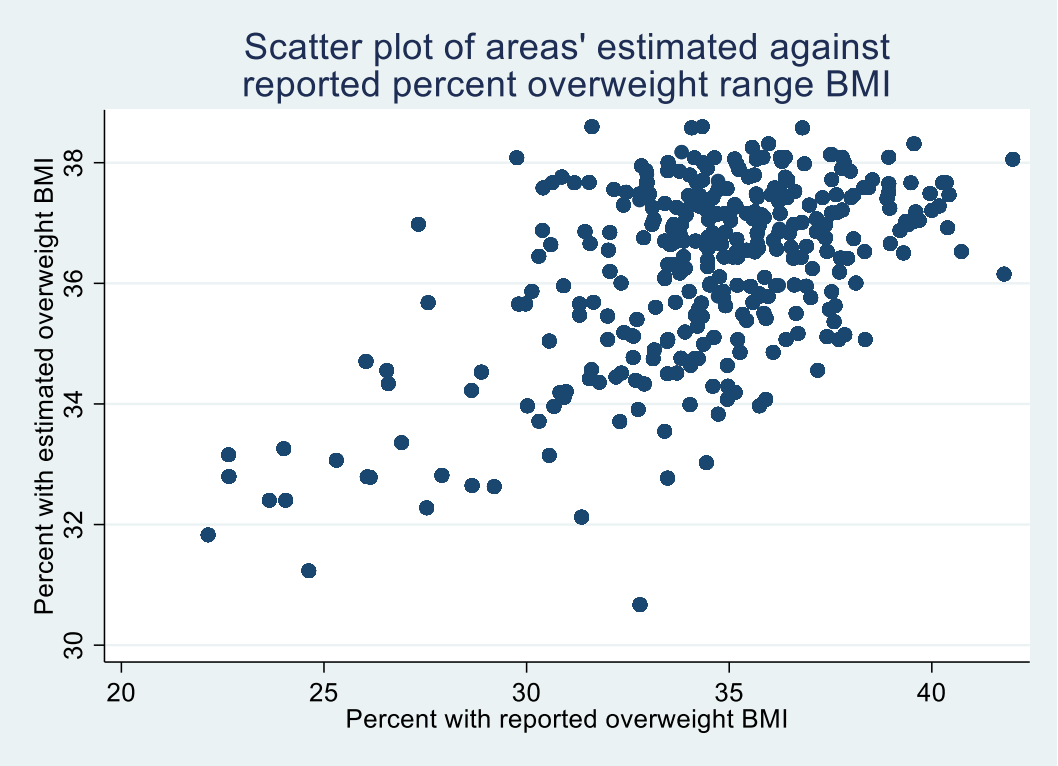
Appendix 1d.6: External validity scatter plots



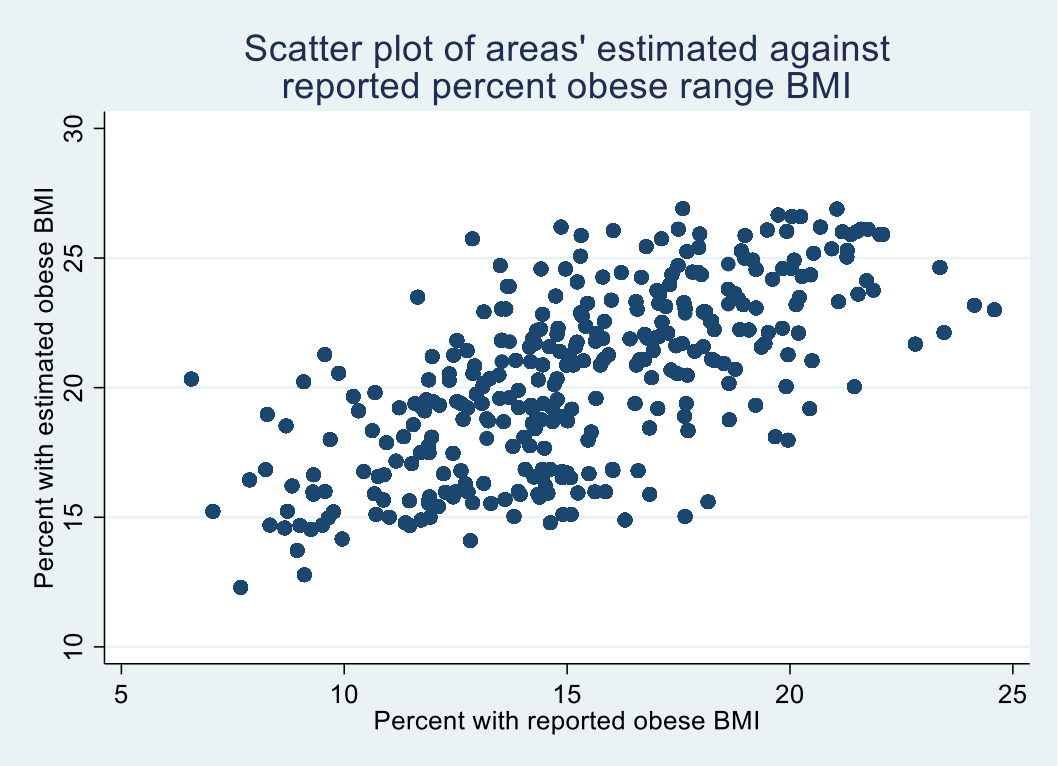
Appendix figure 1d.6a: Scatter plot of the relationship between local areas' estimated versus reported percent of people with BMI <18.5.



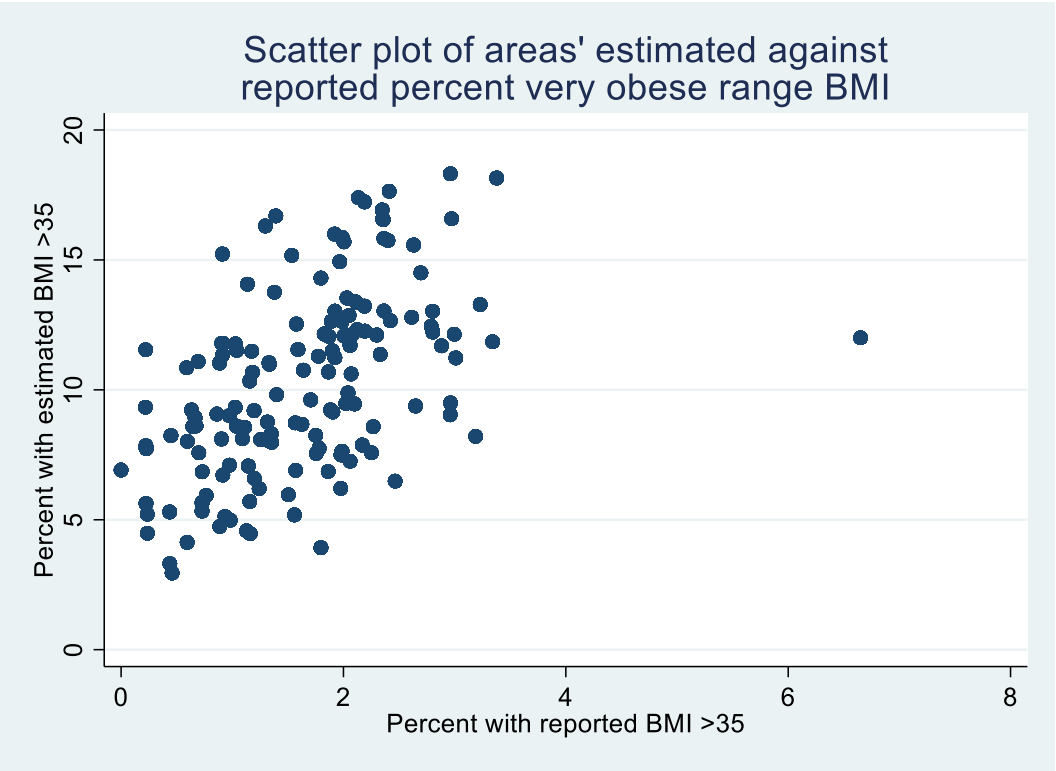
Appendix figure 1d.6b: Scatter plot of the relationship between local areas' estimated versus reported percent of people with BMI in the normal range.



Appendix figure 1d.6c: Scatter plot of the relationship between local areas' estimated versus reported percent of people with BMI in the overweight range.



Appendix figure 1d.6d: Scatter plot of the relationship between local areas' estimated versus reported percent of people with BMI in the obese range.

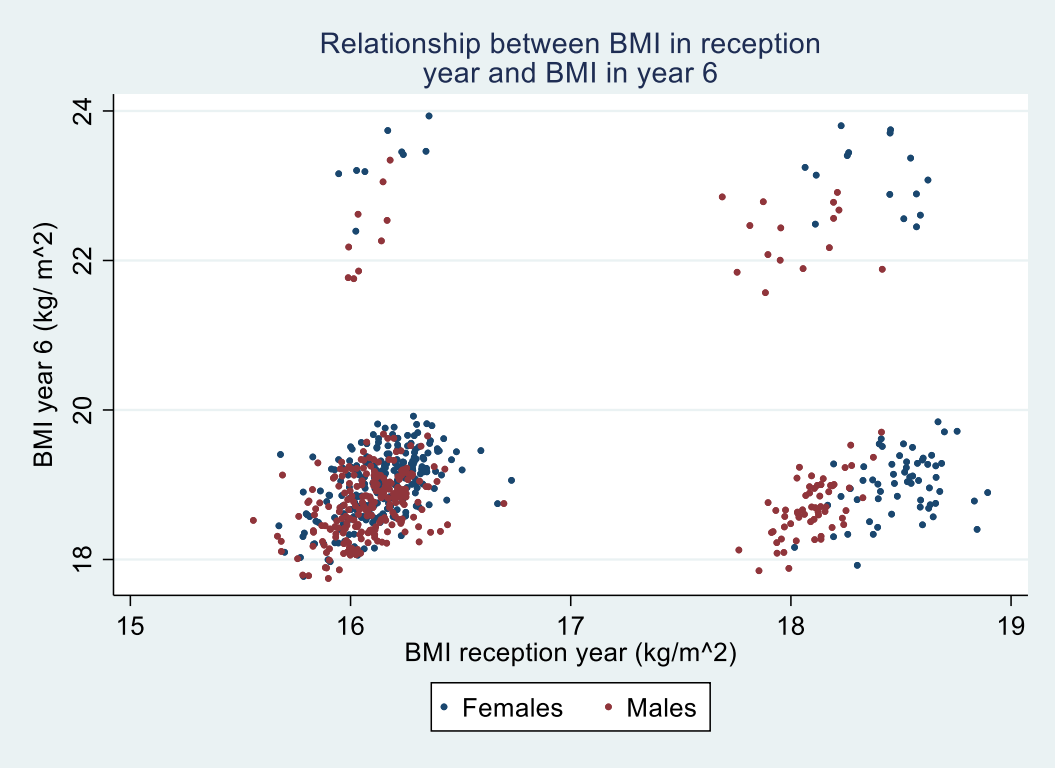


Appendix figure 1d.6e: Scatter plot of the relationship between local areas' estimated versus reported percent of people with BMI >35.

Appendix 1e: Chapter 6 appendices

Data censorship

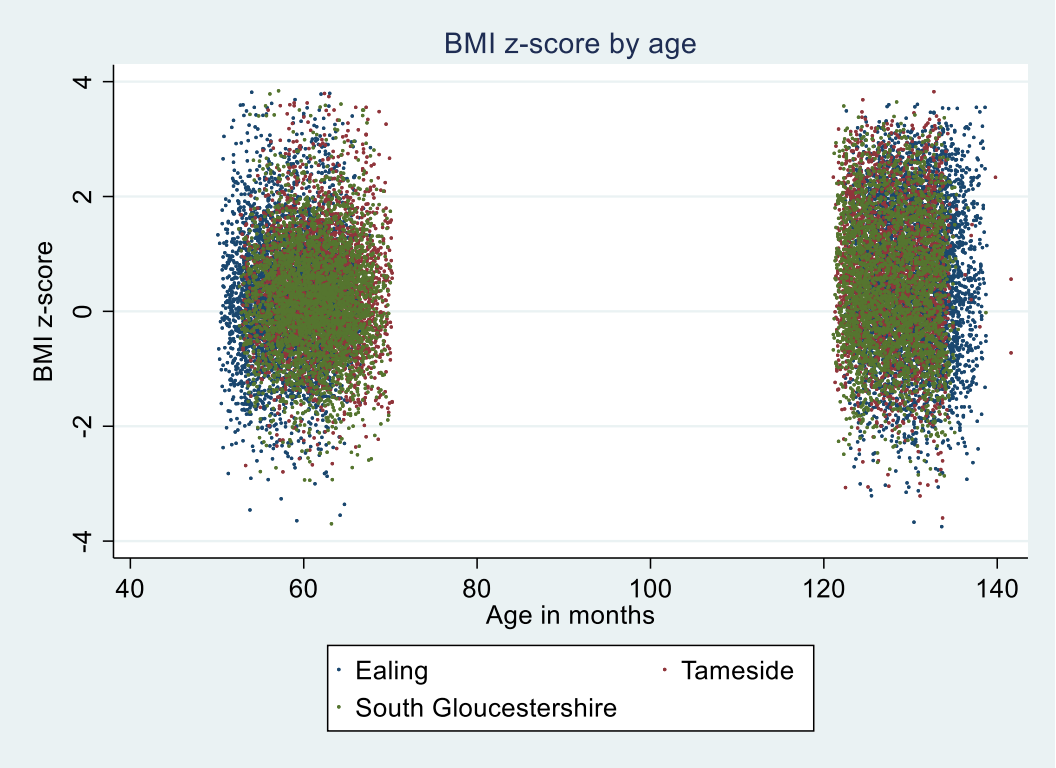
Figure 1e.1 shows the relationship between mean area BMI in reception year with mean area BMI in year 6 for the 2017-18 NCMP dataset. These show clearly bimodal distributions at both ages, lacking the smoothly continuous nature of natural data we would expect. On further inspection, this arises from censoring of microdata demonstrated in Figures 2-5. These show z-scores bounded to be no less than 1. The numbers of observations in that datasets and the numbers of censored values are consistent with those in the user guides.(7,151) Numbers of data points in each cluster are shown in Table 1e.1 labelling the bottom left cluster as 1, top left as 2, top right as 3 and bottom right as 4. The microdata scatter plot in Appendix figure 1e.2 is what we would expect to find: z-scores with a mean of zero and distributions stretching increasingly thinly to the extremes. Appendix figures 3, 4 and 5 show that observations in clusters 2, 3 and 4 have been censored below a z-score of 1 for one or both school years. NHS Digital were contacted and provided an explanation that where a category has fewer than five respondents it is censored, and the next category also censored to prevent the numbers <5 from being back-calculated.



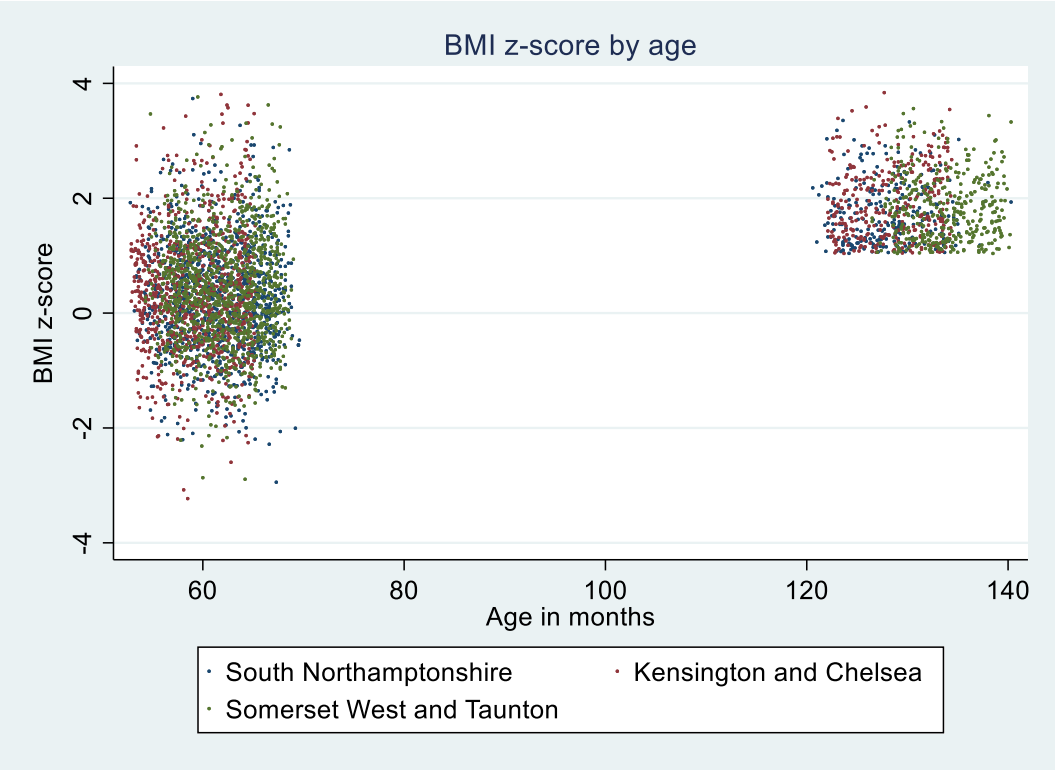
Appendix figure 1e.1: Scatter plot of area mean BMI in reception year against area mean BMI in year 6.

Appendix table 1e.1: numbers and percent of data points in each BMI cluster in each wave (labelling points on the bottom left of figure 1e.1 as cluster 1, top left as 2, top right as 3 and bottom right as 4).

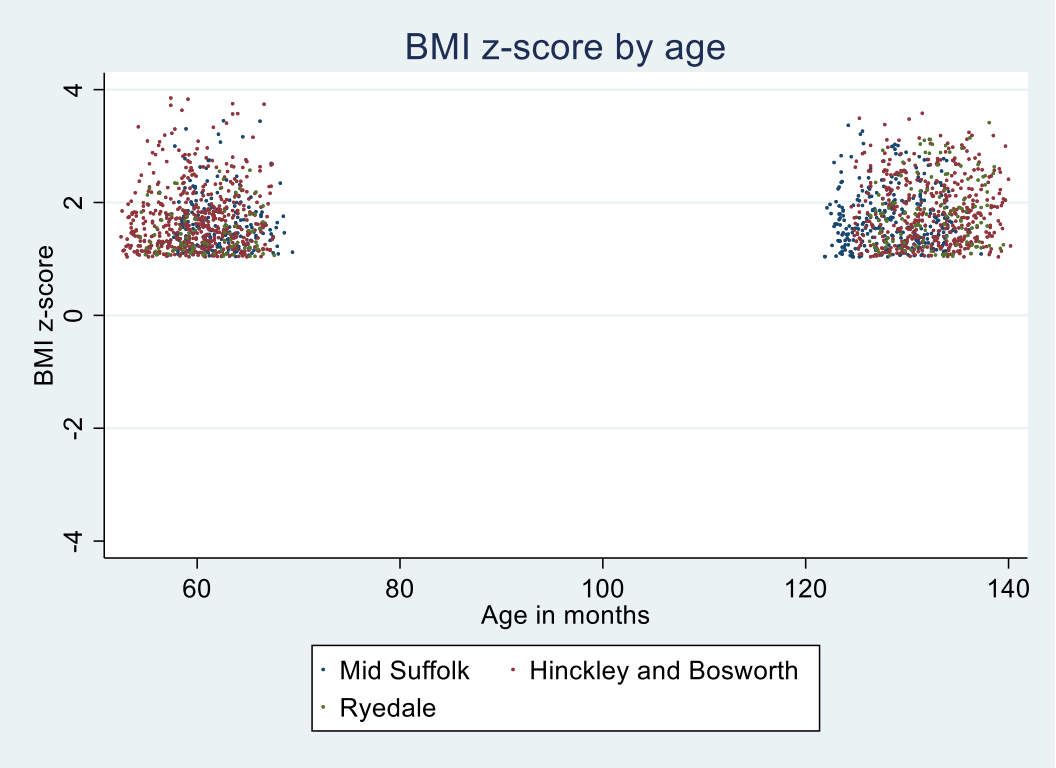
	2017-18		2018-19		Combined 2017-18/ 2018-19	
Cluster	Numbers	Percent	Numbers	Percent	Numbers	Percent
1	231	73%	221	70%	266	84%
2	9	3%	3	1%	0	0%
3	15	5%	13	4%	6	2%
4	60	19%	78	25%	43	14%



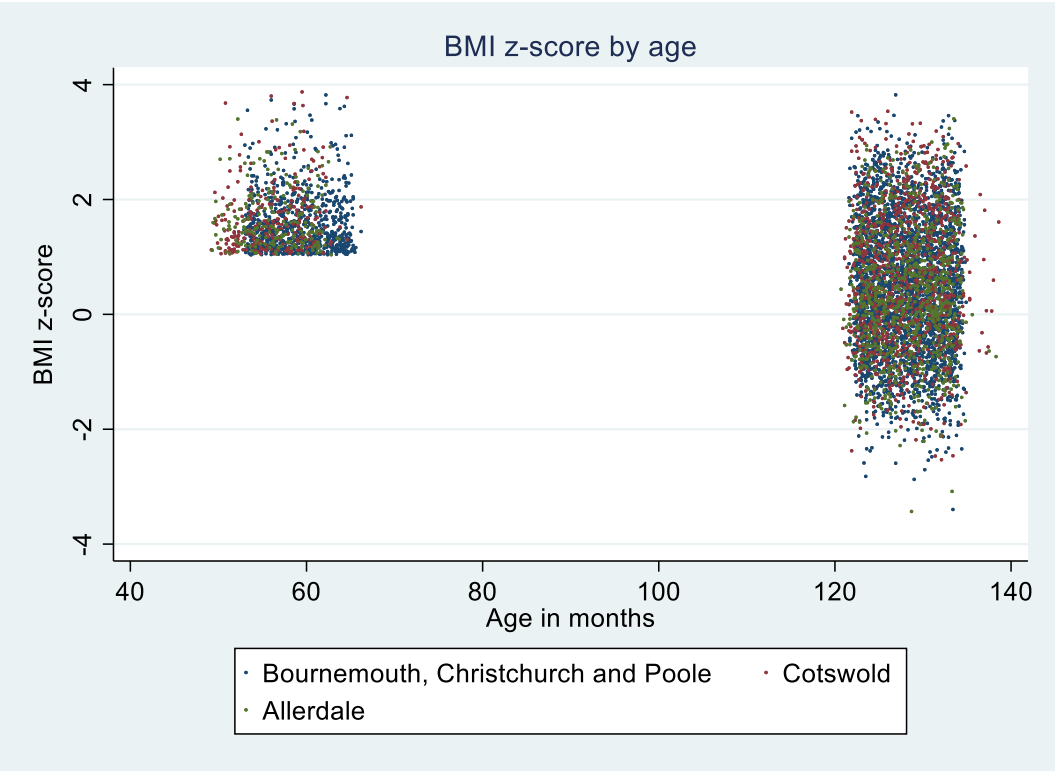
Appendix figure 1e.2: Scatter plot of age against BMI z-score for microdata from 3 areas selected from cluster 1.



Appendix figure 1e.3: Scatter plot of age against BMI z-score for microdata from 3 areas selected from cluster 2.



Appendix figure 1e.4: Scatter plot of age against BMI z-score for microdata from 3 areas selected from cluster 3.



Appendix figure 1e.5: Scatter plot of age against BMI z-score for microdata from 3 areas selected from cluster.

Appendix 1f: Chapter 7 appendices

Table 1f.1 shows outputs from crude regression (OLS) of average PLICS cost (crude mean cost (£) across the 12 disease states) for each UT local authority against its RCI (centred on 1), MFF index (centred on 1) and IMD score (ranging 5.85 to 54.0) in turn. The outputs of the regression model between RCI and MFF is also given for comparison. Figures 1f.1 - 1f.4 show these relationships, respectively, graphed as scatter plots.

The regression outputs show that unadjusted associations between PLICS costs with RCI or IMD are unlikely, though there was a moderate positive relationship between PLICS cost and MFF. There was not a significant relationship between RCI and MFF. That the results of the PLICS methods of estimating trust-level costs, but not the RCI, appeared to be associated with the independently-calculated measure of supply-side costs for those trusts (ie. the MFF) implies an improvement of data quality moving from RCI to PLICS, as it implies that randomness may have been reduced by the method, as intended. However, other factors clearly remain important: average PLICS costs were £3430, while an increase of one unit of MFF (representing 100% of costs) had a coefficient of £4980, implying that MFF over-predicts impacts of its weighted supply-side factors on costs overall. Other factors may therefore be important in driving variation in costs, such as demand-side factors (eg. complexity), unmeasured supply-side factors (such as unmeasured dimensions of skills) or simply non-random misreporting (though misreporting would usually be expected to be random and therefore symmetrical, and not biasing the overall prediction). Adjusting with combinations of two or all three of these factors did not reveal further insights.

Table 1f.1: Regression model outputs for models between PLICS, RCI, MFF and IMD.

				95% confidence intervals:		
Model:	Coefficient	SE	p-value	Lower	Upper	Adjusted R-squared
Model 1:						
PLICS cost RCI	-498	1640	0.762	-3740	2750	-0.0062
Model 2:						
PLICS cost MFF	4980	1280	<0.000	2450	7520	0.0868
Model 3:						
PLICS cost IMD	3.98	11.4	0.732	-18.6	26.3	-0.006
Model 4:						
RCI MFF	-0.113	0.670	0.094	-0.245	0.0195	0.0123

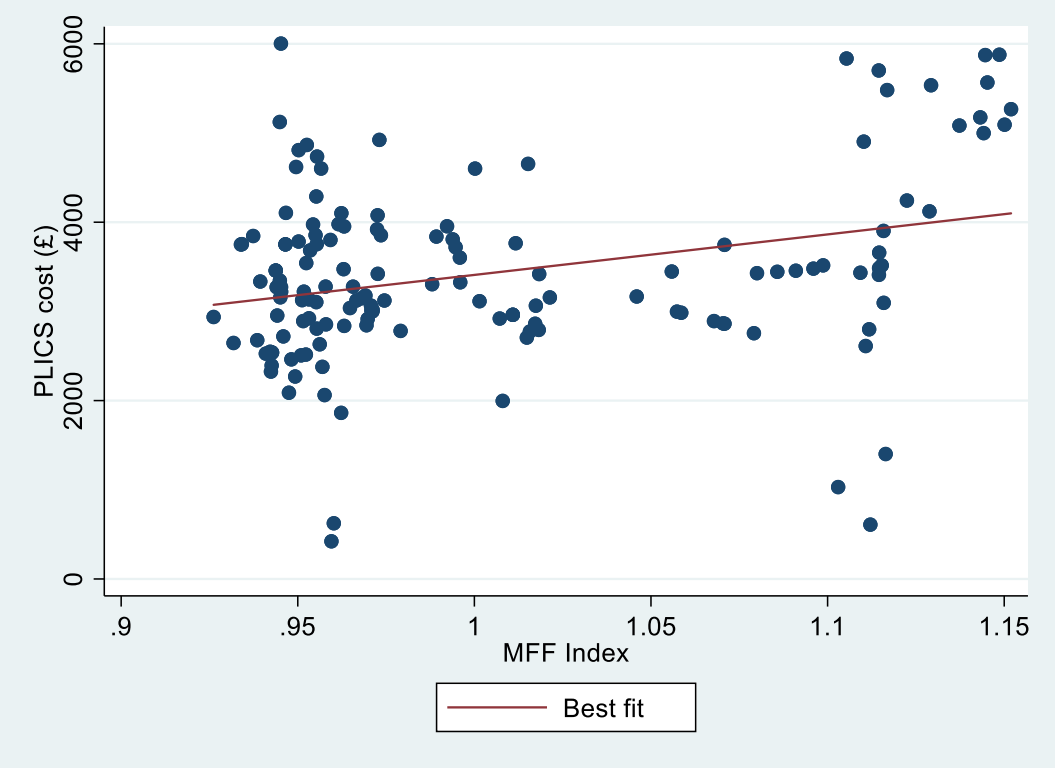


Figure 1f.1: Association between estimated PLICS costs and MFF.

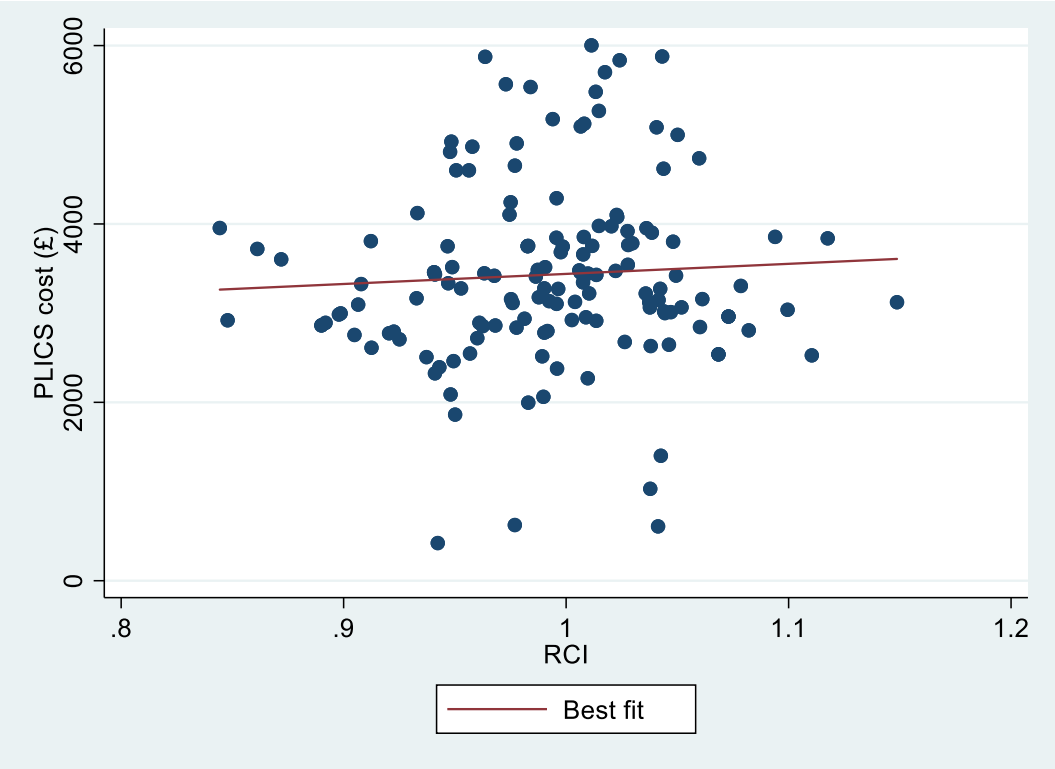


Figure 1f.2: Association between estimated PLICS costs and RCI

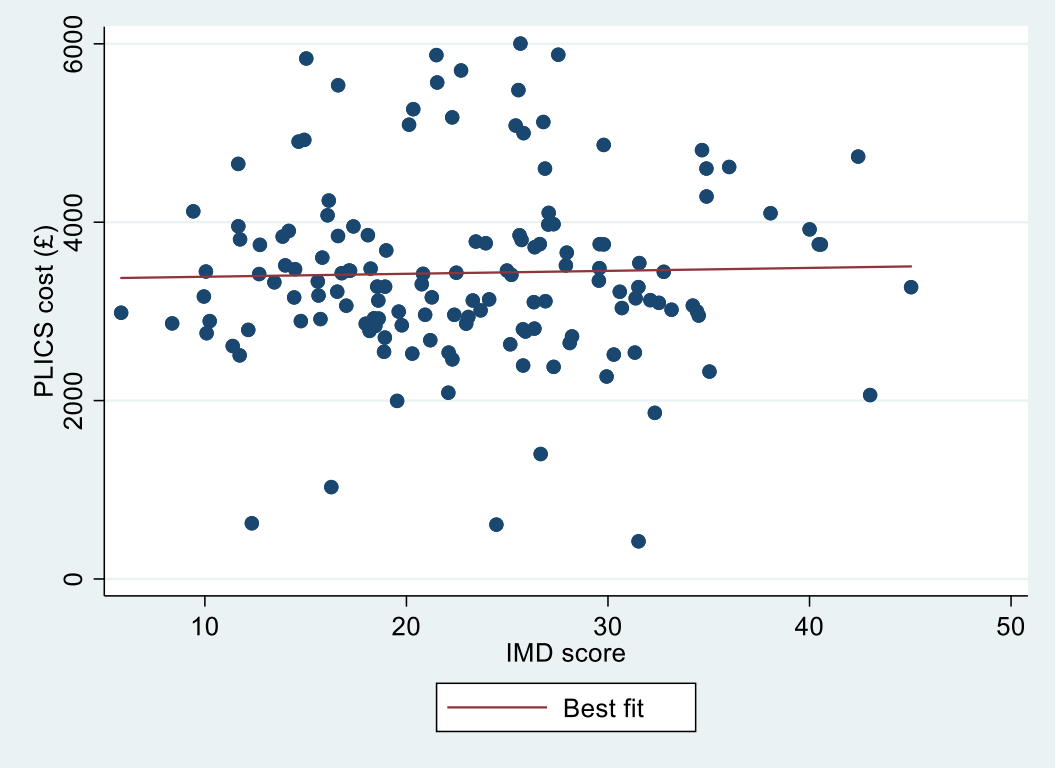


Figure 1f.3: Association between estimated PLICS costs and IMD score

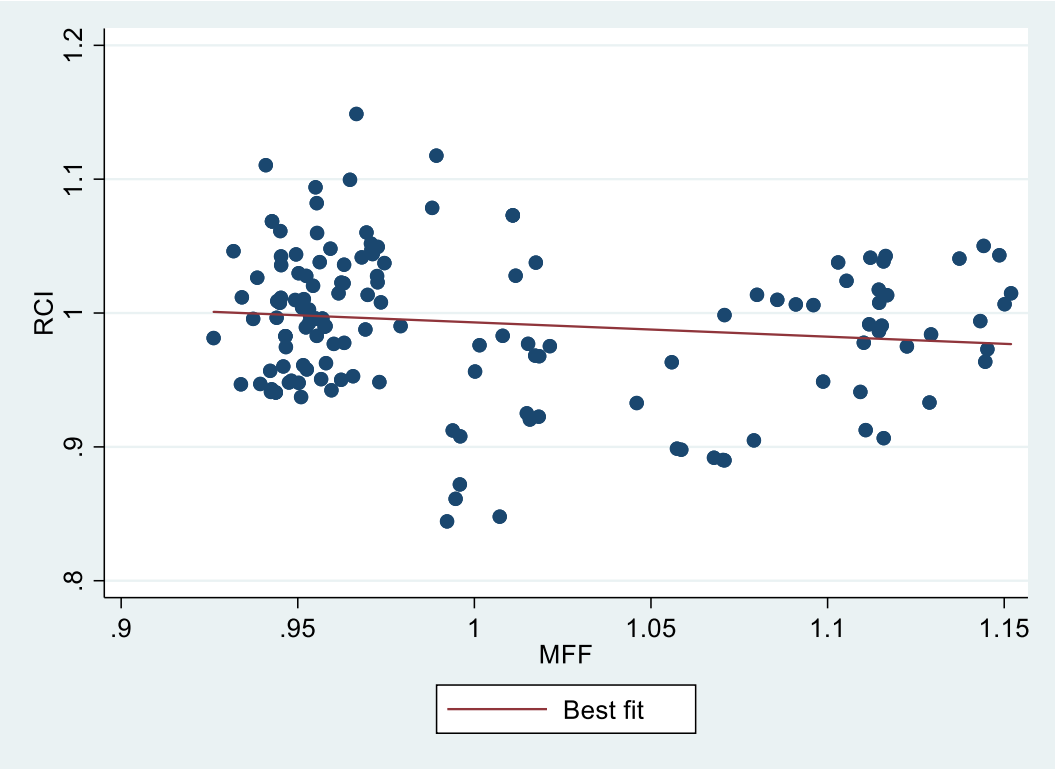
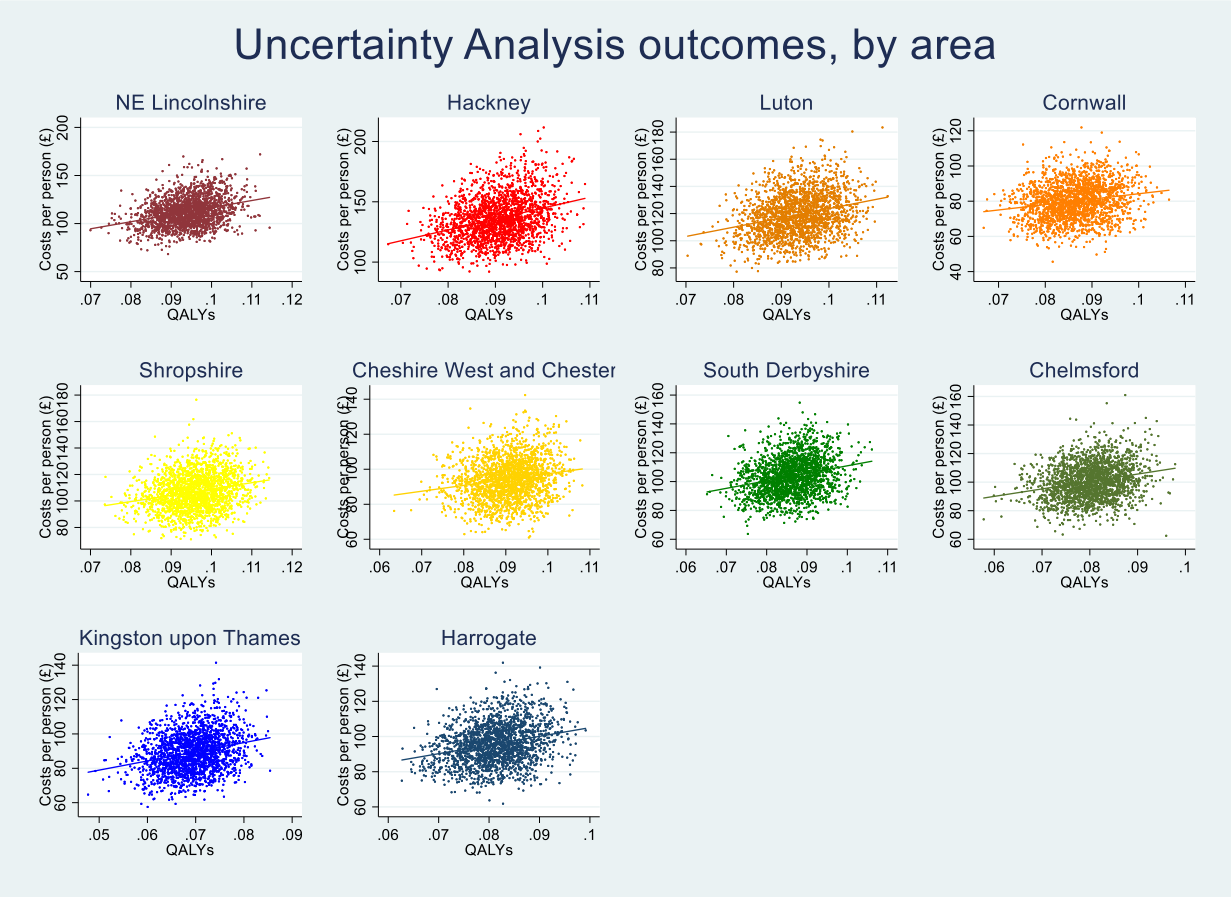
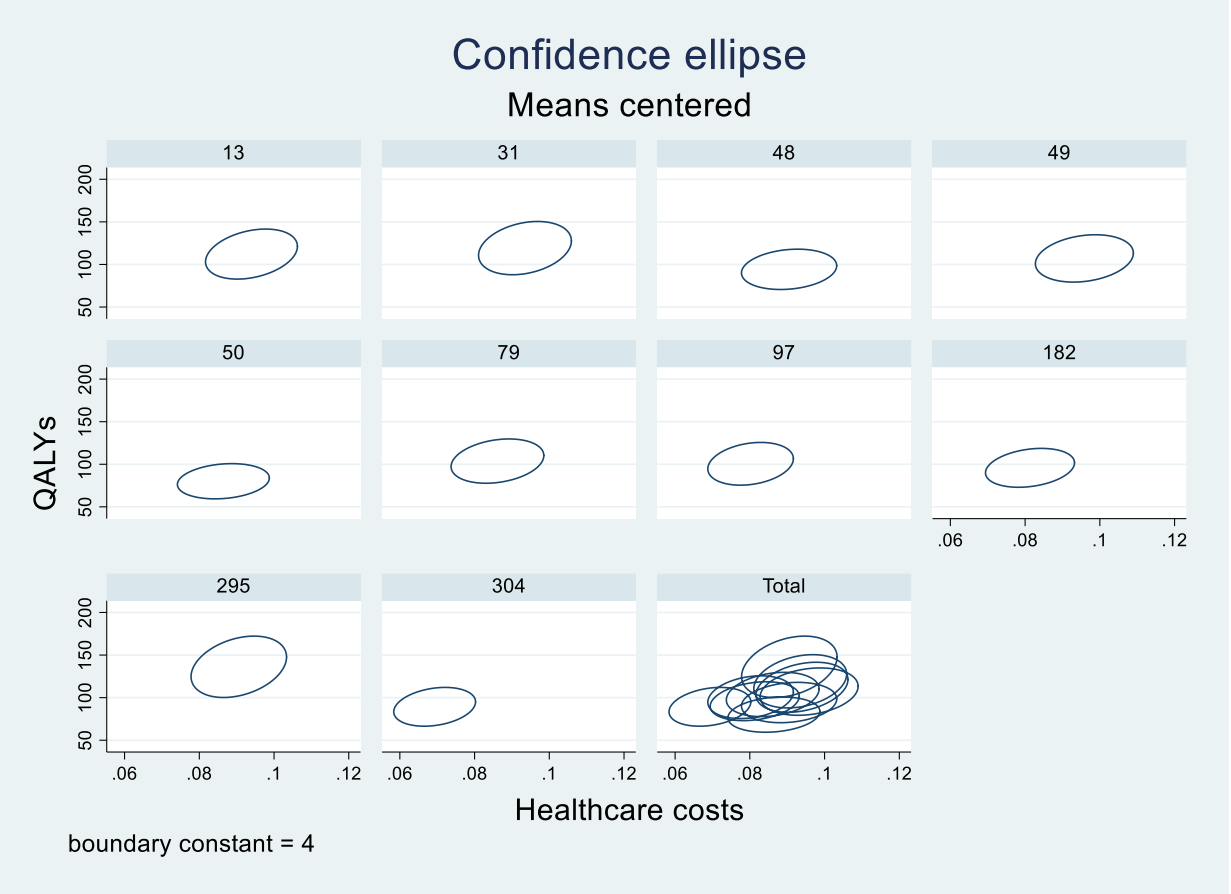


Figure 1f.4: Association between RCI and MFF.



Appendix figure 1g.1: Scatter plots of Probabilistic Sensitivity Analysis on QALY benefits against healthcare cost savings.



Appendix figure 1g.2: Confidence ellipses of Probabilistic Sensitivity Analysis on QALY benefits against healthcare cost savings for each of the ten areas. Area number 13 = North East Lincolnshire, 31 = Luton, 48 = Cheshire West and Chester, 49 = Shropshire, 50 = Cornwall, 79 = South Derbyshire, 97 = Chelmsford, 182 = Harrogate, 295 = Hackney, 304 = Kingston upon Thames.

Appendix 2: Further methods

Appendix 2a: Disease outcome inclusion framework

Background

The aim of a new local-level BMI model needs to have relevant disease burden represented so that as BMI is simulated to rise or fall, the modelled disease states appropriately capture the health implications of those changes. There are many potential ways to prioritise diseases to include in the model.

The aim was for a decision rule that accounted for the maximum modifiable burden of BMI-related disease. A pragmatic balance was required to account for as much burden as possible without adding unduly to model complexity and workload associated with programming and data collection.

Methods

The Global Burden of Disease study was chosen as the data source, as it offers the most comprehensive set of comparable disease burden estimates in England.(1) The GBD disease ("cause") taxonomy has a four-level hierarchical structure, with each level made up of mutually-exclusive and collectively exhaustive disease groups. Group A is for communicable diseases, maternal, perinatal and nutritional diseases, Group B is Non-Communicable Diseases and Group C is Injuries. The nutritional diseases in Group A are diseases of

nutritional deficiencies (such as iodine or vitamin A deficiency) that mainly affect children in the developing world and not relevant to the aims here around overnutrition in the UK.

Data were downloaded from the GBD interactive data tool (Global Health Data Exchange, GHDx).(97) To allow for multiple approaches of prioritising diseases for inclusion in the model to be compared, search criteria in the GBD search tool were defined as:

1. Deaths or Disability-Adjusted Life-Years (DALYs)
2. From each NCD (corresponding to group B in the GBD taxonomy)
3. With burden quantified as that attributed to BMI by the GBD
4. At the level of England
5. For both sexes combined
6. In the most recent available year (2017)

The list of NCDs that this returned then needed to be aggregated at appropriate level for each disease. Solid cancers were aggregated by site, but blood/ bone marrow cancers by type (leukaemias, multiple myeloma, Hodgkin and non-Hodgkin lymphoma) as these are dealt with separately in the literature linking them with risk factors.(274–276) Strokes were aggregated as a group (of ischaemic stroke, subarachnoid and intracerebral haemorrhages) because the literature allows these to be pooled.(264,272) The GBD aggregates the lower respiratory cancers (tracheal, bronchial and lung) together as 'lung cancer' as it is common to use lung cancer and bronchial cancer interchangeably, while tracheal cancer is extremely uncommon.(1) This left 38 disease groups for consideration. To account for the smaller populations susceptible to sex-specific cancers (in this case breast, uterine and ovarian

cancers as the GBD does not link any male-only cancers to the risk factors of interest), disease burdens were expressed as rates per 100,000 susceptible persons.

Inclusion of diseases was decided on the basis of DALYs to allow morbidity and mortality to both contribute to represented disease burden. PRIMEtime generally represents disease burden outputs as QALYs, so DALYS (which are very similar) make a better basis for prioritising the diseases to include than morbidity or mortality burden alone. Results by deaths are included here for comparison. Diseases were ranked in terms of which cause the greatest DALY burden per susceptible person.

Results

Figure 2a.1 shows how many diseases need to be included to make up a given percentage of disease burden related to BMI. Figure 2a.2 demonstrates the crude differences between the burden of disease. There is a large gap between DALYs and deaths, indicating that BMI is responsible for disability as well as deaths.

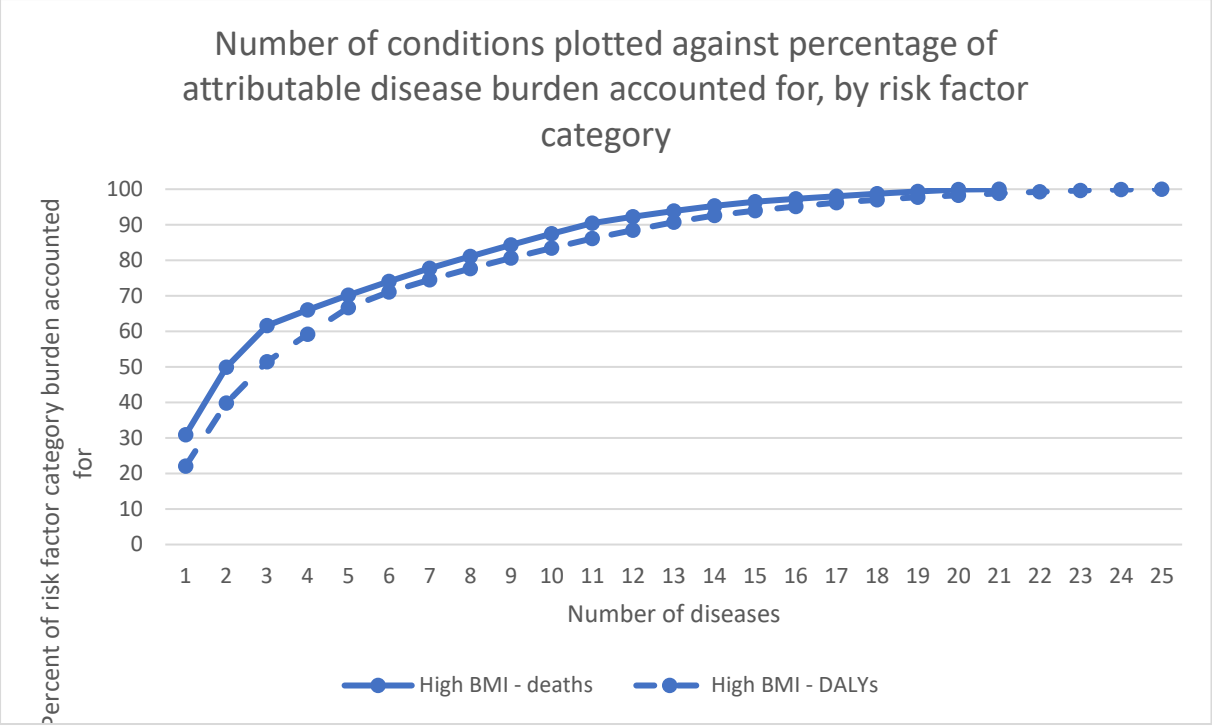


Figure 2a.1: Cumulative percentage of each risk factor group accounted for by including one more disease at a time, from greatest to smallest disease burdens.

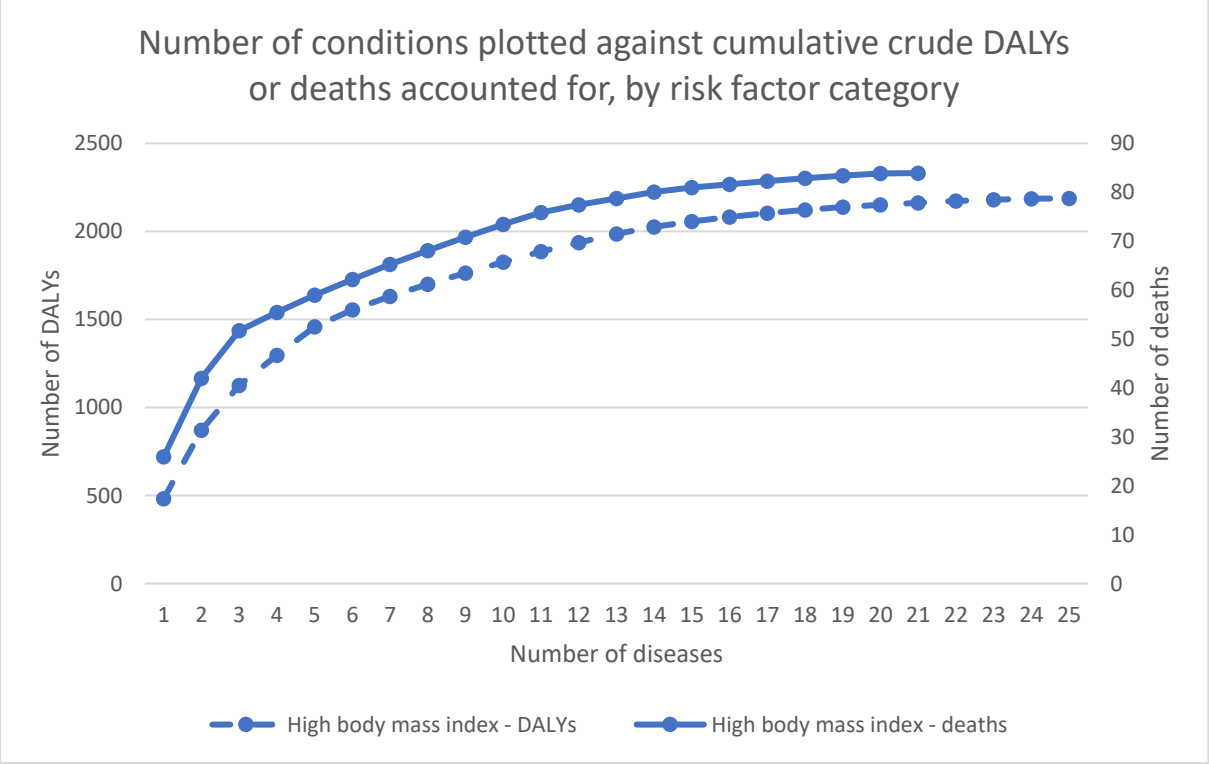


Figure 2a.2: Cumulative crude disease burdens of each risk factor group accounted for by including one more disease at a time, from greatest to smallest disease burdens, in terms of DALYs or deaths.

Table 2a.1 shows the 12 diseases with the greatest burden by risk factor group.

Appendix table 2a.1 of the 'top 10' biggest burden diseases for each risk factor category, listing biggest modifiable burdens separately by DALYs and deaths, with numbers of modifiable DALYs or deaths per 100,000 susceptible individuals in brackets and included diseases in italics.

Rank	Deaths attributable to high BMI	DALYs attributable to high BMI
1	*Ischaemic heart disease* *(25.94)*	*Ischaemic heart disease* *(482.45)*
2	*Dementias* *(15.99)*	*T2DM* *(388.92)*
3	*Stroke* *(9.77)*	*Stroke* *(254.61)*
4	Oesophageal cancer (3.74)	*Low back pain* *(170.03)*
5	Atrial fibrillation and flutter (3.51)	*Dementias* *(163.03)*
6	*Colorectal cancer* *(3.25)*	*Asthma* *(96.06)*
7	*Chronic kidney disease* *(3.06)*	*Chronic kidney disease* *(75.17)*
8	Hypertensive heart disease (2.82)	Oesophageal cancer (70.84)
9	*T2DM* *(2.72)*	Atrial fibrillation and flutter (64.41)
10	*Breast cancer* *(2.63)*	Osteoarthritis (60.98)
11	*Uterine cancer* *(2.45)*	*Colorectal cancer* *(60.2)*
12	*Kidney cancer* *(1.55)*	*Hypertensive heart disease* *(50.2)*

As discussed in Chapter 3, it was chosen not to include dementias as the disease burden is very different in type to the other diseases, and not to include CKD as most of its disease burden is as IHD. Breast cancer was included as interventions operating on PA may also find

this relevant as it is the only one of the five diseases in the GBD that is linked to PA not already included. It was later decided to separate 'osteoarthritis' into 'osteoarthritis of the hip' and 'osteoarthritis of the knee', due to a relative risk not being available for the aggregated disease state.

Conclusion

The final list of diseases to be included is: ischaemic heart disease, stroke, T2DM, low back pain, asthma, atrial fibrillation and flutter, osteoarthritis of the hip, osteoarthritis of the knee, hypertensive heart disease and oesophageal, breast and colorectal cancers.

Appendix 2b: Methods for the estimation of local populations aged 90 years and over

Background

The basis for new local modelling will be populations that reflect the basic demographics of those local areas. The PRIMEtime-CE model population is structured by age and sex demographics, so estimates of populations structured the same way will be necessary. The ONS publishes these from aged 0 to 89 years by local authority but the populations at local authority level aged 90 years and over are aggregated into a single grouping. These estimates are based on the Census, with their inter-Census mid-year estimates accounting for estimated birth, death and net internal and external immigration.(109) The cut-off is age 89 as it is felt that using the Census to make inter-Census population estimates is too inaccurate above this age due to high mortality and small numbers. The England-level population estimates extend to age 104 years by age and sex, but this relies on a different method called the Kannisto-Thatcher survivor ratio method, which, briefly, extending the logic that if a birth was registered at time *X* and no registration of death has been made for that individual by time *X+t*, then they are still alive. As this relies on the registrations of births and deaths done at the national level, this cannot be used to estimate local populations.(103) These estimates are consistent with national population estimates(103) but it under-estimates populations where mortality rates are falling as it relies on previous survival ratios of 'extinct cohorts'.(142) The ONS also rounds these estimates to the nearest

10.(142) As no official estimates exist of local populations above the age of 89 years, a method was required to estimate these. The aim was to estimate the distribution of the populations of females and males above the age of 89. Uncertainty will not be estimated and existing uncertainty estimated around ONS figures will not be included as this cancels out in the modelling process so would not affect the modelled estimates.

Methods

The method used two steps, undertaken separately for females and males. These used the 2018 ONS official population estimates for local areas of England.(109) First, the proportion of the national population of over-89s living in each local area was calculated by dividing the local population of over-89s by the national population of over-89s, using the ONS estimates of these figures. This proportion for each local authority was then used to apportion national population estimates to local areas. The proportion for each area was multiplied by the national estimate for each age from 90-99 years, in turn. The proportion was also multiplied by the ONS estimates of the national population of over-99s nationally to produce local estimates of over-99s. The City of London and Isles of Scilly were excluded due to small numbers (in line with many other local sets of statistics including the GBD work on local areas in England).(2)

Validation of the estimates was then performed by comparing the number of 90 year-olds estimated in the method with the trend line of population between ages 80 and 89 years.

Results

Estimates were produced for local authority populations by year and sex for those aged 90-99 years and by sex for those 100 years and over.

Worked example: Oxford is home to 0.002328 of over-89 year-old females in England (0.2328%). There are 69,330 90 year-old females in England, so; Oxford is estimated to be home to 69,330 x 0.002328 = 161.4 90 year-old females, or rounded to the nearest 1, = 161 individuals.

For example, in figure 2b.1a and 2b.1b, the ONS population estimates of population aged 80-89 can be seen, followed by the new estimates aged 90 years and over.

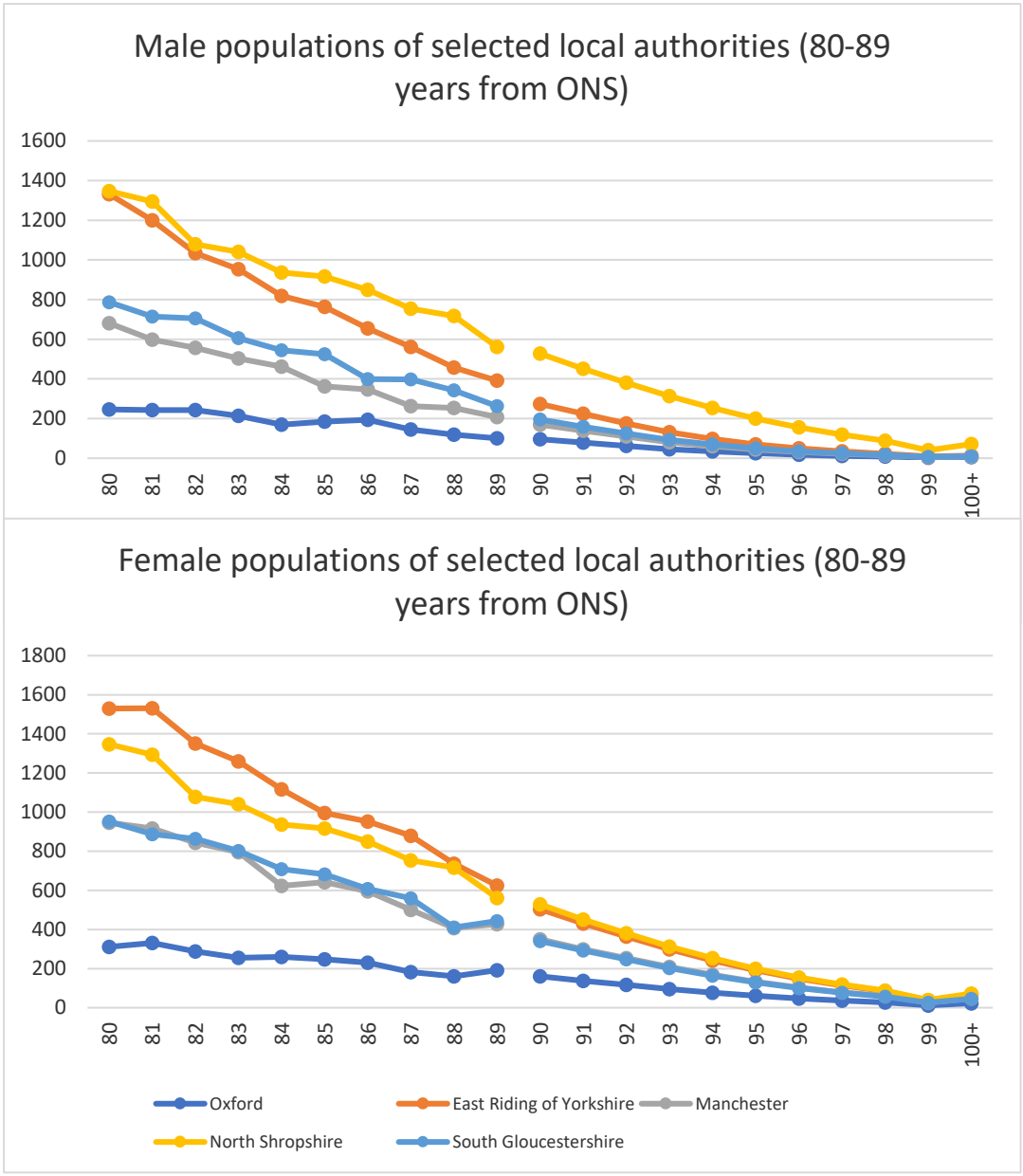


Figure 2b.1a (above) and 2b.1b (below) for females and males, respectively: the populations of selected local authority areas, with ONS estimates up to age 89 and new estimates from 90-100+.

Validation

There is no external dataset to validate these estimates against, but the ONS estimates up to age 89 years were not directly used to make the estimates, so provide a helpful comparison. A linear regression model for each area was used to generate predicted values

of population size at age 90 years, based on the trend between populations aged 80 and 89 years. The predicted value of 90 year-olds from the regression model was compared against the population estimate of 90 year olds.

For each sex and local area, population size was regressed on age between ages 80 to 89 years. This was compared against a model regressing population size against age and age squared. Likelihood ratio tests were used to test model superiority and in a small minority of cases did the test support the inclusion of the quadratic term, so the univariate model was used for all areas. The proportional difference was calculated between the main estimate of the population of 90 year-olds with the validation estimate (so that: $Proportion = \frac{P-R}{R}$ where *P* is the estimate of 90 year-olds and *R* is the predicted value). Histograms of these differences are shown in Appendix figure 2b.2a and 2b.2b. For females, the mean difference in estimates was 0.036 and SD 0.099, the mean proportional difference for males was -0.160 and SD 0.24.

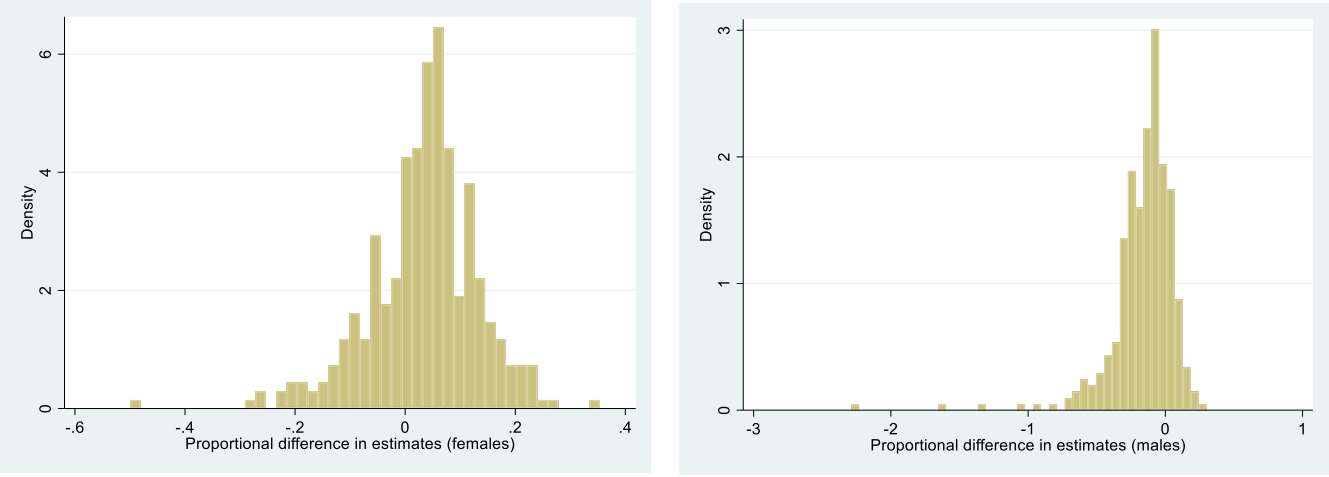


Figure 2b.2a (left) and 2b.2b (right) for females and males, respectively: The distribution of differences between new estimate of the population of 90 year olds and estimate based on linear trend, for females.

The top five biggest breaks from prior trend at age 90 years for each sex at either end of these distributions have been plotted for ages 80-89 years (from ONS estimates) and 90-100+ years (new estimates) in figures 3 and 4. In both the cases of increases and decreases relative to trend at age 90, even the most extreme examples do not appear greatly out of proportion to the volatility in the formal ONS estimates.

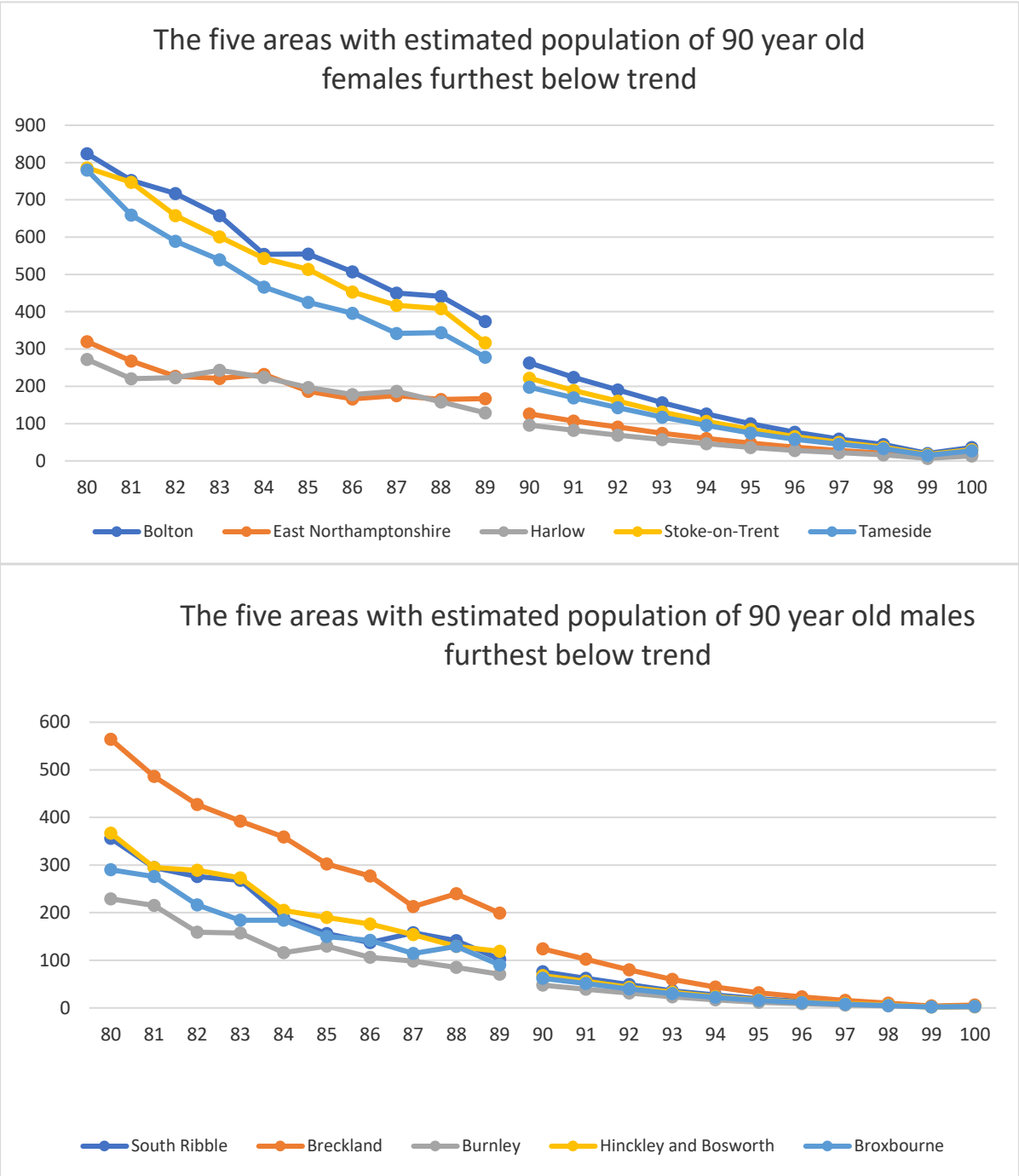
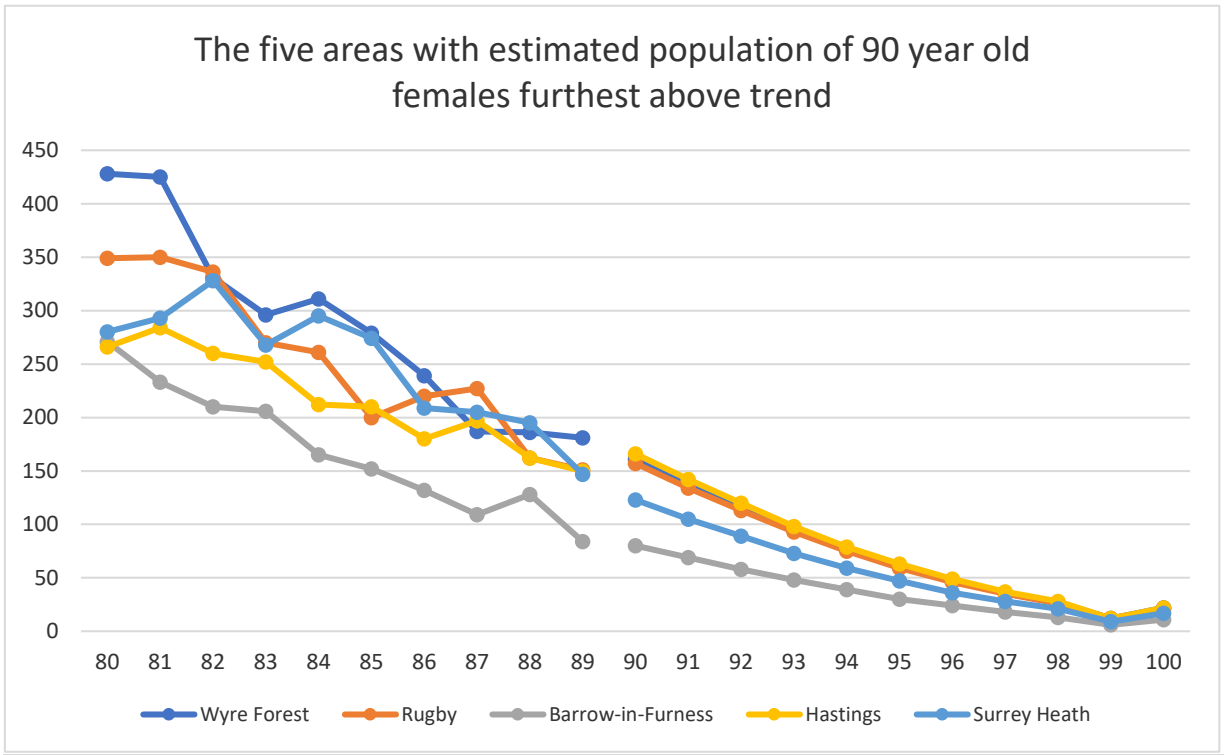


Figure 2b.3a (above) and 2b3b (below) for females and males, respectively: Age plotted against population for the five areas with estimated population of 90 year-olds furthest below the trend between ages 80-89 years.



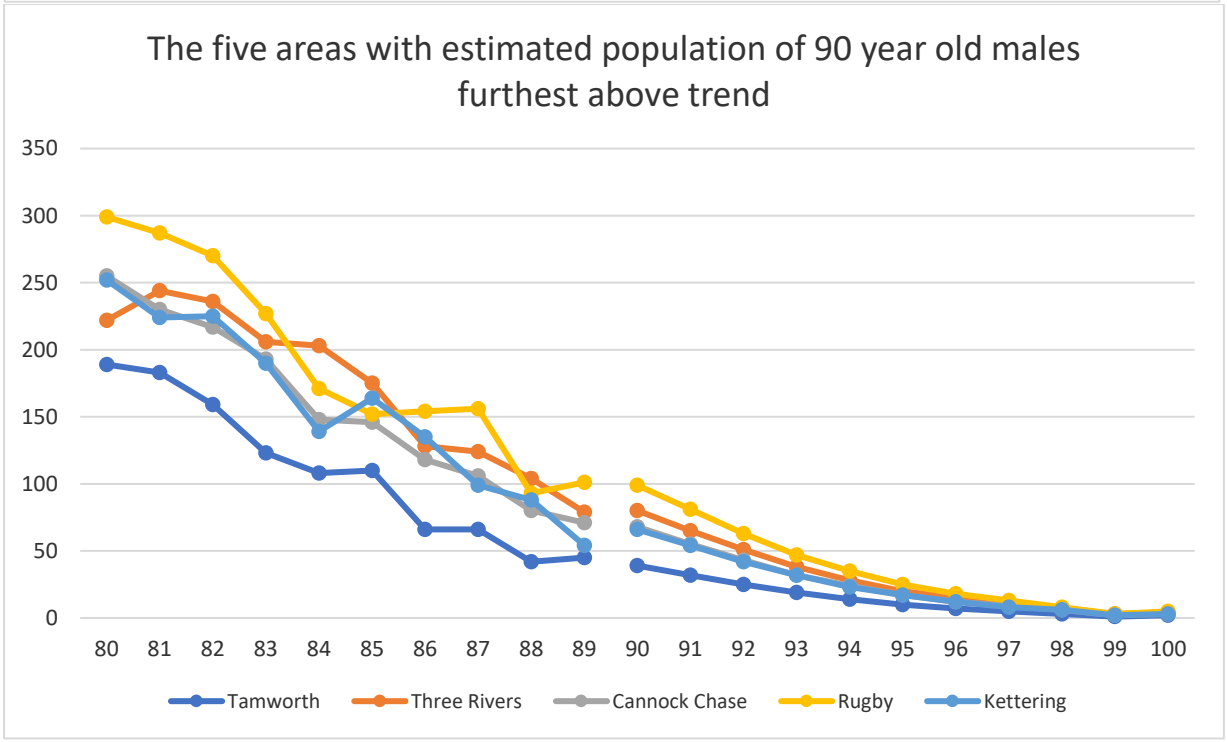


Figure 2b.4a (above) and 2b.4b (below) for females and males, respectively: Age plotted against population for the five areas with estimated population of 90 year-olds furthest above the trend between ages 80-89 years.

Discussion

Estimated populations had smoothly monotonically declining distributions. Each population was estimated separately for ages 90-99 years and a '100 years and over' group.

Validation compared population estimates either side of the breakpoint at age 90 years. Comparing the new estimate of the population of 90 year-olds with an approximation for that figure based on numbers of 87-89 year-olds allowed direct comparison between the year-by-year ONS estimates and the new estimates. This showed even the largest over- and under- estimates are not clearly greater than the natural volatility in populations year-on-year in the official estimates.

There are a few weaknesses of this approach that are worth mentioning. The ONS only publishes uncertainty on local authority populations at the total population level and not by age and sex, so this cannot be incorporated into cumulative modelled uncertainty. These populations cannot be externally validated; this estimation exercise is needed as external data to validate against do not exist. Finally, it does not account for net internal migration between local authorities of different ages in the over-90s group. The ONS does not publish migration figures by age, but the contribution of net migration to the populations of over-89s is likely to be small.

Appendix 2c: Assessment of modelling best practices

Introduction

As mentioned in Chapter 9, it is usually not possible to validate the modelled outcomes of scenario modelling on NCD interventions against empirical evidence. Effects are small at the individual level, thinly distributed, take years or decades to translate from the intervention to the risk factor and from the risk factor to the disease outcome. The population variation and cumulative impact of other interventions may also be much greater than the effects of the intervention, making it difficult to isolate the effects of an intervention on outcomes.

For these reasons, different approaches to validating health economic models and scenario models have been proposed, based on scrutinising the approach to modelling, appropriateness for the phenomenon being modelled and the quality of the data inputs, before asking if any of the outputs can be validated against any comparable studies.

Methods

This validation approach follows that used for the literature review in Chapter 2. This used an adapted version of the ISPOR-SMDM approach to evaluate the models identified in the literature. An adapted version of the full list of best practice recommendations from the ISPOR-SMDM Taskforce is proposed here on the basis that the full protocol contains many items that are intended for individual-level health economic evaluation(80) and do not apply

to population-level public health interventions and the PRIMEtime model. Instead, the core recommendations were taken and merged into fewer items, capturing the most important issues across the domains of validation: 1. conceptualising the model, 2. state transition models, 3. parameter estimation and uncertainty, and 4. transparency and validation. For each of these, the expectation is to explain if certain considerations have been taken when building and explaining the model, so the aim here is to briefly explain whether or not each criterion has been met, in what ways, and explain where it has been met in full detail elsewhere in the thesis.

Results

Conceptualising the model

Stakeholder engagement

The PRIMEtime local model had two streams of stakeholder engagement, one with the public and patients, the other with expert professional stakeholders. This supported the development of the model and provided input into what scenarios were of interest. Previous rounds of stakeholder engagement have also been undertaken and included an assessment of face validity, as detailed in Chapter 3.

Statement of the decision problem, including perspective, objective and scope of modelling

The model is designed to estimate local authority level variation in the effectiveness and cost-effectiveness of body weight interventions. Specifically, in Chapter 8 this is stated as "the effectiveness, cost-effectiveness and equity implications of restrictions on television advertising restrictions is modelled at the local level on disease burden (cases and QALYs) and healthcare costs".

Explanation for the choice of modelling structure

This is detailed under "Strengths and limitations of the PRIMEtime structure" in Chapter 9, in summary that it is ideally designed to evaluate approaches to NCD prevention due to its ability to capture the impact of a risk factor (or multiple risk factors) on multiple disease outcomes, over a lifetime, accumulating disease burden and cost implications. It is designed to allow the exploration of scenarios through varying multiple assumptions.

States represented can capture the impact of the intervention

The model is explicitly designed so that the model disease states capture the impact of the intervention. As explained under "Strengths and limitations of the PRIMEtime structure" in Chapter 9, the model "does not privilege one particular disease state, instead intending to capture as much disease burden as feasible". Disease states are chosen from those that contribute most to BMI-attributable burden, as detailed in Appendix 2a.

Is the right balance of simplicity and complexity made?

One of the strengths of PRIMEtime is the elegance and transparency of the structure and its parameters. The correct balance made between simplicity and complexity is clearly a complex judgement, but overall there has been effort made to include no complexity beyond the necessary relationships between inputs, important disease states and outcome measures. Levels of complexity that are implemented in other versions of PRIMEtime but that were not possible to be included at the local level have been omitted, such as social care costs or BMI trends. These issues relate to data limitations at the local level.

The estimated parameters are calculated transparently and effort is being made to peer review these, including Chapter 5 which has been published in a high-quality peer-reviewed journal.(51)

State transition models

Is a cohort simulation used where number of states is manageable?

Yes. PRIMEtime is explicitly designed to do so.

Is the model structure consistent with decision problem?

The model is intended to reflect the exposure of populations to risk, the development of disease then mortality. This process is well represented by cohorts of ageing population, parameterised by age- sex- and locally-specific population, epidemiology, BMI and costs. Cycle length is 1 year, which is standard, appropriate to the process and consistent with the availability of good-quality data. Relevant disease endpoints chosen as described in

Appendix 2a: disease inclusion framework are based on BMI-attributable disease burden. The 12 diseases account for over 90% of BMI-related morbidity and mortality. Disease epidemiology is calculated from GBD data aggregated to the IMD level. The time horizon is up to 100 years.

Calorie change impacts health via BMI but not via other associated physiological intermediates (lipids, BMI or HbA1c more granularly than T2DM cases). BMI distributions are parameterised as continuous then used to calculate proportions in each BMI risk group. Then using cohorts, RRs are applied to those numbers to generate baseline and scenario numbers of cases, QALYs and costs. Correction factors applied for T2DM-IHD and T2DM-stroke to account for the association between these states without double counting.

Clearly communicated

The model is extensively described through this thesis, in addition to the other work describing the structure and function of PRIMEtime. The aim is for individual components to be peer reviewed and published separately, of writing Chapter 5 estimating local adult BMI distributions.

Transition probabilities calculated transparently

Relative risks taken from gold standard Meta-Analyses of Prospective Cohort Studies and Case-Control Studies. Epidemiology is taken from the Global Burden of Disease study 2019:(96,97) all-cause mortality rates are taken directly and secondary modelling is used to estimate consistent sets of incidence and case fatality rates, explained in detail in Chapter 4.

Parameter estimation and uncertainty

Parameters estimated on the basis of unbiased evidence, with formal synthesis

Populations were taken and adapted from Office for National Statistics Appendix 2b, a well-known and trusted source. Utility values are taken from Sullivan *et al*,(98) the best quality utility values available for disease states applying to the UK. Local-specific adult and child BMI, disease costs and intervention parameters were also estimated, laid out in detail in Chapters 4, 5, 6 and 7, respectively. The adult BMI estimation chapter has been published in a peer-reviewed journal(277) and the aim is to submit remaining parameter estimates for peer review in future. The most up to date data sources were used throughout, with the exception of children's BMI data from the NCMP, due to acute changes in the COVID-19 pandemic, as explained in Chapter 6. Although at the time the Census 2011 was the most up-to-date data source for estimating adult BMI and the 2021 microdata is not yet available, there is a risk of the 2011 data being outdated. It is unfortunately never possible to exclude any risk of bias, even in good-quality surveys, as issues such as response bias may make data more representative of healthier populations.

Clear distinction between uncertainty analysis and sensitivity analyses

Yes, this distinction is clearly made in Chapter 8.

Appropriate deterministic and probabilistic uncertainty analysis included, with appropriate distributions, communicated in standard statistical methods

Deterministic sensitivity analyses were performed in Chapter 8 on the effects of changing discount rates, on the size/variation in the effect of ad exposure-calorie change, and the effect of the regulations resulting in adverts being delayed rather than cancelled. PSA was performed on 10 areas with uncertainty incorporated for Relative Risks, utilities, healthcare costs and the intervention effect.

Transparency and validation

Technical and non-technical documentation available

Extensive technical explanations are presented through this thesis on the type of model, its intended applications, funding, structure, inputs, outputs, and their relationships to model structure, data sources, validation methods and limitations. Non-technical documentation is not provided.

Process for assessing face validity

Previous work has examined face validity(72,88) with expert stakeholders, with 9 of 12 respondents supporting face validity, as explained in Chapter 3.

Structural verification

Throughout model building, checking was performed on example lines to check that data from the risk factor module, disease models and lifetable were connecting accurately.

Comparison with similar modelling analyses

Comparison of modelled outputs against Mytton *et al*(203) is provided. Use of the same discount rates results in comparable aggregate QALY effects being estimated.

External validation

Modelled outputs at the local level cannot easily be validated, but input parameters were validated externally where possible. Adult BMI distributions were validated against NDNS and ALS data, with close concordance. Local cost estimates and intervention effect cannot be externally validated. Previous work(47) has externally validated results against two other sources. This set PRIMEtime inputs to 2005 and estimated disease rates 10 years later, then compared outcomes to estimates from the Global Burden of Disease 2015 and a paper by Kuan *et al*. estimating disease rates in 2015. This work found PRIMEtime estimates falling between the two external sets of estimates (higher than the GBD but lower than Kuan *et al*). This provides good assurance that the outputs are comparable to other high-quality work.

Discussion

Assessment against comparable models

The extent that PRIMEtime local met best practice recommendations set out can be compared with the models identified in the literature review in Chapter 2:

- The PHE CVD Prevention model 2018
- The SPHR Diabetes Prevention model 2015
- The NICE Physical Activity model 2014
- The PHE Weight Management model 2016

The former two of these (both developed by the team at SCHARR, University of Sheffield) were more successful at meeting the set best practice recommendations than the latter three.

Conceptualising the model:

The PRIMEtime local model had less stakeholder engagement for its development than the two Sheffield models, but the remaining two did not have any stakeholder engagement. The model structure is justified, along with the Sheffield models. In terms of simplicity vs complexity and the appropriateness of states modelled, there was a risk that the PHE Wight Management model did not represent any BMI heterogeneity, only modelling on the group average BMI. In common with the PRIMEtime local model, the four reviewed models each had a well-defined decision problem.

State transition models:

PRIMEtime local met comparable recommendations on state transition models to the other four models. It uses a cohort model, meeting the criterion to do so where feasible along

with the remaining four local models (it was not feasible for the two Sheffield models). This thesis is predictably more detailed than the remaining technical documents, though it was strongest for the Sheffield models and weakest for the PHE Weight Management model. The model structures were all satisfactorily consistent with the decision problems, half cycle corrections were not implemented and state transition parameters were transparently calculated and of good quality.

Parameter estimation and uncertainty

Deterministic sensitivity analyses were performed for the PRIMEtime local model and Diabetes Prevention model, with suggestions of deterministic sensitivity analyses for users to implement for the PHE Weight Management model (though their impacts on results are not presented). PSA is performed for PRIMEtime local and the Diabetes Prevention model, while the NICE PA tool documentation did not meet the expectation to distinguish uncertainty and deterministic sensitivity analysis clearly. The included parameters for each model were calculated transparently and appropriately.

Transparency and validation

Detailed technical documents are available for the PRIMEtime local model and Sheffield models, but less documentation is available for the remaining models. The PRIMEtime_local model had no process for assessing face validity, though previous work on PRIMEtime has done this well.(72,88) Of the other local models the two Sheffield models were the only two to do so. Manual checking of calculations was done for the PRIMEtime local model, Sheffield

models and PHE Weight Management model. No directly comparable local modelling analyses were available for PRIMEtime local outputs, but one was compared at the England level. The Sheffield models all presented some comparison against similar models and external data sources, the PHE Weight Management model had good external comparisons, but the remaining two models did not present any.

Synthesis

Overall, the PRIMEtime local model was successful at meeting the modified set of ISPR-SMDM modelling best practice recommendations – with slightly lower success than the two Sheffield models (the PHE Diabetes Prevention model and the SPCR CVD Prevention model) but greater success than the remaining two (NICE PA model 2014, and the PHE Weight Management model 2016). This is a good level of success, considering that the two Sheffield models were commissioned to teams of researchers with substantial resources. Indeed, none of the reviewed models were published primarily as an academic work.

There are a wide variety of best practice and validation frameworks available, with the ISPOR-SMDM being chosen for this assessment on the basis that it is generally considered the gold standard, with a wider variety of recommendations than others and explicitly encompassing a frame of 'best practices' rather than validation alone.

Direct comparison of model best practices is complex. Its assessment involves a two-way relationship between a technical understanding of what has been done, knowledge of datasets used and methods for their synthesis as input parameters, with nuance around a model's relationship to its decision problem, simplifying assumptions in how much

complexity of the phenomenon should be captured and a reasonable judgement on what is necessary and appropriate to model given finite resources. Models are clearly never entirely "valid" in that "all models are wrong", while it is also likely very few are entirely *in*valid.

Conclusion

The PRIMEtime local model meets high standards of modelling best practices across the domains of conceptualising the model, state transition modelling standards, parameter estimation and uncertainty, and transparency and validation, as judged using a modified set of ISPOR-SMDM recommendations and compared against the four local authority models identified in the literature review in Chapter 2.

Appendix 3: Manuscripts arising from this thesis

- Amies-Cull B, Scarborough P, Wolstenholme J, Reference costs: a history and critique, Health Economic Review (*in review*)
- Amies-Cull B, Wolstenholme J, Cobiac L, Scarborough P, Estimating BMI distributions by age and sex for local authorities in England, BMJ Open, doi.org/10.1136/bmjopen-2022-060892