Estimating and comparing the cost-effectiveness of primary prevention policies affecting diet and physical activity in England

Adam Briggs

Brasenose College

A thesis submitted for the degree of Doctor of Philosophy

Nuffield Department of Population Health

Medical Sciences Division

Trinity 2017

Abstract

Estimating and comparing the cost-effectiveness of primary prevention policies affecting diet and physical activity in England

Adam Briggs, Brasenose College
For the degree of Doctor of Philosophy, Nuffield Department of Population Health
Trinity 2017

Health and public health services in England are under increasing financial pressure. At the same time, nearly 40% of the total disease burden is potentially amenable to known causes with two of the leading behavioural risk factors being unhealthy diets and physical inactivity. To better inform decision makers and improve health in England, this thesis aims to develop a cost-effectiveness model that can directly compare diet and physical activity interventions.

Published public health economic models were reviewed and the strengths and weaknesses of the modelling structures were explored. A pre-existing multistate life table model, PRIMEtime, was developed into PRIMEtime Cost Effectiveness (PRIMEtime CE). Disease specific NHS England costs were derived from NHS England Programme Budgeting Data and unrelated disease costs from NHS cost curves. Social care costs were quantified using a Department of Health tool for estimating wider societal costs. Disease specific utility decrements were adopted from a catalogue of EuroQoL five dimensions questionnaire scores. The cost effectiveness of reformulating food to have less salt and of expanding access to leisure centres in England were modelled from an NHS and social care perspective over a 10 year time horizon, including government and industry costs.

Salt reformulation was dominant with an estimated cost per quality adjusted life year (QALY) of -£17,000 (95% uncertainty interval, -£40,000 to £39,000), compared with £727,000 (£514,000 to £1,064,000) for increasing access to leisure centres. Sensitivity analyses and cross validation testing of outcomes demonstrated how cost per QALY estimates were sensitive to the choice of model scope, parameters, and structure.

PRIMEtime CE is a tool for decision makers to compare interventions affecting diet and physical activity, enabling them to make better informed choices about how to spend finite resources. Future work will focus on making the model freely available and expanding its risk factors to enable comparisons of other public health interventions.

Acknowledgements

I would like to thank my supervisors, Pete Scarborough and Jane Wolstenholme, for their advice and guidance throughout my DPhil, and for allowing me the autonomy to pursue my academic interests as I wished. I am privileged to have been able to work with the Centre on Population Approaches for Non-Communicable Disease Prevention both during my DPhil and before. It is a hugely positive and encouraging environment within which to work. I will miss it and everybody who works there.

I am very grateful to my funders, The Wellcome Trust and the Nuffield Department of Population Health. I would also like to acknowledge the project stakeholders who have been integral to the design and development of my DPhil, providing feedback on the model's progress and its results.

I would like to thank my parents for normalising science and academia, and for indulging and enabling my multiple years of full time higher education. Finally, I am hugely grateful to my wife, Sarah, and daughters, Iris and Agnes, for their emotional and intellectual support and distractions throughout.

Table of contents

List of tables

List of figures

List of abbreviations

ABS, Agent based simulation

ACE Prevention, Assessing Cost Effectiveness in Prevention

APS, Active People Survey

BENESCO model, Benefits of Smoking Cessation on Outcomes model

BMI, Body mass index

CASH/AoS, Consensus Action on Salt and Health/Action on Sugar

CASP, Critical appraisal skills programme

CBA, Cost-benefit analysis

CCA, Cost-consequence analysis

CCC, Clinical classification category

CCG, Clinical commissioning group

CHD, Coronary heart disease

CHEERS, Consolidated Health Economic Evaluation Reporting Standards

CHOICES, Childhood Obesity Intervention Cost-Effectiveness Study

CISNET, Cancer Intervention and Surveillance Modeling Network

CEA, Cost-effectiveness analysis

CMA, Cost-minimisation analysis

COPD, Chronic obstructive pulmonary disease

CRA, Comparative risk assessment

CUA, Cost-utility analysis

CVD, Cardiovascular disease

CVPPIP, Cardiovascular Patient and Public Involvement Panel

DES, Discrete event simulation

DFID, UK Department for International Development

ERS, Exercise referral scheme

EQ-5D questionnaire, EuroQol five dimensions questionnaire

FSA, Food Standards Agency

GI, Gastrointestinal

GP, General practitioner

GPRD, General Practice Research Database

HALY, Health adjusted life year

HCHS index, Hospital and community health services index

HES, Hospital episode statistics

HR, Hazard ratio

HRQoL, Health related quality of life

HSE, Health Survey for England

HTA, Health technology assessment

ICD-10, International Statistical Classification of Diseases and Related Health Problems, 10th revision

ICER, Incremental cost effectiveness ratio

ISBNPA, International Society of Behavioural Nutrition and Physical Activity

ISPAH, International Society for Physical Activity and Health

ISPOR-SMDM, International Society for Pharmacoeconomics and Outcomes Research-Society for Medical Decision Making

RIVM CDM, Rijksinstituut voor Volksgezondheid en Milieu Chronic Disease Model

MCDA, Multi-criteria decision analysis

MET, Metabolic equivalent of task

MI, Myocardial infarction

MOVE model, Measurement of the Value of Exercise model

NCD, Non-communicable disease

NDNS, National Diet and Nutrition Survey

NHS, National Health Service

NICE, National Institute for Health and Care Excellence

OECD CDP model, Organisation for Economic Co-operation and Development Chronic Disease Prevention model

PBC, Programme budgeting category

PCT, Primary care trust

PRIME, Preventable Risk Integrated ModEl

PRIMEtime CE, PRIMEtime Cost Effectiveness

PSSRU, Personal Social Services Research Unit

QALY, Quality adjusted life year

SACN, Scientific Advisory Committee on Nutrition

SES, Socioeconomic status

SG, Standard gamble

SROI, Social return on investment

TTO, Time trade off

UI, Uncertainty interval

US CDC, United States Centre for Disease Control

VAS, Visual analogue scale

WHO, World Health Organization

WHO-CHOICE, World Health Organization CHOosing Interventions that are Cost-Effective project

Publications and presentations arising from this thesis

An adapted version of the background and review of model structures (sections 1.1 and 1.3) has been published in the journal, Population Health Metrics in May 2016:

- Briggs ADM, Wolstenholme J, Blakely T, Scarborough P. Choosing an epidemiological model structure for the economic evaluation of non-communicable disease public health interventions. Population Health Metrics. 2016;14(1):17.

Three further research articles arising from this thesis are planned for publication within the next 12 months:

- Estimating comparable disease costs from NHS England programme budgeting data for use in health economic modelling. Target journals: Value in Health, Medical Decision Making, Health Economics
- The cost-effectiveness of public health policies affecting diet and physical activity using PRIMEtime CE, a multistate life table cost effectiveness model. Target journals: Population Health Metrics, BMC Public Health, PLoS One, BMJ Open
- The challenges of validating public health economic models. Target journals: Journal of Public Health, BMC Public Health, Population Health Metrics

The methodology for estimating costs and utilities (chapter three) was presented at the Health Economists' Study Group meeting in January 2017. Results of this thesis were presented to public health professionals at the 2017 Faculty of Public Health annual conference, and to health economists from the Department of Health and Public Health England in June 2017.

Chapter 1. Introduction

1.1 Background

This thesis aims to improve health and reduce the burden of non-communicable disease (NCD) in England by developing a cost-effectiveness model that can compare diet and physical activity interventions, and help public health decision making. Non-communicable diseases are responsible for 88% of the total disease burden in England, 36% of which is attributable to potentially amenable behavioural, environmental, and metabolic risk factors. This figure varies by disease with as much as 80% of the cardiovascular disease (CVD) burden attributable to known risk factors. The four leading behavioural risk factors for disease in England are tobacco, unhealthy diets, alcohol misuse, and physical inactivity. Of these, poor diet and physical inactivity account for a third of the total attributable disease burden,[1] a burden that could be significantly reduced through public health interventions.[2–5]

Over recent years, there have been increasing pressures on acute health services in England and as a consequence, health providers are arguing for there to be a greater emphasis placed on prevention. In 2014, The National Health Service in England (NHS England) announced in its Five Year Forward View *"that the future health of millions of children, the sustainability of the NHS, and the economic prosperity of Britain all now depend on a radical upgrade in prevention and public health"*.[6] In order for prevention and public health to play its part in maintaining the sustainability of the NHS, public health practitioners and decision makers need to have the information, influence, and resources to make the best decisions about how to spend finite resources. In 2013, the structure of public health in England changed with public health responsibility moving from the NHS to local government.[7] This potentially enables public health professionals to influence the wider determinants of health by working alongside council departments such as social services, housing, and transport. However, at the same time it places local public health departments under the financial control of local governments, which have been enduring

substantial budget cuts. Over the course of the 2010-2015 parliament, core local government funding was reduced by 40% and the Local Government Association estimates a £6 billion funding gap in 2016/17 rising to £10.8 billion by 2019/20.[8] In 2016, the UK Health Select Committee highlighted the challenges faced by public health professionals to deliver on the Five Year Forward View, citing evidence that real-terms funding of public health in England will be cut from £3.47bn to just over £3bn between 2015/16 and 2020/21.[9]

Despite shrinking public health budgets, there is still significant potential to improve population health using prevention strategies that may be cost-effective or cost saving.[2,3,10,11] So that local and national decision makers can make informed choices about how best to spend finite public health resources, they need to be able to quantify the anticipated impact of any intervention, its cost, and the associated opportunity cost. Furthermore, decision makers need to be able to directly compare two or more public health policies.

There are well established methods for modelling the cost-effectiveness of healthcare interventions. The International Society for Pharmacoeconomics and Outcomes Research-Society for Medical Decision Making (ISPOR-SMDM) guidelines for modelling research have published a key set of papers and guidelines on best practice in healthcare decision modelling.[12] The National Institute for Health and Care Excellence (NICE) has related guidance on methods for health economic modelling.[13] The ISPOR-SMDM and NICE guidance aim to standardise methods for health technology assessments (HTAs) meaning that outcomes of different studies can be directly compared with each other against a decision framework, such as the NICE cost effectiveness threshold for implementation of £20,000 to £30,000.[13] However, economic evaluations of public health interventions have their own specific challenges compared with HTAs:[14–22]

- long term health impacts (e.g. the effects on disease burden and health expenditure of increasing the price of unhealthy foods may manifest and persist for many years after the intervention is introduced);

- wider societal costs and consequences (e.g. transport policies to improve physical activity have societal costs and outcomes beyond health; and different societal costs and outcomes are relevant to different modelled scenarios, for example influencing alcohol use has an important effect on crime whereas reducing stroke incidence has an important effect on social care);
- impact on inequalities (e.g. more deprived population groups may respond differently to those who are less deprived following a price increase on unhealthy food);
- multicomponent interventions (e.g. a policy aimed at increasing physical activity may include both additional bicycle lanes and subsidised gym membership); and
- interactions within complex non-health sector systems (e.g. if simulating the effect of upgrading home insulation to reduce winter deaths, it may be important to consider interactions with the housing sector, energy sector, and social care).

In 2005, NICE started a new initiative to produce guidance on public health interventions, including cost-effectiveness analyses (different forms of economic evaluation are discussed in section 1.3).[23] As part of this initiative, NICE and the NICE Decision Support Unit set out guidelines for public health economic modelling.[13,24] The guidelines offer advice on how to address some of the specific challenges of modelling economic evaluations of public health interventions. For example, the NICE reference case for public health economic interventions states that the time horizon should be long enough to incorporate all important costs and effects; the perspective on costs may be public sector, societal, or any other as appropriate; and non-health benefits to local government and other settings may also be included.[13] However, the guidance is not prescriptive meaning that different public health economic modellers use different modelling methods (from here on referred to as the model structure), time horizons, health and economic perspectives, epidemiological data, and outcome measures. Even within academic fields with a well-established history of health economic modelling, such as diabetes, chronic obstructive pulmonary disease

(COPD), and cancer, different model structures and assumptions can produce very different outcomes despite modelling the same intervention.[25–27] Therefore, often it is not possible to directly compare results when prioritising different public health policies. As such, standardised processes are required for assessing and modelling the cost, health impact, and cost-effectiveness of public health interventions.

1.2 Aim of this thesis

The aim of this thesis is to develop a model to directly compare the cost-effectiveness of interventions affecting diet and physical activity. The intention is for the model to help local and national decision makers to allocate limited resources and maximise population health.

1.3 Economic evaluations

Economic evaluation is a tool to help decision makers decide how to spend finite resources when allocating health and public health budgets. Studies evaluating the effectiveness of interventions provide useful information on their own, however to make a fully informed decision about whether to fund a policy or intervention, a decision maker also requires knowledge of possible implications for healthcare and non-healthcare resources. Economic evaluations quantify these and are defined by Drummond et al. as "*the comparative analysis of alternative courses of action in terms of both their costs and consequences*";[28] as such, they can also enable a better understanding of foregone alternatives – the opportunity cost.

An economic approach aims to provide an explicit and theoretically based approach to quantifying costs and outcomes, enabling the comparisons of different interventions to be made.[29,30] There are different approaches to economic evaluation, however they all involve comparing the costs and outcomes of two or more different options (which may include doing nothing or carrying on with business as usual). Economic evaluations in health and

public health can take a variety of perspectives and will ordinarily include the costs and consequences of an intervention to health and healthcare as a minimum, but may also include costs to social care or other settings (as suggested by NICE,[13] see section 1.1).

1.3.1 Cost-consequence analysis

A cost-consequence analysis (CCA) presents a disaggregated list of costs and possible outcomes following an intervention and is a useful way to illustrate the potential impact of interventions on different sectors (including those other than health that might be relevant to public health decision makers). An example is Trueman and Anokye who present a CCA of the possible effects of exercise referral schemes (ERSs, see section 1.3.2.1).[31] However, CCAs make it difficult for decision makers to make comparisons between interventions and across disease areas where the types of costs and consequences differ, for example comparing a road cycling scheme aimed at reducing heart disease and diabetes with a tobacco mass media campaign aimed and reducing heart disease and cancer (further discussion of how to value outcomes for public health interventions is in section 7.3.2).

1.3.2 Cost-minimisation analysis

Cost-minimisation analyses (CMAs) compare the costs of two different interventions, assuming that the health consequences of each are equivalent. However it is rare that outcomes are similar enough to be called equivalent, especially given their associated uncertainty, and therefore CMAs are rare in economic evaluations of both health and public health interventions.[28]

1.3.3 Cost-effectiveness analysis

A cost-effectiveness analysis (CEA) compares the costs and effects of an intervention with those of one or more alternatives where the costs of each intervention use the same perspective and the effects are measured with the same unit, which are commonly natural units e.g. numbers of cancers detected or life-years saved. Results are often presented as an incremental cost effectiveness ratio (ICER) where the incremental cost of one intervention over another is divided by the incremental effect size. An example of a public health CEA is

Dallongeville et al. who estimated the cost per life year saved of different interventions aimed at increasing fruit and vegetable consumption.[32] A limitation of CEAs is that because the unit of effect is often specific to the analysis in question, it makes it difficult to compare results with other economic analyses. This is particularly important when intervention costs are from the same budget meaning it is not possible to estimate the true opportunity cost. This can be solved by using a common and generic outcome measure, such as that employed by cost-utility analyses (CUA).[28]

1.3.4 Cost-utility analysis

A CUA can be viewed as a type of cost-effectiveness analysis where the unit of outcome is generic across different health states, and incorporates life expectancy and quality of life.[28] The most commonly adopted unit of effect in England is the quality adjusted life year (QALY), as recommended by NICE.[13] A QALY is a measure of both quality and quantity of life, with one QALY equalling one year of life in perfect health. Quality of life is estimated using utilities which reflect the degree to which an individual would prefer one health state over another (their quantification is discussed in section 3.1). A QALY represents any number of combinations of quality and quantity of life, for example, one year of life with a utility score of one is equivalent to two years of life with a score of 0.5. Notwithstanding the challenges of comparing different economic evaluations of public health interventions mentioned in section 1.1, using a generic outcome measure makes it easier for decision makers to compare different interventions when disease outcomes differ. Finally, given that national organisations such as NICE prefer CUAs to other forms of economic analyses, CUAs are common in both health and public health economic modelling (see section 1.4).

1.3.5 Cost-benefit analysis

In contrast to both CEAs and CUAs, cost-benefit analyses (CBAs) instead place a monetary value on health outcomes. This then allows health and public health interventions evaluated using a CBA to be directly compared with other areas of spending such as education or defence.[30] However, Marsh et al. highlight that CBAs are rare in evaluations of public health

interventions, in part because of the challenges of how to identify monetary values of different outcomes, citing evidence that the value of preventing a fatality can vary between £0.5m and £64.0m.[14] Approaches to valuing health include the human capital approach which values life based on future earnings, and willingness to pay or discrete choice experiments.[33] Each of these methods have their own inherent methodological limitations, which are compounded by the more fundamental ethical challenge of placing a monetary value on life.[14,33,34]

1.3.6 The theoretical basis for different economic evaluations

Outcomes assessed using each of the approaches to economic evaluation discussed above are likely to be a simplification of the true health and wellbeing consequences following an intervention. A common measure of health such as the QALY is useful but may undervalue the total effects of some interventions, especially public health policies with wider-societal benefits (see section 7.3.2), and diminish the relevance of other areas of societal value that receive investment, such as education and justice.[28] Incorporating such values requires a quantification of social welfare.

The challenge of identifying everything that might contribute to social welfare is impossible both due to the scale of the task and because of inherent differences between individuals' views on what would be relevant to include. Therefore, economic evaluations rely on individuals to decide whether their welfare improves following an intervention using revealed preferences – if someone chooses (prefers) one intervention over another, then this intervention must result in greater welfare for that individual.

Welfare economics is a framework for understanding if one course of action is more socially desirable to another using non-economic outcomes.[35] Generally speaking, the main outcome of interest in welfare economics is whether social welfare is improved following an intervention, using the Pareto principle. This states that:

> *social welfare unambiguously increases only if the welfare of any member of society increases and that of no one falls.*[36]

Although there are various branches of welfare economics, they are consistent in that the only relevant way of valuing welfare is using utilities.[36,37] Welfare economics and welfarism can be applied to economic evaluations of health and public health interventions where the scenarios being compared improve the welfare of a population without affecting anybody else. However, such a scenario is rare, and in resource constrained settings (such as healthcare) it is likely that additional resources in one setting will likely lead to fewer resources invested elsewhere.

To deal with this, a potential Pareto improvement can be estimated by asking whether those who benefit from an intervention would still prefer the intervention were adopted even if they had to compensate those who lost out.[28] This is the basis for CBAs where social welfare is quantified by total individual utility and a decision can be made based on whether net social welfare is increased despite those who gain needing to compensate those who lose. As discussed by Gray et al., the use of CBAs in healthcare go back at least to the 1970s, and assume that by assigning monetary units to health outcomes, net social welfare is reflected by economic balance of costs and outcomes and therefore the full range of health and non-health outcomes can be estimated and presented.[30,38]

However, there are significant criticisms of applying a welfarist approach to healthcare. These include inherent problems of valuing health outcomes, relying on the assumption that individuals' preferences are the best way to quantify social welfare, and the lack of consideration of the effects of an intervention on the distribution of both wealth and health.[28,36,39] Therefore, rather than focusing on CBAs, health and public heath economic analyses have more commonly taken a cost-effectiveness or cost-utility approach.[40,41] Although these are based on maximising utilities and are argued by some to be founded in welfarism,[29,42] modern economic evaluations using CEAs and CUAs are generally viewed of as being aligned with extra-welfarism.[36,39,43]

Extra-welfarism is an extension of welfarism that allows for additional considerations to be included when making decisions. Individual utilities may be included, but so may other

factors that are important when assessing public health interventions such as health gain and its distribution across the population. Under welfarism, utility is valued by the individual affected, however in extra-welfarism, this can be by a delegated authority or a representative sample of the general population (as recommended by NICE[13]). Furthermore, with extra-welfarism utilities can be weighted to reflect societal preferences regarding equity and other relevant outcomes such as capability.[36] The use of other extra-welfarist approaches to public health economic evaluations such as estimating subjective wellbeing and capabilities is discussed in section 7.3.2, and NICE's preferred approach to quantifying utilities is discussed in section 3.2.

1.4 Review of model structures

1.4.1 A taxonomy of modelling structures for the economic evaluation of public health interventions

This section of the thesis presents an analysis of the pros and cons of different modelling structures giving examples of their application to the economic evaluation of public health interventions affecting NCDs. Also highlighted are examples of how different model structures cope with some of the specific challenges of public health economic modelling listed in section 1.1.

Authors have previously developed taxonomies to help modellers to decide on the most appropriate health economic model structure, however, these are primarily aimed at HTAs[21,44,45] and are difficult to apply to public health economic evaluations of interventions affecting NCDs. Brennan et al. published a taxonomy of model structures for HTAs, and to make it more relevant to public health economic modelling, Squires et al. adapted it to include agent based simulation (ABS).[45–47] Squires et al.'s adapted version of Brennan et al.'s taxonomy is used in this thesis as a framework for discussing public health economic modelling of NCDs. Comparative risk assessment (CRA) models, which can simultaneously model multiple disease processes and risk factors without large increases in model

complexity, are added to the taxonomy (table 1.1). Also included is a discussion of how microsimulation techniques and multistate life tables can be a useful adjunct to some of the model structures in table 1.1. Although other model combinations are possible, for example using the outputs of a decision tree to parameterise a Markov model,[48] these are not individual model structures in themselves and all such combinations are therefore not discussed.

In table 1.1, columns A to D divide models into population or individual level, and separate model structures by how they deal with random events, expected values, and heterogeneous populations. Broadly speaking, by using population level model structures, population heterogeneity (differences between population subgroups, for example, in terms of age, gender, or risk factor distributions) can be simulated by re-running the model for different cohorts, whereas individual level model structures use multiple samples of different types of people. The ability to incorporate randomness allows for Monte Carlo simulations to estimate stochastic uncertainty (describing the uncertainty in individual level models resulting from two individuals being in the same situation but, by chance, having different outcomes), and parametric uncertainty (in either population or individual level models and describing the uncertainty in the estimates of model parameters).[49] Rows one to four categorise model structures by whether they allow for interactions to occur between entities within the model and between entities and the environment, and how the model deals with time (*untimed* means the models do not explicitly include a temporal component). Row five describes ABS modelling, which allows for multiple interactions governed by rules affecting individuals within the model, rather than affecting the system, as is the case in rows three and four. Table 1.2 summarises the main advantages and disadvantages of the public health economic modelling structures discussed in this review.

Table 1.1 Revised version of Brennan et al.'s taxonomy of model structures[45,47]

			A	B	C	D
			Cohort/aggregate level/counts		Individual level	
			Expected value, continuous state, deterministic	Markovian, discrete state, stochastic	Markovian, discrete state	Non-Markovian, discrete state
1	No interaction	Untimed	Decision tree rollback or comparative risk assessment	Simulation decision tree or comparative risk assessment	Individual sampling model: Simulated patient-level decision tree or comparative risk assessment	
2		Timed	Markov model (deterministic)	Simulation Markov model	Individual sampling model: Simulated patient-level Markov model	
3	Interaction between entity and environment	Discrete time	System dynamics (finite difference equations)	Discrete time Markov chain model	Discrete time individual event history model	Discrete time discrete event simulation
4		Continuous time	Systems dynamics (ordinary differential equations)	Continuous time Markov chain model	Continuous time individual event history model	Continuous time discrete event simulation

5	Interaction between heterogeneous entities/spatial aspects important	x	x	x	Agent-based simulation

Table 1.2 Summary table of modelling structures for the economic evaluation of public health interventions

Corresponding section of chapter and table 1.1	Modelling method	Advantages	Disadvantages	Public health examples
Section 1.4.2.1 Table 1.1: A1, B1, C1, D1	Decision tree	Can be easy to construct. Relatively easy to interpret. Can be adapted for cohorts and individuals.	No explicit time component. Exponentially more complex with additional disease states. No looping/recurring. Poorly suited to complex scenarios.	Comparing exercise referral schemes with usual care to increase physical activity.[31]
Section 1.4.2.1 Table 1.1: A1, B1, C1, D1	Comparative risk assessment	Can model multiple diseases and risk factors simultaneously. Can be used for individuals or cohorts.	More complex to build than decision trees. No explicit time component. No looping/recurring. Unable to model interactions between individuals, populations, or their environment.	Return on investment of workplace interventions to improve physical activity.[50]

Section 1.4.2.1 Table 1.1: A2, B2, C2, D2	Markov models without interaction	Relatively straightforward to construct and to communicate. Can model populations or individuals. Has time component. Allows looping/recurring.	The Markovian assumption – individuals have no memory of (are independent of) previous disease states. Modelled populations can only exist in one disease state. Exponential increase in complexity with increasing numbers of disease states.	Investigating the cost effectiveness of different smoking cessation strategies using the Benefits of Smoking Cessation on Outcomes (BENESCO) model.[51–54]
Section 1.4.2.2 Table 1.1: A3, A4	System dynamics models	Allows for interactions between populations and the environment. Allows for feedback and recurring.	Models populations rather than individuals.	Modelling the effects of policies aimed at increasing bicycle commuting rather than travelling by car.[55]
Section 1.4.2.2 Table 1.1: B3, B4, C3, C4	Markov chain models and Markov individual event history models	Can model individuals or populations. Allows for interaction between populations or individuals within the model.	Markovian assumption still exists (although its impact can be reduced – see main text). Becomes rapidly more complex with added disease states.	A US Centre for Disease Control (CDC) model evaluating the cost-effectiveness of different diabetes prevention strategies.[56,57]

Section 1.4.2.2 Table 1.1: D3, D4	Discrete event simulation	Allows for interaction between individuals and between individuals, populations, and their environment, governed by system rules. Allows for modelling of complex scenarios.	Model structure can be difficult to communicate and interpret. Computationally demanding both in terms of designing the model and running it.	Evaluating the cost-effectiveness of screening programs.[58]
Section 1.4.2.3 Table 1.1: D5	Agent-based simulation	Allow for interactions within and between individuals, populations, and the environment, governed by rules applied to individuals. Allows for individuals to learn. Allows modelling of complicated systems.	More complex than discrete event simulation. Requires large computational power. Difficult to communicate and interpret model structure.	Scaling up the availability of secondary prevention medications for acute myocardial infarction in India.[59]
Section 1.4.3.1 Table 1.1: adjunct to A1, B1, C1, D1, A2, B2, C2, D2	Multistate life tables	Can be used with comparative risk assessment and decision tree models to add a time component. Can be combined with Markov models to increase the numbers of possible disease states without	Assumes diseases are independent of each other. Model limited by underlying model structure, for example, if combined with a Markov model, the Markovian assumption remains.	The Australian Assessing Cost Effectiveness in Prevention (ACE Prevention) project.[60,61]

		exponentially increasing model complexity.		
Section 1.4.3.2 Table 1.1: adjunct to C1, C2, C3, C4, D1, D2	Microsimulation	Can be combined with decision tree, comparative risk assessment, and Markov models to make it easier to model heterogeneous populations or multiple disease states.	Data requirements and simulations can become computationally challenging with complex models. Model limited by underlying model structure, for example, if combined with a Markov model, the Markovian assumption remains.	The NICE obesity health economic model used by Trueman et al. to estimate the cost-effectiveness of primary care weight management programs.[62]

1.4.2 Descriptions of model structures with examples

1.4.2.1 Rows 1 and 2 – no interaction

Decision trees (row 1, columns A, B, C, and D)

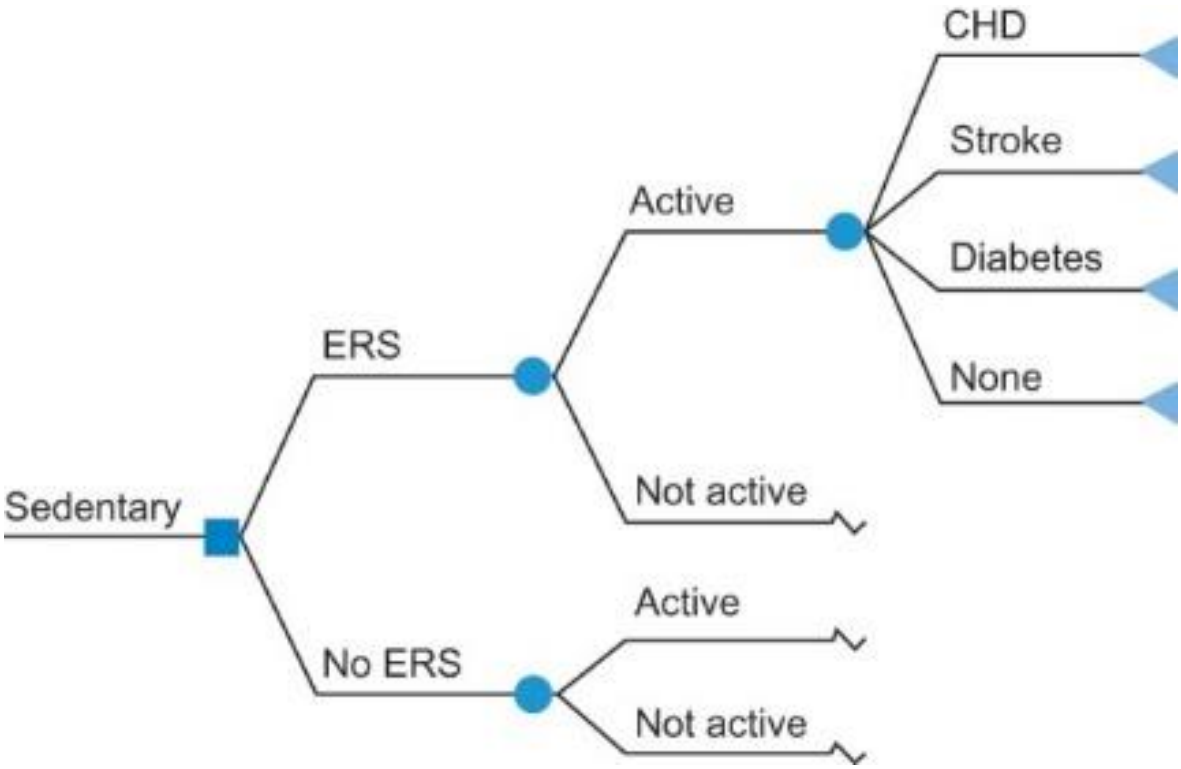
Decision trees simulate possible decisions and outcomes using branches to represent each potential event. Branch points are usually described as nodes and can represent a decision (i.e. whether or not a public health education campaign takes place), or a chance event (i.e. development of disease). The options at each chance node are assumed to be mutually exclusive and the probability of each option occurring must sum to one. Each branch eventually ends with a terminal node against which the associated morbidity and costs of that patient journey can be attached (figure 1.1).[63] In order to calculate the cost-effectiveness of an intervention, the costs and morbidity of each patient journey are multiplied by the probability of that journey occurring before being summed for each intervention and compared.

Decision trees are one of the simpler model structures available for public health economic modelling, and as such they are transparent as well as relatively straightforward to construct and to analyse. Decision trees can operate at the cohort level or at the individual level but do not have any explicit time component, do not allow for looping (recurring) of events, and do not accommodate interactions between individuals or populations.[45] This makes it particularly challenging to model long term chronic conditions or relapses in disease without becoming overly complicated.[63]

Although often used by health economists, decision trees are less commonly used in public health economic models as they are constrained by their lack of a temporal component, which other model structures can handle more efficiently.[47,64] Trueman and Anokye used a decision tree to model ERSs promoting physical activity in the UK (figure 1.1).[31] Alongside a CUA, Trueman and Anokye estimated the effects of ERSs on a variety of health outcomes and work absenteeism with a CCA (see section 1.3.1. In terms of simulating the effect of ERSs, the model has various assumptions meaning it fails to fully capture the true

complexity of the intervention. For example, the population of 40-60 year olds simulated are assumed to be homogeneous in terms of both their baseline characteristics and their response to the intervention, are defined as being either active or inactive without any greater resolution of physical activity levels, and can only develop one disease outcome in their lifetime (either coronary heart disease [CHD], stroke, or type two diabetes). Using a CRA model (see below) would have been one mechanism for Trueman and Anokye to include additional diseases without significantly increasing model complexity.

Figure 1.1 Example of a decision tree, from Trueman and Anokye[31]



Reproduced from Trueman and Anokye under the terms of the Creative Commons CC BY license.[31] Squares represent a decision node, circles represent a chance event, and triangles represent a terminal node. The symbol ( ) signifies that branches indicating the outcomes (CHD, stroke, diabetes, and none) apply. ERS, exercise referral scheme; CHD, coronary heart disease

Comparative risk assessment (row 1, columns A, B, C, and D)

Comparative risk assessment models have also been used for public health economic modelling of NCDs.[32,50,65] These are commonly aggregate level models that estimate how

parameters describing the relationship between a risk factor and disease outcome would change following an intervention using population impact fractions. Individuals can be simulated by combining CRA models with microsimulation (see section 1.4.3.2). Models using CRA do not allow for interactions but can simulate the age and sex specific effect of an intervention on multiple risk factors and disease processes simultaneously without becoming as complex as the equivalent decision tree. Other population strata can be simulated in the same way, thereby estimating the impact of an intervention on health inequalities. Furthermore, CRA models can be linked to multistate life table models to include a time component (see section 1.4.3.1). Costs and morbidity associated with each disease process can be applied to the modelled population to compare total cost and health outcomes with and without an intervention. Examples are the World Health Organization (WHO) Comparative Risk Assessment project,[65] a related global health model estimating the cost effectiveness of sodium reduction strategies,[66] a French model simulating changes to fruit and vegetable consumption,[32] and a US model by Trogdon et al. estimating the potential return on investment of workplace interventions to improve physical activity and reduce obesity.[50]

Trogdon et al. compared the costs of staff absenteeism and medical expenditures for US employees with and without physical activity interventions aimed at reducing staff body mass index (BMI).[50] The model estimated absenteeism and medical costs as functions of BMI for the US workforce aged 18-65 years (whilst controlling for multiple potential confounders), and using these data, predicted what the overall costs would be were BMI to be reduced, based on the number of employees existing in each BMI half unit. The use of this model to US employers and decision makers is enhanced by ensuring the model's structure is transparent, and including costs from the employer perspective as well as from a health perspective. Furthermore, the model's input variables (costs, BMI distribution, and effects) are generic to the US population meaning the model is transferable between employers in the US. Trogdon et al. have adapted the model to allow the user to estimate costs and outcomes for interventions with different effects on BMI in the first year

compared with subsequent years. To do this, the model calculates the annual costs and benefits in year one with year one specific data separately to calculating costs and benefits for year two onwards, thereby partly solving the problem that CRA models do not have an explicit time component.

Trogdon et al.'s model estimated some outcomes and costs outside of the health sector by capturing costs from the employer perspective. This approach of estimating wider societal costs and consequences directly from the health outcome (such as lost days at work as a result of sickness) can be applied to all model structures where data on the societal implications of a given health outcome are available. If data were available and it were useful to the model's users, Trogdon et al.'s model could also be adapted without changing its fundamental structure to include a greater number of disease processes, calculate age and sex specific effects, and incorporate a wider range of costs, such as avoided future costs through obesity prevention, costs associated with improved productivity, and changes to insurance and disability costs.[50] The model could also be developed to estimate parametric uncertainty using Monte Carlo analyses. Finally, the model currently assumes that individuals exposed to the intervention will have the same percentage impact on their body weight irrespective of their baseline age, gender, BMI, employee status, work environment, exercise levels, diet, or any other potentially modifying variable. In order to incorporate the effect of these variables (assuming data are available), a more complicated CRA model could be developed within its current structure, or the model could be adapted into an individual level model using microsimulation techniques (section 1.4.3.2). Finally, if it were thought important to include interactions, such as a cohort effect where more people engage with the intervention as larger proportions of the workforce begin to see benefits, then a different model structure which allows for interactions would be needed (from rows 3, 4, or 5 of table 1.1). However, all methods that increase the complexity of the model require additional data and may make it less transparent and user-friendly for decision makers.

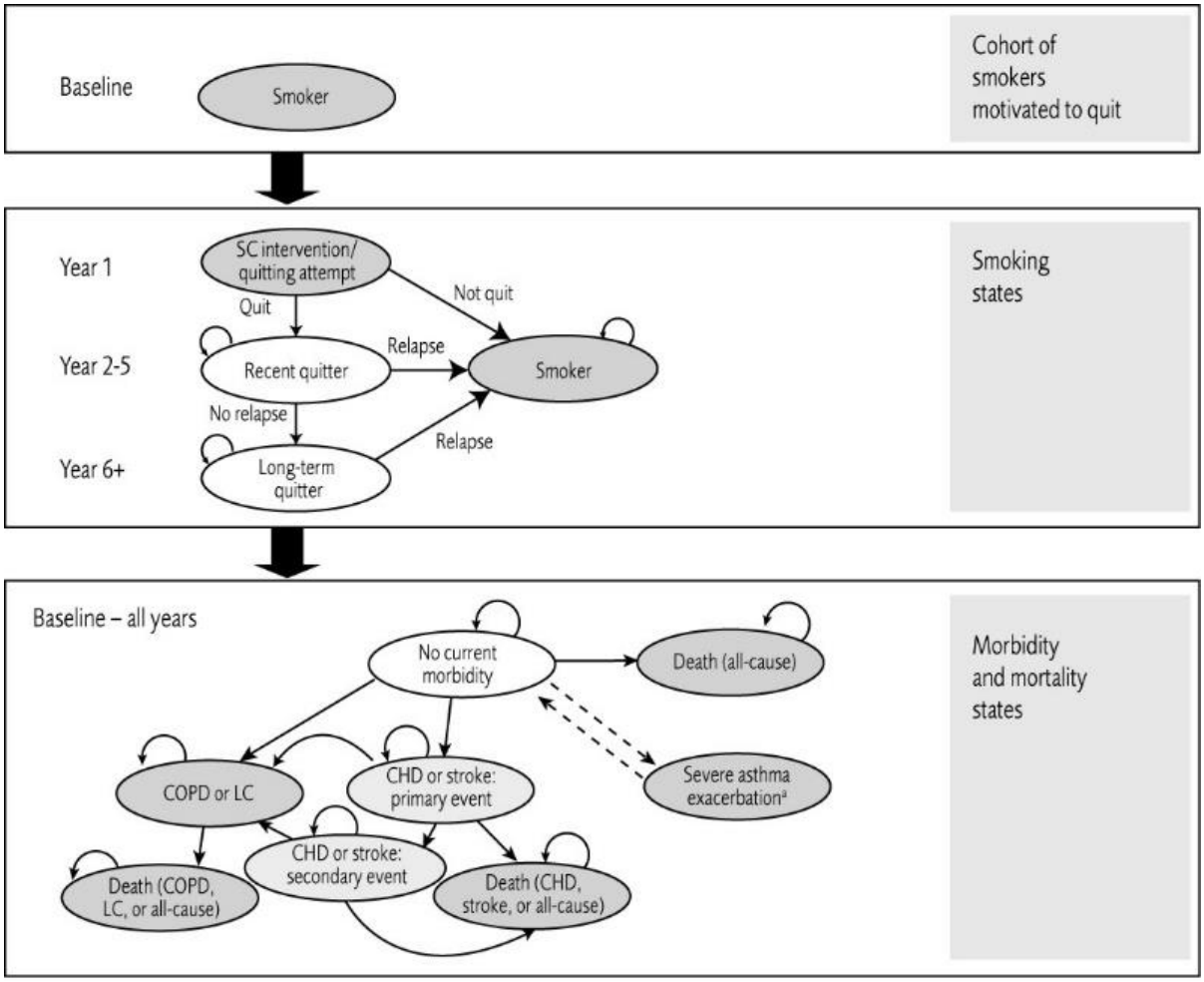
Markov models without interaction (row 2, columns A, B, C, and D)

In comparison to decision trees and CRA models, Markov models are much more commonly used in public health economic modelling of NCDs, and are able to simulate more complex scenarios.[64] For example, a Markov model may be used to model the cost-effectiveness of different smoking cessation methods incorporating multiple disease outcomes, the probability of relapse, and outcomes over different time horizons.[51–54] Markov models simulate how a population or individual moves between pre-defined health or disease states at a specific time interval (for example, annually). This incorporates a time component and allows for modelled populations to remain in a health state from one time interval to the next, and to loop back from a diseased state to a healthy state (recur), all based on transition probabilities. The model can then be run for either a pre-defined number of cycles or, if using a population cohort, until the entire population has reached a certain age or died. In this way, long-term effects of interventions on disease outcomes can be estimated. When computing cost-effectiveness from a Markov model, each health state is assigned a measure of morbidity and a cost. As the model runs through cycles, the costs and morbidity (or utility) are then summed for the numbers of individuals in each state at each time cycle.

Markov models can be relatively straightforward to develop and to represent graphically thereby making them transparent to peer-reviewers and decision makers (figure 1.2). Furthermore, transitional tunnel states can be added to increase complexity and make the model more realistic. For example, in order to more accurately capture the increased costs and morbidity associated with the first year of having a heart attack compared to subsequent years, all individuals who have a heart attack may spend one cycle in a tunnel state with associated high morbidity and cost, before having a 100% probability of leaving that state. Markov models have been commonly used in a range of different public health economic analyses of NCDs in different settings and countries. For example, the Dutch Rijksinstituut voor Volksgezondheid en Milieu (RIVM) Chronic Disease Model (CDM),[67–70] the Benefits of Smoking Cessation on Outcomes (BENESCO) model,[51–54] the Australian Quit

Benefits Model,[71,72] the US Centre for Disease Control (CDC) Measurement of the Value of Exercise (MOVE) model,[73] the Australian Coronary Heart Disease Prevention Model,[74] the US CHD Policy Model,[75,76] and bespoke models for analysing public health interventions internationally,[77] in the Netherlands,[78] Switzerland,[79] Finland,[80,81] Germany,[82] US,[83,84] UK,[85–87] and in Australia.[88,89]

Figure 1.2 Example of a Markov model, taken from the BENESCO model published by Bolin et al.[52]



SC, smoking cessation; LC, lung cancer.
[a]Asthma exacerbation requiring a physician visit.

Reproduction of figure 1 in the supplementary material from Bolin et al., Cost-effectiveness of varenicline compared with nicotine patches for smoking cessation-results from four European countries. European Journal of Public Health. 2009; 19: 650–4. With permission of Oxford University Press and the European Public Health Association. The ovals indicate health or disease states with arrows indicating how the population, or proportions of the population move between states. Arrows looping back on themselves indicate that the

population remains in the same state. SC, smoking cessation; LC, lung cancer; CHD, coronary heart disease; COPD, chronic obstructive pulmonary disease

Markov models have some important assumptions. Firstly, population level models have no memory (the Markovian assumption) meaning that the morbidity, cost, and transition probabilities associated with a given health state are the same irrespective of previous health states or how long an individual has been in a health state. This can, in part, be dealt with by using tunnel states, as used by the Australian Quit Benefits Model,[71] or by using microsimulation (see section 1.4.3.2). Secondly, individuals can also only exist in one state at a time. This means that to add a new disease to a model that can co-exist with the previously modelled diseases, the number of health states included in the model must be exponentially increased (as new health states are required for each disease combination). Further complexity is introduced if trying to model a heterogeneous population where population subgroups have different transition probabilities. This can be addressed by using weighted average costs, disabilities, and transition probabilities at the aggregate level. Alternatively, multiple cohorts, each with different characteristics, can be run through the model with cohort specific transition parameters to give results by population subgroup, which can also be useful when estimating effects on inequalities; however, this leads to additional model complexity and run time. An example of using multiple Markov models to model a heterogeneous population with multiple disease states is the US CDC diabetes prevention model, discussed in more detail in section 1.4.2.2.[56,57]

Howard et al. used the BENESCO model to simulate the lifetime health perspective costs and the disease consequences of a population of US smokers making a single quit attempt in year one of the Markov model's simulation (figure 1.2).[51] The model includes four disease states: asthma exacerbation, following a first event of CHD or stroke, following a second (or further) event of CHD or stroke, and COPD or lung cancer; and three terminal absorbing states: death (CHD or stroke), death (COPD or lung cancer), and death (all other causes). The population simulated through the model is divided by gender and three age categories,

and transition probabilities vary for each age-sex group, as well as by their smoking status (current, former, never). The model has an explicit time component, allows for recurrence of disease and relapse of exposure, simulates multiple diseases, and estimates parametric uncertainty. The limited number of disease states makes the BENESCO model straightforward to illustrate and communicate, however this also limits how well the model reflects reality.[51] For example, the model simulates an unrealistic disease hierarchy such that once in the COPD or lung cancer state, the simulated population cannot enter a CHD or stroke disease state. Furthermore, the model suffers from the Markovian assumption meaning that the costs and morbidity associated with the COPD or lung cancer disease state are the same irrespective of whether the population entering the state comes from either the healthy state or the CHD disease state. A more realistic model could be developed using a more complex Markov model with greater numbers of single and multiple disease states. To avoid the model becoming overly complicated, this could be combined with a multistate life table (section 1.4.3.1). Further complexity, as well as avoiding the Markovian assumption, could be achieved using discrete event simulation (DES) or ABS (sections 1.4.2.2 and 1.4.2.3), however this would be at the cost of transparency and ease of use by decision makers.

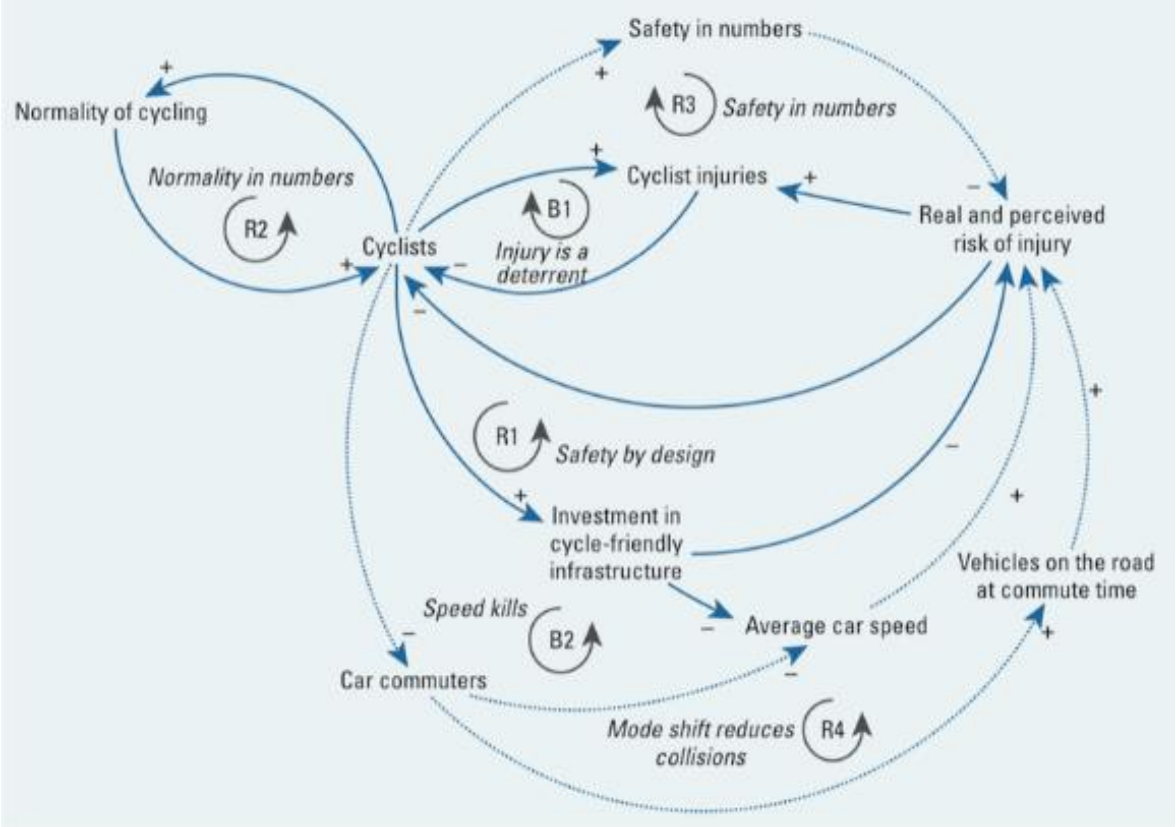
1.4.2.2 Rows 3 and 4 - interaction allowed

System dynamics models (rows 3 and 4, column A)

System dynamics models allow for populations to interact both with each other and with their environment. The probabilities of events occurring in the model (the system) change through feedback as the model runs, governed by algebraic or differential equations.[90] Such a model can be made increasingly complex as more factors influencing the system are added (requiring additional data). This makes system dynamics models better able to simulate interactions within complex non-health sector systems, and to estimate effects of multicomponent interventions, than previously discussed model structures. Costs can be

applied to either the disease state, or to all factors within the model, and then cost and health outcomes with and without the intervention can be compared. System dynamics models can usually be represented graphically, facilitating communication of the model with stakeholders. Such models are well-established for communicable diseases[91,92] and are increasingly being applied to NCD risk factors, such as Macmillan et al. who used a continuous time system dynamics model to explore the potential effect of different transport policies on bicycle commuting in Auckland, New Zealand (figure 1.3).[55] The use of a system dynamics model enabled Macmillan et al. to include complex interactions, represented by positive or negative feedback loops, and to estimate health, social, and environmental outcomes over a period of 40 years. The ability to represent diagrammatically the model makes it possible to communicate and discuss its face validity with relevant stakeholders. System dynamics models operate at the population level, rather than individual, making it more difficult to model heterogeneous populations. Were results required for population subgroups such as different ages or deprivation level, the model would either need to be replicated and appropriately parameterised for each subgroup, or an individual level model allowing interaction could be used, such as DES. A potential limitation of system dynamics models is that the dynamic element of the model (the rate of change of parameters with time) is deterministic, although parametric uncertainty can be estimated (as was done by Macmillan et al.).

Figure 1.3 Example of a system dynamics model, taken from Macmillan et al.[55]



Reproduction of figure 1 from Macmillan et al. in Environmental Health Perspectives.[55] This shows the causal loop diagram developed by Macmillan et al. for their system dynamics model. Dotted lines indicate loops identified by stakeholders and the literature but local data suggest they are inactive. A (+) sign indicates a change in the variable at the end of the arrowhead in the same direction as the originating variable, a (-) sign indicates the opposite. R and B illustrate the feedback loops in the model, where R indicates reinforcing or positive feedback, and B indicates balancing or negative feedback.

Markov chain models and individual-level Markov models with interaction (rows 3 and 4, columns B and C)

In discrete or continuous time Markov chain models, state transition probabilities can depend on (interact with) the proportion of different populations in different disease states, and on the time that has elapsed in the model. These interactions are the key difference between Markov chain models and Markov models without interaction discussed in section 1.4.2.1, and provide the model with some degree of memory, in part overcoming the Markovian assumption. However, when simulating cohorts using time-dependent transition factors, it is not possible to completely negate this assumption because the model

cannot know how long different proportions of the population in a given health state have been in that state. By contrast, individual-level Markov models simulate individuals separately (microsimulation, see section 1.4.3.2), making it possible for the model to "know" how long an individual has been in each state and to alter transition probabilities as a function of time. The simulation may or may not retain memory of which states an individual has been in previously. Markov chain and individual-level Markov model structures with interaction are better able to cope with system complexity than CRA and decision tree model structures, however they are unable to explicitly model non-health sector system interactions.

Examples of discrete time aggregate-level Markov models allowing for interaction are a version of the RIVM CDM (also cited in section 1.4.2.1) which includes disease incidence parameters that depend on time from smoking cessation,[93] a cohort model used in Australia and New Zealand where CVD transition probabilities depend on the time elapsed in the model,[94–96] and the US CDC diabetes prevention model, described in detail in a technical report by Hoerger et al., available from Herman et al. as an online supplement.[56,57]

Herman et al. estimated the cost-effectiveness of metformin versus lifestyle interventions for the prevention or delay of type two diabetes among adults with impaired glucose tolerance. The model differs from the BENESCO Markov model[51] discussed in section 1.4.2.1 in that many of the transition probabilities between disease states are functions of time since diagnosis of diabetes, levels of glycaemia, and hypertension. Furthermore, the model simulates multiple disease processes simultaneously by allowing the cohort to coexist in five different disease pathways which are linked to the overall Markov model by bridge models (see online supplement from Herman et al. for full description of the model).[57] Bridge models allow the overall Markov model to collect accumulated data on the number of events that have occurred, and keep track of the proportion of the cohort remaining in each disease state in any given year and the proportion who have left either through death or remission. Finally, Herman et al. accounted for a heterogeneous

population by simulating 560 different cohorts, each with individual state transition probabilities dependent on age, sex, race, hypertension, cholesterol, and smoking status. The model is limited, as with all Markov models, by the Markovian assumption – once in a disease state, the associated costs and disability are independent of any previous disease state. Also, although Herman *et al.* estimated costs from both a health system and societal perspective, the model does not include interactions between individuals and society. This could be relevant if, for example, a more health conscious population following the lifestyle intervention creates a demand for healthier food choices or increased access to recreational space leading to further population health benefits. If this degree of complexity were required then the model would be better suited to DES or ABS (which would also overcome the Markovian assumption, discussed below). Finally, due to the complexity of the Markov model used by Herman et al., it is more difficult to illustrate the model's detail than simpler Markov models.

Discrete event simulation (rows 3 and 4, column D)

Discrete event simulation (DES) is an extremely flexible modelling structure that simulates a system changing over time with a sequence of discrete individual events.[97] Rather than simulating populations or individuals through states for a fixed length of time, multiple future events are in competition and the model jumps to whichever event occurs next based on predefined probabilities. The occurrence of an event can directly lead to a series of contemporaneous events, as well as affect the probability of future events. The probability of a given event occurring can also vary with time and be affected by interactions between individuals, populations, and their environment at each event. A set of system rules and probabilities govern the behaviour of the population or individuals in a DES model, and these can be changed based on the intervention being modelled. As each event occurs, costs and utilities based on the event, consequence, and time to event are estimated. Along with other model structures that can simulate interactions between the population being modelled and the environment, DES is well equipped for addressing interactions within complex systems. However, due to the large number of variables possible in DES and the

need to simulate many individuals, models can be computationally slow to run (particularly when estimating uncertainty) and require large amount of data for each disease outcome. Within public health economic modelling of NCDs, this approach has been commonly adopted for evaluating screening programs,[58] and is the basis of the Archimedes model, built to simulate a wide range of clinical, financial, and administrative scenarios affecting the US health system.[98,99]

Stout et al. estimated the costs and health outcomes associated with screening mammography for breast cancer in the US using DES.[100] The authors modelled 64 alternative screening scenarios using discrete time DES to estimate their cost-effectiveness. Women in the US aged 20 years or over between 1990 and 2000 were simulated through four model components: the natural history of breast cancer as it develops, detection of breast cancer via screening mammography or clinical signs, improvements in treatment and treatment diffusion, and incidence of death from breast cancer or other causes. The model was calibrated such that interactions between the four model components over time resulted in modelled breast cancer outcomes that mimic actual US cancer registry data from 1990 to 2000. The model was then used to simulate the 64 alternative screening scenarios over the same time period, offering valuable information on the cost-effectiveness of alternative breast cancer screening regimes in the US (such as varying the time between screening and the offered age-range).

Stout et al. could develop their model, assuming data were available, to estimate a wider range of outcomes such as societal costs and consequences of targeted screening among women at different baseline disease risks. The main challenge for decision makers wanting to use this model is in communicating its complexity. To this end, the Cancer Intervention and Surveillance Network (CISNET) developed by the US National Cancer Institute publishes standardised model profiles to assist with cancer model interpretation and comparisons (for the model used by Stout et al. see the Wisconsin-Harvard model available from: http://cisnet.cancer.gov/breast/profiles.html).[101] If it were useful to allow individuals

to learn through the model, for example individuals moderating their risk behaviour (such as smoking or screening attendance) in response to different screening policies or to other individuals developing symptoms, then an ABS model could be used.

1.4.2.3 Agent-based simulation (row 5, column D)

Models using ABS have many similarities to DES in that they allow for the probability of events occurring within the system being modelled to change as a consequence of interactions between individuals (agents), between agents and the environment, and with time. Therefore, ABS can deal with the challenges of multicomponent interventions and interactions within complex non-health sector systems. The difference between ABS and DES, however, is that ABS models apply rules to agents or groups of agents, and responses depend on individual agent characteristics which can change either over time or following interactions with other agents or the environment. This is compared to system-based rules found in DES.[46,102,103] A heterogeneous population of agents is therefore able to learn over time and affect the system, which, as Squires discusses in her thesis, allows for more accurate representation of spatial or social effects, such as social networks.[47,104] Costs and morbidity can be applied to each event and disease state in order to derive estimates of cost-effectiveness of interventions that affect the behaviour of agents (other societal costs and outcomes could also be estimated in a similar fashion). Population-level results then emerge from the aggregate of all agent-level results. However, ABS is again more complex than DES and can require large amounts of data to represent heterogeneous population groups.

The use of ABS is rare in health economic modelling,[103] and even more so in public health economic modelling. The IndiaSim model has been used to simulate the cost-effectiveness of policies affecting the prescribing of secondary prevention drugs for myocardial infarction (MI) in India.[59] Only an abstract of this work has been published, with the model otherwise being used for analysing the impact on infectious disease prevention in India of vaccination and improved sanitation.[105,106]

1.4.3 Using multistate life tables and microsimulation to increase model flexibility

1.4.3.1 Use of multistate life tables with decision tree, comparative risk assessment, and Markov models with no interaction (rows 1 and 2, columns A, B, C and D)

In this thesis, multistate life tables are defined as life tables that model an individual's, or proportion of a population's, probability of developing a given disease at different ages and subsequent case fatality rates once the disease is acquired. These can simulate multiple diseases simultaneously and can be used to add a temporal component to decision tree or CRA models. For each intervention scenario being analysed, the decision tree or CRA model can generate population impact fractions to alter the multistate life table disease incidence, case fatality, and overall mortality rates. Rerunning the multistate life table model then allows scenarios to be compared over the number of years in the life table allowing for long-term health effects and economic impacts to be estimated.

In a similar way, multistate life tables can be used with population or individual Markov models to simulate multiple diseases without the need for large increases in the number of disease states. To do this, proportions of the population can coexist in more than one disease state in the multistate life table and for each disease a Markov model can simulate the probability of moving between diseased and non diseased states (thereby performing a similar function to bridge models discussed in section 1.4.2.2). These properties usually assume that diseases are independent of one another (for example, the probability of developing ischemic heart disease does not change with a concurrent diagnosis of cancer). Published examples include the Australian Assessing Cost Effectiveness in Prevention (ACE Prevention) program of research (which estimates the effects of various interventions over the lifetime of the Australian population),[60,61,107–110] the WHO PopMod model (which estimates effects over a 100-year time period),[111,112] the US model used as part of the Childhood Obesity Intervention Cost-Effectiveness Study (CHOICES, estimating outcomes over a 10 year period),[113] and a New Zealand model estimating the effects of increasing

tobacco taxation (which estimates effects over the lifetime of the New Zealand population).[114]

An example of research from the ACE Prevention programme is Cobiac et al.'s model assessing interventions aimed at improving physical activity.[60] The authors modelled the impact of six interventions (such as general practitioner referrals to an exercise physiologist or mass-media campaigns) on life-long health outcomes and health-sector costs for the Australian population. An overall multistate life table described population age and gender-specific disability and mortality rates. These rates were altered following an intervention based on five disease-specific Markov models (one for each of the diseases included in the model: ischaemic heart disease [IHD], ischaemic stroke, type two diabetes, breast cancer, and colon cancer), each parameterised by a disease specific multistate life table in which the modelled population simultaneously exists. The degree to which the rates of disability and death are changed in the overall life table is proportional to the amount of the population in each simulated disease state, and the degree of disability associated with that disease. A major assumption of this model is that diseases are assumed to be independent of each other, however Cobiac et al. partly overcame this limitation by increasing the risk of the modelled cohort developing CVD based on the proportion of the population with type two diabetes.

1.4.3.2 Use of microsimulation with individual-level decision tree, comparative risk assessment, and Markov models (rows 1 and 2, columns C and D; and rows 3 and 4, column C)

In order to overcome the complexity of modelling multiple diseases and heterogeneous populations in decision tree, CRA, and Markov model structures, an alternative approach is to use individual patient simulation models (microsimulation). These allow for a population of heterogeneous individuals to move through the model based on probabilities appropriate to their characteristics (such as age, gender, and socioeconomic status [SES]). The model is run at the individual level with all members, or randomly selected members of a predefined population, being simulated until either a pre-specified outcome occurs or

a certain length of time has elapsed (e.g. death or reaching age 100). Once completed, individual results can be aggregated to calculate either a single population-level result or percentile (or other) variations in results across individuals (thereby also allowing reporting of inequalities). Microsimulation is also useful when estimating population means of skewed effects (such as the growth rates of different cancers when a few may be very quick-growing and have different subsequent events compared to the majority being slow-growing), and is very flexible at modelling interactions.

Model parameters can be changed for different scenario analyses in the same way as multistate life tables (section 1.4.3.1) by using decision tree, CRA, or Markov models. However, in microsimulation this is at the individual level and parameters are specific to the individual's characteristics. Drawbacks of individual-level simulation are that they are computationally intense although modern computers and software are increasingly able to cope with the many thousands of iterations often required. Examples of patient-level simulation Markov models within public health economic modelling of NCDs include the UK Health Forum model,[115] the WHO and Organisation for Economic Co-operation and Development (OECD) Chronic Disease Prevention (CDP) model,[116,117] and the NICE Obesity Health Economic Model,[62,118] as well as examples from Australia,[119] Korea,[120] and Sweden.[121,122]

The NICE Obesity Health Economic Model was used by Trueman et al. to estimate the long term cost-effectiveness a primary care weight management programme called the Counterweight Programme. The micro-simulation model was run for the life-time of 10,000 individuals with age, gender, and BMI distributions representative of the UK population. The Markov model upon which the microsimulation model was based estimated the probability that at each six month time period an individual either gains weight, loses weight, or remains the same weight. There are then associated probabilities of moving into one of five health or disease states: developing diabetes, colon cancer, or CHD, dying, and having no weight related co-morbidity. At each six month interval, the quality of

life and healthcare costs for each individual were calculated depending on their BMI and disease state. The age, sex, and BMI related probability of moving into a disease state was then re-calculated every six months. Using a micro-simulation model means that results can be easily reported by individual-level characteristics. The Markov model underlying the NICE Obesity Health Economic Model is limited by the number of BMI related diseases included and baseline demographic data. Although it would be possible to add more diseases states and demographic details (such as SES), data were not available with which to parameterise the associated disease outcomes, highlighting the high data requirements of more complex models.

1.5 Addressing the challenges of public health economic modelling

Modelling the cost-effectiveness of NCD-related public health interventions is an expanding academic field that is starting to embrace more sophisticated modelling structures. Modellers have thus far primarily used Markov models, with or without the use of life tables and microsimulation, to investigate the health impacts of a given intervention and its cost-effectiveness. However, there is scope for more complex systems to be modelled with a wider range of outcomes by using model structures such as DES and ABS. Balancing transparency and parsimony with complexity when developing such models is crucial for model results to be readily interpreted and used by decision makers.[45,123–125]

Many of the models cited in this chapter address some of the specific challenges of public health economic modelling; namely assessing long-term health and economic impacts, quantifying societal costs and consequences, identifying the potential impact on inequalities, simulating multicomponent interventions, and simulating interactions within complex non-health sector systems. The solutions presented are either specific to a given model structure, or can be applied across all structures.

For example, long-term health and economic impacts are impossible to estimate using cross-sectional model structures such as CRA and decision tree models structures.

However, combining these with a multistate life table can add a longitudinal component to the model thereby enabling long-term outcomes to be simulated.[60,112]

The challenge of quantifying societal costs and consequences (e.g., air pollution or effects on employment) is less a problem of choosing the appropriate modelling structure, but more an issue of identifying robust cost and non-health data with which to parameterise the model. For example, if data are available it may be possible within any model structure to estimate some of the effects on productivity costs of an intervention directly from a model's health outcomes, such as inferring what may happen to work absenteeism as a consequence of ill health.[31] Some model structures are better suited than others for estimating the costs and consequences of other non-health related outcomes or where interactions are required. For example, Macmillan et al. were able to simulate non-health related outcomes using a system dynamics model and relate air pollution (and its cost) to part of the causal pathway (number of vehicles on the road) between the intervention and health outcome,[55] something that would be possible to do using any model structure where the number of vehicles on the road was part of the causal pathway. For outcomes that cannot be directly estimated from a step on a model's causal pathway or result from interactions with other model outcomes (for example, if increased air pollution affected weather patterns which in turn affected cycling habits) then model structures allowing for interactions are likely to be more appropriate.

The potential impact of a simulated intervention on inequalities can also be estimated using any modelling structure by simulating population groups of interest separately if using a cohort model (for example, Blakely et al. estimated the effects of a tobacco tax by ethnic group in New Zealand using a multistate life table model[114]) or aggregating results fromtwo or more population groups of interest using microsimulation (such as that used by Feldman et al. when estimating the effects of lifestyle interventions on different risk groups for metabolic syndrome[122]). Estimating wider societal costs and consequences is again dependent on available data.

The challenge of multicomponent interventions can be addressed both using model structures that do and do not allow for interaction. Without interaction, a model assumes that each component of a multicomponent intervention acts independently on a disease outcome or risk factor. If interaction between multiple interventions is necessary to simulate, then model structures that allow interactions are required. Both for multicomponent interventions and for simulating interactions within complex non-health sector systems, microsimulation models (such as ABS and DES) offer more flexibility than population-level models.

Outside of the economic model structures discussed in this chapter, regression models may be useful to estimate the impact of an intervention on one or more outcomes. An example is the CBA model developed by New Economy designed to help local authorities and other public agencies in Greater Manchester evaluate possible regional interventions.[126] The model reports how interventions affecting outcomes such as employment or antisocial behaviour may have financial implications for various agencies including the police, NHS, and local authorities. Financial benefits are separated into fiscal benefits arising from changes to taxed revenue, and public value benefits such as economic growth and improved health and wellbeing. These are quantified using unit costs taken from a range of sources. After adjusting results for benefits that would be expected under business as usual, the potential impact of an intervention on costs is further adjusted based on the quality of the underlying data used to estimate its impact and cost consequence. This CBA model, and others like it, has the benefit of being able to compare costs and savings across multiple interventions and multiple sectors of society. It also includes a temporal component to provide an estimate of costs and return on investment over a period of up to 25 years. However, health outcomes are only assessed in terms of their potential economic savings to the NHS without any attempt to quantify the change to quality of life either in terms of utilities or monetary units. Therefore it is not possible to compare results with other health or public health economic models. Furthermore, estimating the impact of interventions on

population subgroups and inequalities would require separate analyses for each subgroup, and there is no ability to model interactions.

Finally, it is important to note that when dealing with any of the challenges of public health economic modelling, a model is only as good as the data used to parameterise it and adding more complexity may only serve to make the model more uncertain and more difficult to communicate. As Whitty discusses, a model that is simple, timely, and lays bare its problems is far more useful to a policymaker than one that is more detailed, more complicated, possibly more accurate, but less interpretable and arrives after the policy decision is made.[125]

1.6 Conclusion

This chapter presents a revised version of Brennan et al.'s taxonomy of model structures for the economic evaluation of HTAs[45] and discusses the advantages and disadvantages of these structures in relation to NCD public health economic models. Of the studies cited, no one model addresses all the challenges of modelling economic evaluations of public health interventions, and the choice of which model structure to use will depend on what is being modelled: the interventions being evaluated, the outputs required, and the needs of the decision maker. This chapter is therefore a guide to the choice of model structure to use in this thesis, discussed in chapter two.

Chapter 2. Model Structure

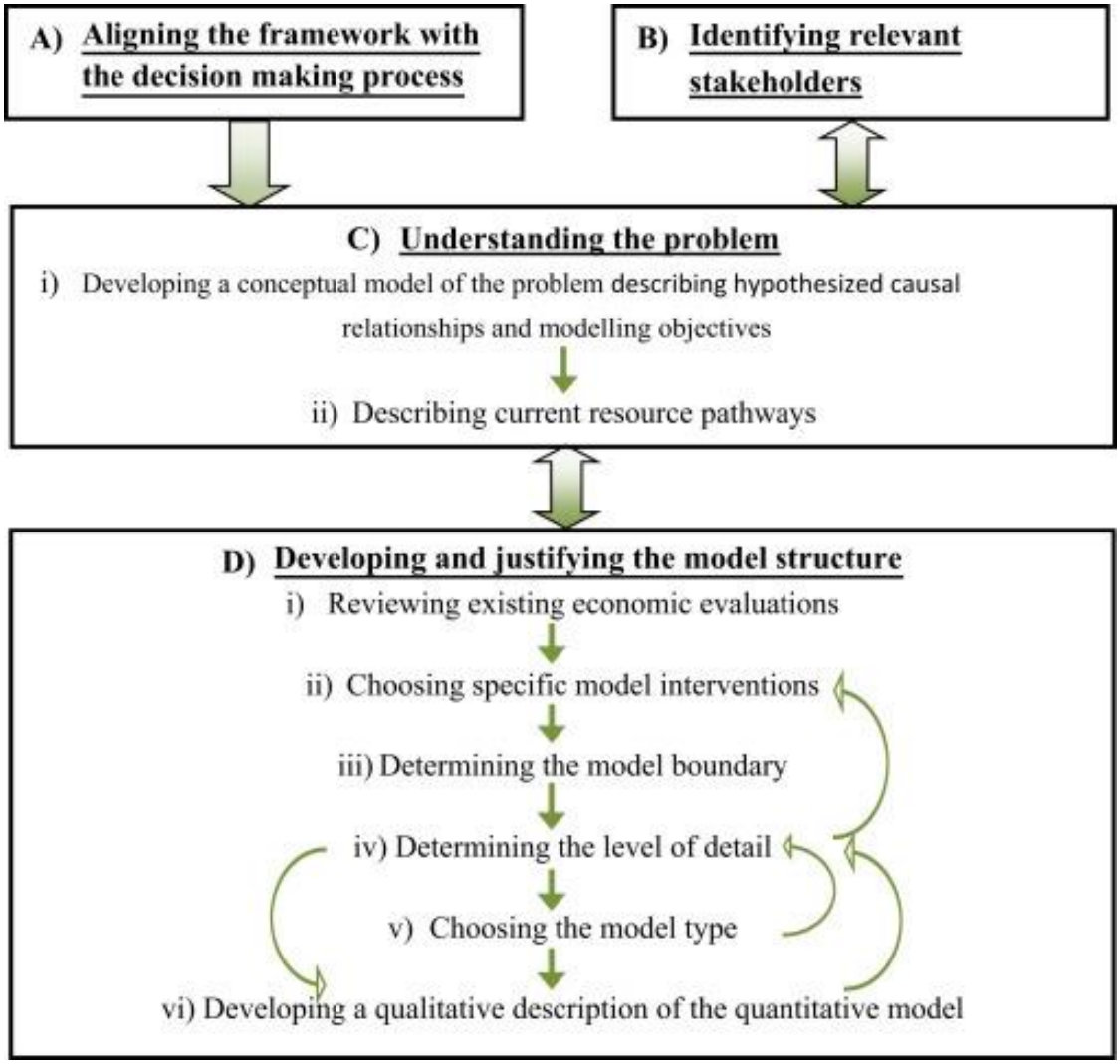
The previous chapter classified public health economic model structures and outlined their major strengths and weaknesses. This chapter describes how the model structure used in this thesis was chosen and developed.

2.1 Structure and scope of the conceptual model

In order to better understand the challenge of estimating the cost-effectiveness of public health interventions affecting diet and physical activity, a conceptual model was developed. A conceptual model is a qualitative description and visual representation of the proposed simulation model and has two principle uses in this thesis. Firstly, it is a tool for thinking through the potential relationships and interactions affecting the system being simulated, thereby acting as a reference point and guide for the model's subsequent development. Secondly, it provides stakeholders with a visual description of the model so that they can understand and critique the model during its development and when interpreting its outputs.[46]

When developing the conceptual model, Squires et al.'s framework was used as a reference (figure 2.1).[46]

Figure 2.1 Squires et al.'s conceptual modelling framework[46]



Following a review of the literature and interviews with stakeholders, Squires et al. identified four key principles of good practice when developing a conceptual model:

i. a systems approach to public health modelling should be taken;
ii. a documented understanding of the problem is imperative before and alongside developing and justifying the model structure in order to develop valid, credible, and feasible models;

iii. strong communication with stakeholders and members of the team throughout model development is essential for model transparency, validity, and credibility;
iv. a systematic consideration of the determinants of health is central to identifying all key impacts of the interventions in public health economic modelling.[46]

Where appropriate, these four principles have been used to guide the development of the conceptual model using Squires et al.'s framework, which is detailed in sections 2.1.1 and 2.1.2.

Squires et al. suggest that the development of the conceptual model should have four phases (A to D, in figure 2.1) and use an iterative approach, as demonstrated by the double headed arrows. For example, as understanding of the problem increases more stakeholders may be identified, and as the model structure develops the problem may become better understood.

This chapter discusses each phase in figure 2.1 before describing the structure of the final model.

2.1.1 Phases A and B, aligning the framework with the decision making process and identifying relevant stakeholders

Several steps were taken to align Squires et al.'s framework with the decision making process underlying this thesis' aim of comparing the cost-effectiveness of different public health policies. Firstly, a three-year research plan was developed as part of the project proposal and shared with the project funders and DPhil supervisors. This ensured that the project team was happy that the proposed scope of the DPhil was realistic within the timeframe and resources available, and that there was consensus that it would remain relevant and timely to decision makers whilst in progress. Secondly, stakeholders were identified and consulted both at the start of the project when choosing which interventions to test and throughout the development of the conceptual model.

Project stakeholders were identified and categorised using a stakeholder analysis (methodology from Mind Tools).[127] Identified stakeholders were divided into five main categories:

i. national governmental organisations;

ii. local governmental organisations;

iii. charitable organisations;

iv. health professional and academic organisations;

v. patients and public.

Stakeholders within each category were selected based on either having an interest in the project or influence (power) on its outcomes and their application. In total, 25 stakeholder groups and organisations were identified (table 2.1) and arranged in a stakeholder analysis grid according to their potential level of interest and level of power (figure 2.2).

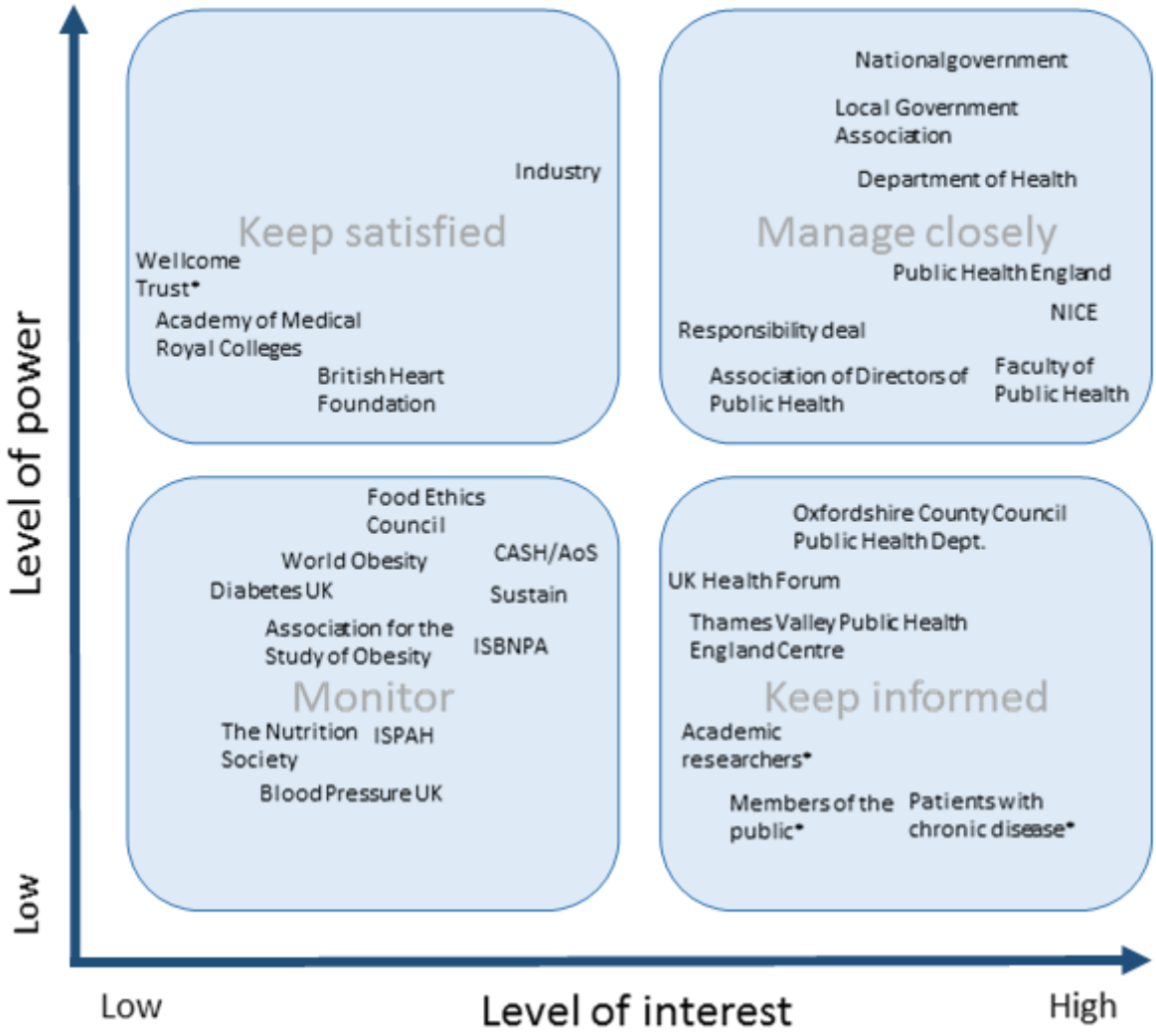
Table 2.1 List of stakeholders

Stakeholder category	**Stakeholder**
(i) National governmental organisations	National government*†
	Department of Health*†
	National Institute for Health and Care Excellence*†
	Public Health England†
	Physical activity and diet responsibility deal†
(ii) Local governmental organisations	Local Government Association†
	Thames Valley Public Health England Centre†
	Oxfordshire County Council Public Health Department†
(iii) Charitable organisations	The Wellcome Trust (project funder)
	UK Health Forum*†
	British Heart Foundation*†
	Food Ethics Council*†
	Consensus Action on Salt and Hypertension / Action on Sugar*†
	World Obesity Federation*†
	Diabetes UK*†
	Sustain*†
	Blood Pressure UK*†
(iv) Health professional and academic organisations	Association of Directors of Public Health†
	Academy of Medical Royal Colleges*†

	Faculty of Public Health*†
	International Society for Physical Activity and Health†
	International Society of Behavioural Nutrition and Physical Activity†
	The Nutrition Society*†
	Association for the Study of Obesity*†
(v) Patients and the public	Members of the public*
	Patients with chronic disease*

*contacted to identify scenarios to test; †contacted to feedback on model structure

Figure 2.2 Stakeholder analysis grid, organising stakeholders by their potential level of interest and level of power



Asterisk indicates that the stakeholder was not asked for feedback on the model's structure. CASH/AoS, Consensus Action on Salt and Health/Action on Sugar; ISBNPA, International Society of Behavioural Nutrition and Physical Activity; ISPAH, International Society for Physical Activity and Health; NICE, National Institute for Health and Care Excellence

When defining the model scope and its boundary, a balance was required between the breadth of the model's scope and its complexity on one hand, and ensuring the model would be completed within the timeframe and resources available on the other. Therefore, in discussion with stakeholders and DPhil supervisors, diet and physical activity were chosen as the two risk factors to include because of their significant attributable burden of disease in England, and because of the current relative lack of cost-effectiveness tools available to

policy makers aiming to reduce this burden. As discussed in chapter one, unhealthy diets and physical inactivity are estimated to account for a third of all attributable disease burden in England.[1] The other two leading behavioural risk factors are tobacco and alcohol misuse; it was decided to not include these in the model at this stage because there is already an internationally agreed suite of tobacco control policies that the UK has signed up to,[128] and the bespoke Sheffield Alcohol Policy Model exists to understand impact of UK alcohol policy to different sectors.[129] Increasing access to medication as part of a primary prevention strategy was also considered for inclusion in the model at the initial stakeholder engagement, however this was subsequently removed from the model's scope because of time and resource limitations. There would still be value in including tobacco control, alcohol misuse, and increasing access to medications in a comparable cost-effectiveness model, and this will be a focus of future work (see section 7.3.1).

At the start of the project, selected stakeholders (those with an asterisk next to them in table 2.1) were asked to identify which intervention they would prefer to be evaluated from a selection of four interventions aimed at improving diet and four aimed at increasing levels of physical activity (box 2.1). Patients and the public were identified through the Oxford University Hospitals NHS Trust Cardiovascular Patient and Public Involvement Panel (CVPPIP).

Box 2.1 Intervention options given to stakeholders

Physical activity

1. Free access to swimming pools and/or leisure centres
2. Employees given time to exercise during their working day
3. Physical activity advice given to all adults by their GP
4. Prescriptions for exercise available to all adults who are overweight

Diet

1. Taxing or subsidising food to help people choose healthier options (such as taxing sugar-sweetened drinks or subsidising fruit and vegetables)
2. Food producers changing food to be healthier (such as reducing salt, sugar, and unhealthy fat content)
3. Limiting serving size of unhealthy food and drink (such as limiting the size a soft-drink that can be bought at restaurants)
4. Providing free fruit and vegetables to those who are worst off

The eight policy options presented to stakeholders (box 2.1) were selected based on there being data available with which to parameterise the relationship between the intervention and the change in risk factor (diet or level of physical activity), and on the intervention either being recommended by NICE public health guidance or of stated importance to the UK government in 2014. Participants were asked to prioritise one intervention from each group of four.

Ten out of a possible 60 CVPPIP participants responded, as well as five charities, two national governmental organisations, and two professional organisations.

The two selected interventions were:

- free access to swimming pools and/or leisure centres;
- food producers changing food to be healthier.

Increasing access to medication as part of a primary prevention strategy was still included as a possible intervention at this stage and therefore, using the two selected interventions above as well as increasing prescriptions of beta-blockers to prevent CVD as proposed interventions, the initial conceptual model was developed.

2.1.2 Phases C and D, understanding the problem and developing the model structure

To better understand the problem, the initial conceptual model included all hypothesised relationships between the proposed interventions and possible outcomes. This detailed approach helps to document the problem and aims to ensure that wider intended and unintended consequences are included before practical decisions are made as to which aspects should be built into the model and which should be left out. It also means that the model is generalisable and flexible enough to simulate other interventions and changes to other behavioural risk factors that may subsequently be simulated. The initial conceptual model is shown in figure 2.3.

Figure 2.3a Detailed version of the initial conceptual model

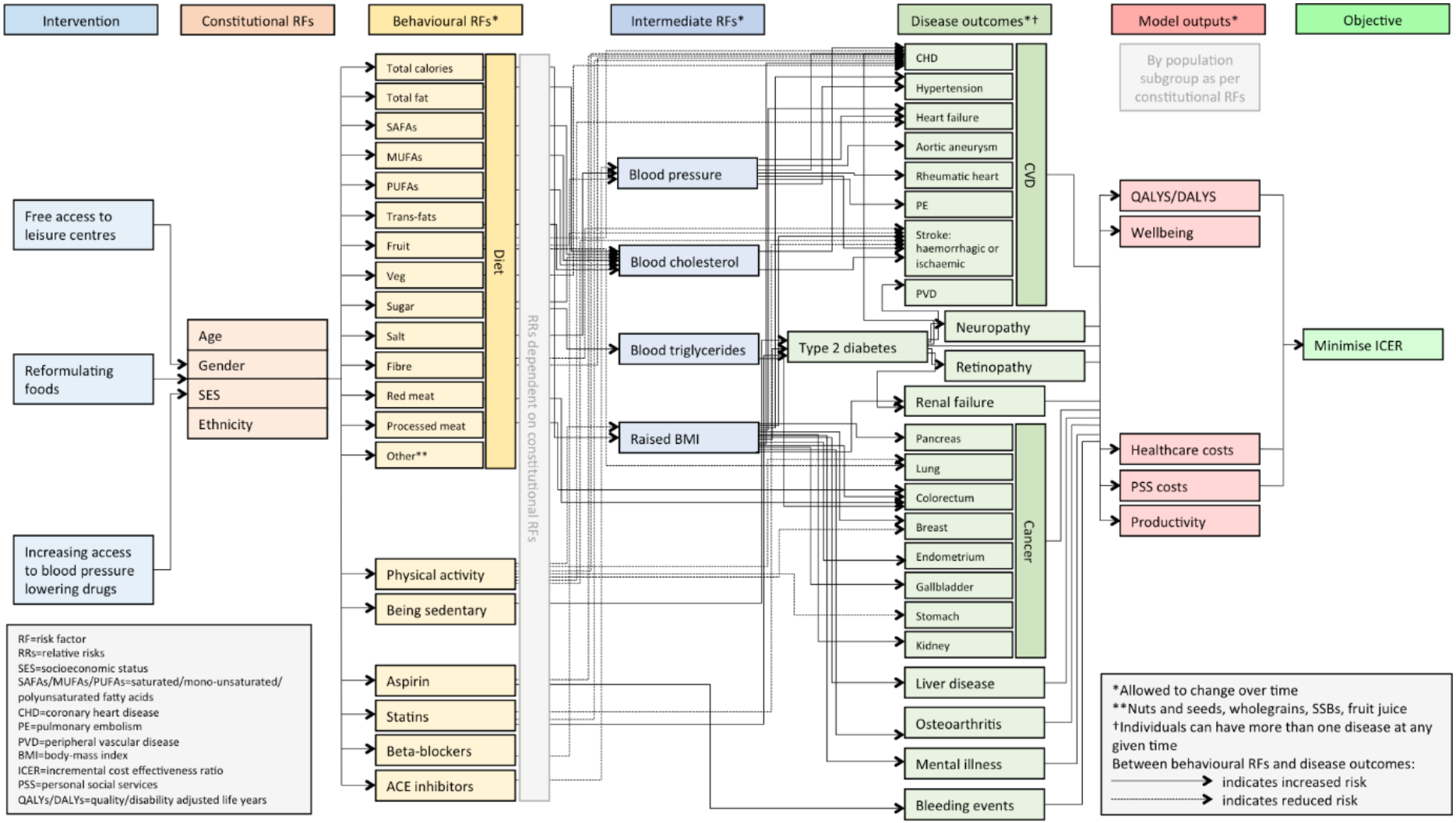
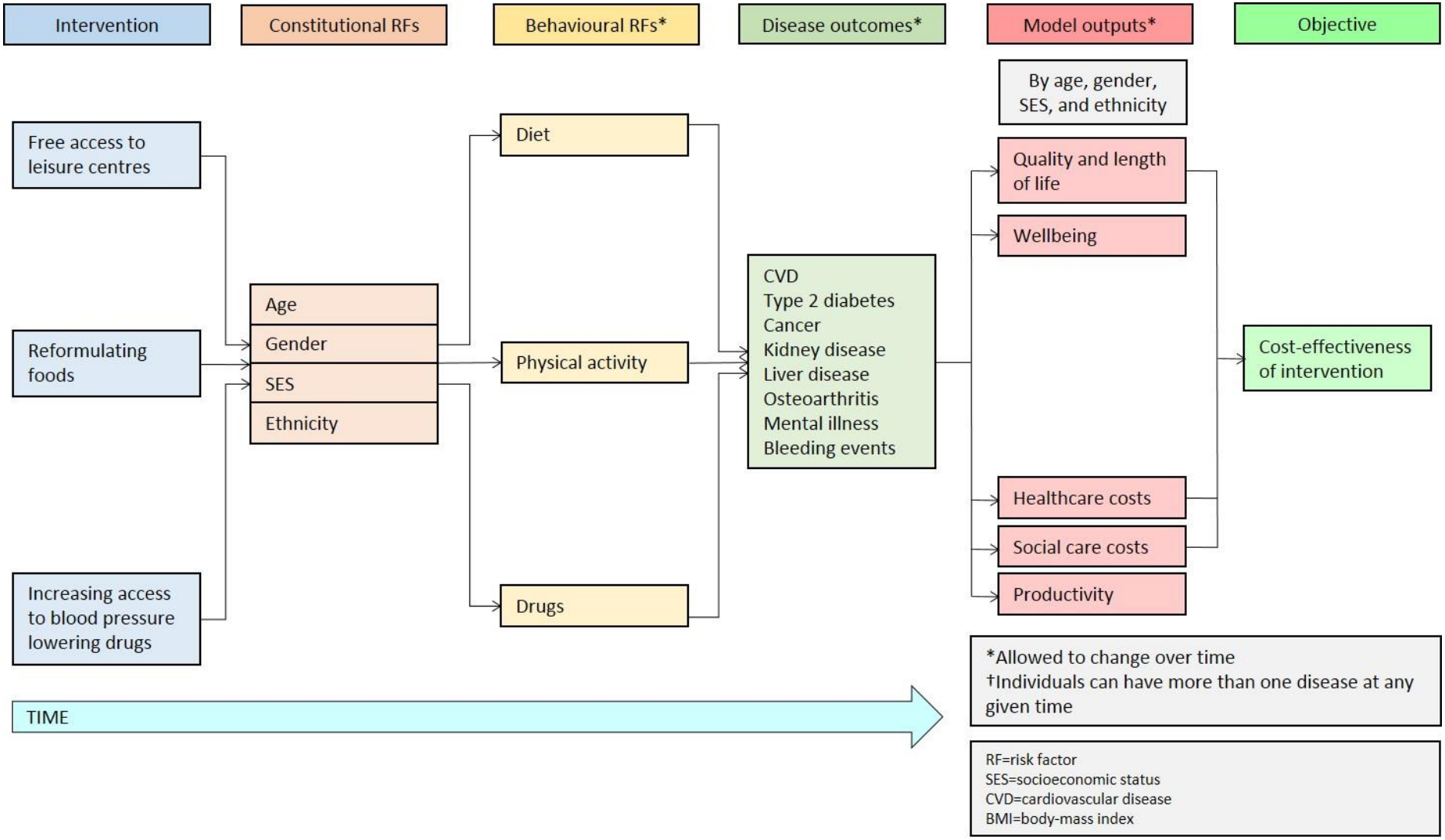


Figure 2.3b Simplified version of the initial conceptual model



A protocol for stakeholder involvement was written (see appendix one) and in accordance with this protocol, the simplified initial conceptual model (figure 2.3b) was shared with stakeholders for feedback on the model's proposed outcomes and its face validity (ethics approval reference: MS-IDREC-C1-2015-052, see appendix two for the ethics approval letter). Stakeholders were asked whether they think that the steps in the model made sense and whether any crucial steps were missing, see appendix three for the full list of questions and responses. Twelve of a possible 29 stakeholders responded with nine saying that the logic of the model is plausible and valuable, and all 12 saying it is a model that they would potentially use.

Of the three respondents who replied that they did not think the model's logic was valid and plausible, one felt that the range of interventions was too limited and should be expanded, one queried how productivity would be estimated, and one suggested that social value outcomes should be added including losses to earnings. In response to these concerns, it is intended that the range of interventions will be expanded as part of future work (see section 7.3.1), and estimating the effect on productivity and wider societal costs is discussed in more detail in section 3.1. Quantifying social value such as societal wellbeing or using a capability approach is part of wider debate about how best to evaluate public health interventions.[16,17,130] This is discussed in more detail in section 7.3.2 but based on stakeholder responses discussed below and on what can be practically achieved within the scope of this thesis, it was decided not to include measures of wellbeing in the model.

Following stakeholder feedback (appendix three), the model's primary outcome is cost-effectiveness from a health and social care perspective over a 10-year time horizon; the most selected diseases to include were heart disease, diabetes, and stroke. Less important to stakeholders were reporting results by population subgroup, estimating the impact of interventions on inequalities, and including a measure of population wellbeing or productivity. Also included in the primary analyses are all costs of the intervention, whether they are incurred by government organisations such as the NHS and local authorities, or by

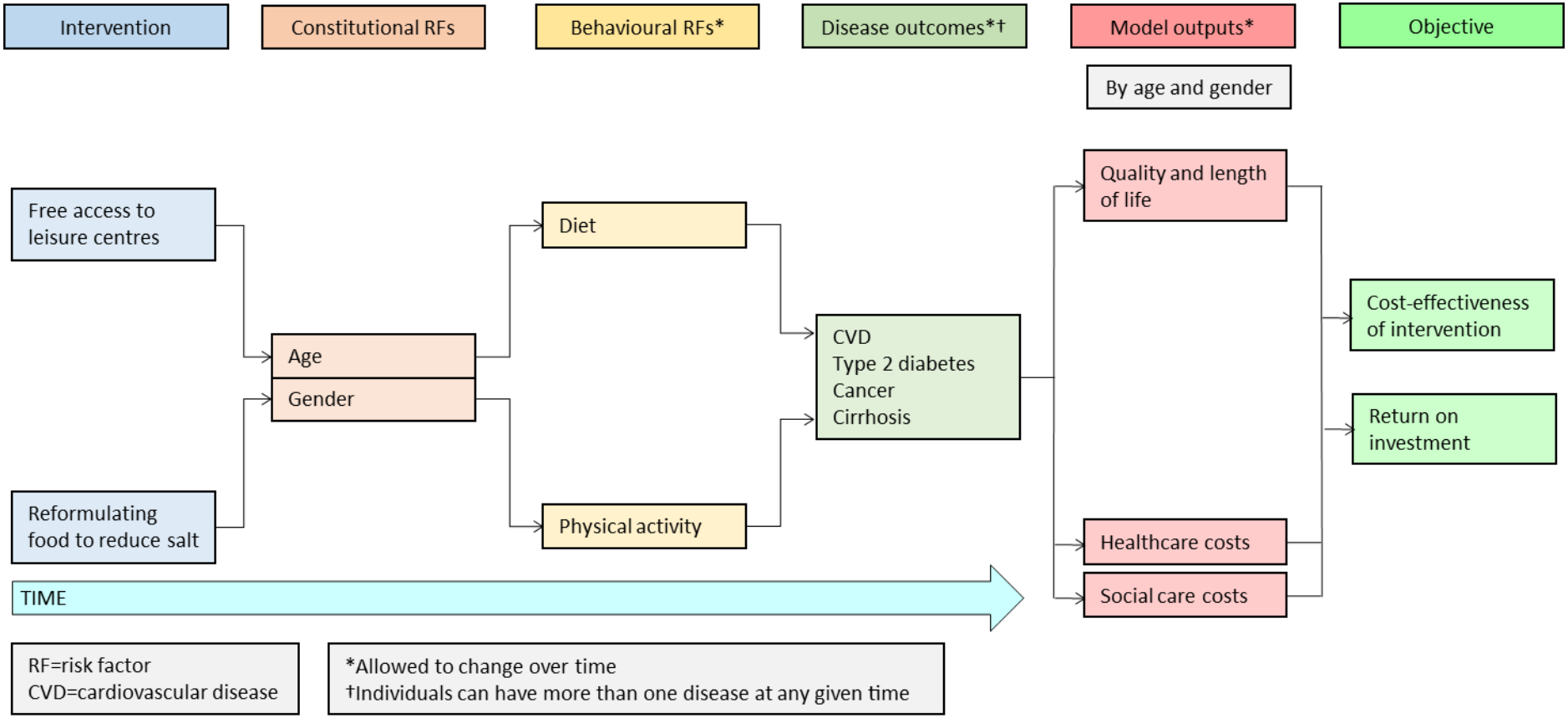
industry where appropriate, as these costs may be relevant to decision makers (industry costs are removed in sensitivity analyses, see appendix five). Stakeholder responses were used to help determine the model's boundary and level of detail. In discussion with DPhil supervisors, it was decided not to include estimates of the effect of medications on health or to simulate outcomes by SES and ethnicity in order to ensure the model can be delivered within the resources and timeframe available. Finally, the number of simulated diseases was limited to IHD, stroke, type two diabetes, and cancers, as prioritised by stakeholders, as well as liver cirrhosis for reasons explained below. Although mental health was also ranked relatively highly by stakeholders, it was not included in the model because of the lack of meta-analyses of randomised controlled trials or prospective observational studies describing the dose-response relationship between incidence of mental illness and diet or physical activity (the evidence threshold for inclusion in the model).

The review of public health economic models in chapter one illustrated that the majority of published public health economic models are population level Markov models but these rapidly become complex when modelling multiple diseases as each added disease leads to an exponential increase in the number of modelled disease states. Other cohort modelling structures such as multistate life table models can simulate the effects of interventions on multiple diseases without dramatically increasing model complexity, but have other limitations such as assuming that included diseases are independent of each other (see section 2.2 for a discussion of this limitation).

One such multistate life table model is PRIMEtime, which estimates the effect of changing diet on morbidity and mortality from IHD, type two diabetes, stroke, cancers, and liver cirrhosis in the UK population.[131] As suggested by Squires, it is inappropriate to build a novel model if an existing model is available that can be adapted to fit a new user's purpose.[47] Therefore this thesis reports on the development of PRIMEtime to achieve the DPhil's aims. PRIMEtime is the ideal model to work with because it includes all the relevant disease outcomes (plus cirrhosis, hence its inclusion in the final conceptual model), and it

includes diet as a risk factor. Furthermore, the model is freely accessible from the model's developers and there are not any more appropriate models that are readily available to be shared and used. Throughout this thesis, the adapted model is referred to as PRIMEtime Cost Effectiveness (PRIMEtime CE). The structures of both PRIMEtime and PRIMEtime CE are discussed in more detail in section 2.2. The final conceptual model is shown in figure 2.4. This was shared with stakeholders (at time T=1, see appendix one) with further feedback from national government organisations requesting that in addition to cost-effectiveness being estimated, for interventions that are dominant within the lifetime of the modelled cohort, time to return on investment and absolute return on investment should be quantified.

Figure 2.4 The PRIMEtime CE conceptual model



2.2 Final model structure and discussion

The model used in the this thesis is PRIMEtime CE, developed from PRIMEtime, a multistate life table model described in detail in the supplementary file of Cobiac et al.'s paper estimating the health impact of achieving the diet in the new Eatwell guide.[131] Briefly, PRIMEtime combines the PRIME (Preventable Risk Integrated ModEl) CRA model[132] with multistate life table methods developed by Cobiac and colleagues in Australia and New Zealand.[60,109,114] It estimates the effect of changing population diet on the incidence of CHD, stroke, type two diabetes, cirrhosis, and seven different cancers (breast, colorectum, lung, stomach, pancreas, kidney, and liver). PRIMEtime simulates a closed adult population cohort (aged 15 and above) by single year of age and sex over the lifetime of the cohort or until individuals reach 100 years of age, with UK specific data by age and sex (where available) on baseline disease incidence, prevalence, and case-fatality rates, and on disease trends. The relationships between diet and disease are parameterised using meta-analyses of randomised controlled trials or prospective observational studies and are modelled either as direct effects or via one of three intermediate risk factors: blood pressure, BMI, and total cholesterol. PRIMEtime is built in Microsoft Excel and uses Ersatz and EpiGearXL add-ins from EpiGear International to run Monte Carlo probabilistic sensitivity analyses and quantify the uncertainty in model input data.[133–135] The 2.5th and 97.5th percentiles of 2,000 model runs are used to estimate 95% uncertainty intervals (UIs). The sources of input data and their uncertainty included in PRIMEtime and PRIMEtime CE are shown in appendix four.

To create PRIMEtime CE from PRIMEtime, costs and utilities were added (chapter three) and interventions were added and parameterised (chapter four). Furthermore, the effect of changing physical activity levels on health was added because although this is included in PRIME,[132] it is not included in PRIMEtime. PRIMEtime CE was also updated with functions to allow the discount rate for costs and health outcomes, and for the time horizon to vary (in the primary analyses, costs and health outcomes are discounted at 1.5% as

recommended by NICE for interventions likely to have long-term health benefits,[13] although other guidelines on discount rates for the appraisal of UK Government policies are available[136]).

2.2.1 Limitations of using PRIMEtime and addressing the specific challenges of public health economic modelling

As discussed in chapter one, a limitation of multistate life table models is that diseases are assumed to be independent meaning that changing the prevalence of one disease does not affect the incidence of another. For example, the model cannot distinguish between whether the population with cancer is the same as that with heart disease. Multistate life table models may therefore over- or under-estimate utilities and costs due to differences occurring between two diseases existing in two separate individuals or being co-morbid within the same individual. This limitation can partly be addressed by allowing the incidence of one disease to vary based on the prevalence of another if the independent relationship between the two diseases is known. This technique is used in PRIMEtime to allow the incidence of stroke and IHD to change with the prevalence of type two diabetes. Furthermore, for risk factors such as physical activity and BMI that increase the incidence of both diabetes and CVD, over-estimating subsequent effect sizes can be avoided by calculating and using the relative risk of CVD attributable to the risk factor independent of diabetes. The overall effect of the risk factor on CVD is then the sum of the direct effect of the risk factor on CVD plus the effect mediated by diabetes. This adjustment is already made in PRIMEtime for BMI[131] and was added for physical activity using results from Cobiac et al., described in more detail in section 4.2.1.[137]

In PRIMEtime, changing a risk factor only impacts on disease incidence rates with case fatality rates assumed to remain constant. If case fatality rates were also reduced following a change in the burden of a given risk factor (for example, there is evidence that physical activity improves survival rates for some cancers[138]), then health and economic outcomes would be under-estimated.

When developing the conceptual model and engaging stakeholders, a systems approach was not taken. This was a pragmatic decision based on the time and resources available to complete this thesis. Instead, the initial conceptual model in figure 2.3 is a linear illustration of the possible impact of the interventions on health outcomes, without any potential feedback loops. Furthermore, representatives from industry were omitted from the stakeholder analysis and would have potentially had useful contributions to make to the model's design. If as a consequence there are key industry-relevant features left out of the model, industry may be less likely to accept modelled outcomes or engage with the simulated public health policies.

When considering the specific challenges of public health economic modelling, multistate life table models are well suited to modelling long terms health impacts and PRIMEtime CE can model outcomes over a population's lifetime. Furthermore, a health and social care perspective is taken, thereby incorporating some wider societal consequences of the interventions modelled (see section 3.1). Possible wider impacts of the interventions on non-health outcomes such as public attitudes to exercise or the environmental impact of people being more active was not included. The potential effect on model results of valuing these wider societal impacts is discussed in more detail in chapter seven but was beyond the scope of this thesis. A limitation of multistate life table models (and cohort models more generally) is that they are less flexible than microsimulation models at modelling heterogeneous populations (see chapter one). In order to obtain results by population subgroup and assess the impact of the modelled interventions on inequalities, cohort models need to use multiple runs, each with a different population subgroup and using subgroup specific model parameters. The relevance of this limitation to this thesis is reduced because socioeconomic group and ethnicity are not included within the model boundary, thereby reducing the number of population subgroups modelled. Modelling multicomponent interventions is possible using PRIMEtime CE, aided by the ease of including multiple disease outcomes in multistate life table models. However, PRIMEtime CE is limited in that interactions between populations and between a population and the

environment cannot be estimated. Any such interactions are beyond the model boundary and are not quantified.

2.2.2 Conclusion

As requested by stakeholders, the primary outcome of the model will be the cost effectiveness and return on investment of each intervention from a health and social care perspective over a 10-year time horizon, including costs to industry where appropriate (see section 2.1.2 and table 2.2). Sensitivity analyses will explore the effects on cost-effectiveness of changing PRIMEtime CE's principle assumptions including time horizon and discount rate, and removing industry costs (see appendix five for the list of sensitivity analyses).

Table 2.2 Values and settings used for PRIMEtime CE primary analyses

Variable	**Value or setting**
Annual discount rate for health outcomes	1.5%
Annual discount rate for costs	1.5%
Economic perspective	NHS England and social care costs for both modelled and unrelated diseases
Intervention costs	Costs to government and to industry where appropriate
Time horizon	10 years

Chapter three details how the model's costs and utilities are derived before chapter four describes how the interventions are modelled.

Chapter 3. Disease costs and utilities.

3.1 Disease costs

3.1.1 Background

Disease specific health and social care costs can be calculated using one of two different approaches: adding together unit costs of all components of a patient's care (a micro-level approach, from here on called a bottom-up approach), or allocating an overall healthcare budget to specific diseases (a macro-level or top-down approach).

Guidelines exist for how to estimate costs for health economic cost-effectiveness models, however these do not provide specific advice on how to identify costs when evaluating the impact on multiple diseases simultaneously where the comparability of data sources for different diseases is important.[13,139] This is a particularly relevant problem to public health economic models where an intervention may affect many diseases. For example, ISPOR-SMDM Good Practices for Outcomes Research include guidelines for estimating drug costs but not the costs of simulating multiple diseases.[139] By contrast, a range of recommendations for public health economic evaluations are provided by NICE. For example, evaluations should take place from a public perspective (including NHS, central and local government organisations, and other societal perspectives as appropriate), all health effects should be included, non-health benefits should be included where appropriate, and costs and health outcomes should be discounted at either 3.5% or 1.5% depending on whether it is likely that there will be long term health outcomes.[13] However, although NICE suggest a bottom-up approach is appropriate for estimating resource use, there is no guidance on how to quantify comparable costs across multiple diseases. A comparable approach is important given the variation in costs that can result from using a mixed approach.[140–143] Examples of how different researchers have estimated disease-specific costs (taken from the review of public health economic modelling studies in chapter one) are shown in table 3.1, demonstrating the variety of methods used.

Table 3.1 Sources of UK or English healthcare costs used by studies modelling multiple disease outcomes

Study with UK data	Diseases modelled	Sources for NHS costs	Sources for non-NHS costs
Trueman and Anokye [31]	CHD; stroke; type two diabetes	Different secondary sources used for each disease	Cost of intervention to participants and providers included but no other non-NHS costs
Bolin et al.[52]	COPD; heart failure; lung cancer; stroke	Different secondary sources used for each disease	Only cost of intervention included
Frew et al.[144]	CHD, stroke, type two diabetes, breast and colorectal cancer	Different secondary sources used for each disease	Only cost of intervention included
Trueman et al.[62]	Type two diabetes; colorectal cancer; CHD	Different secondary sources used for each disease	Only cost of intervention included
Cecchini et al.[117] and Sassi et al.[116]	Cancers; stroke; CHD	Used Adam et al.[145] and the WHO-CHOICE project[146] which estimated bottom-up UK hospital costs	Only cost of intervention included
Gulliford et al.[85]	Diabetes; CHD; stroke; colorectal cancer; depression	Estimated bottom-up costs by applying health sector unit costs to healthcare utilisation for different diseases from General Practice Research Database records (now called the Clinical Practice Research Datalink)[147]	Only cost of intervention included

Divajeva et al.[115], Bhimjiyani et al.[148]	CHD; stroke; COPD; oesophageal, laryngeal, renal, lung, breast, colorectal, hepatic, endometrial, and pancreatic cancer; type two diabetes; osteoarthritis; hypertension	Estimated top-down costs using NHS England programme budgeting data[149]	Bhimjiyani et al. estimated productivity costs using Department of Health tool[150]

CHD, coronary heart disease; COPD, chronic obstructive pulmonary disease; WHO-CHOICE, WHO CHOosing Interventions that are Cost-Effective

Table 3.1 illustrates that there is little consistency in how different studies identify costs for multiple disease outcomes. This can present a problem as bottom-up costs can be estimated differently for different diseases as a result of either the inclusion or exclusion of different cost components, following patients for different lengths of time, or locally driven variations in patient journeys.

Rather than using a range of secondary sources for disease-specific healthcare costs, Gulliford et al. first extracted data from the General Practice Research Database (GPRD, now the Clinical Practice Research Datalink) to estimate healthcare utilisation.[85,147] The GPRD includes recorded data on the use of primary care, secondary care, and prescriptions. Personal Social Services Research Unit (PSSRU) unit costs for health and social care and First Data Bank Europe prescription unit costs were then used by Gulliford et al. to apply unit costs to healthcare utilisation by patients with different diseases in the GPRD records. This bottom-up approach allows for comparable costs from multiple diseases to be estimated. An advantage of this over a top-down approach is that heterogeneous population groups can be analysed in greater detail because high and low healthcare utilisation by different individuals with the same disease can be measured. However, the method is data intensive and GPRD does not record secondary care use beyond referral and admission meaning total NHS disease costs are underestimated. Furthermore, with regard to quantifying wider public expenditure, it will likely underestimate NHS costs from hospital transport, community care, end of life care, and running costs, and does not include non-NHS resources such as local government services.

Cecchini et al.[117] and Sassi et al.[116] used data compiled by Adam et al.[145] to estimate the health sector costs of a variety of diseases in multiple countries alongside intervention costs. Adam et al. estimated a set of healthcare unit costs for 49 countries as part of the WHO-CHOICE (CHOosing Interventions that are Cost-Effective) project.[146] The WHO-CHOICE project aims to estimate standardised healthcare unit costs for multiple countries, using modelling to predict costs where data are not available. This is useful for comparing the cost-

effectiveness of the same intervention in multiple countries, however the range of included unit costs are limited making it difficult to accurately quantify disease specific UK costs.

The UK Health Forum have published two reports that estimate the costs of multiple diseases using a top-down approach.[115,148] Authors used 2012/13 NHS England programme budgeting data to assign NHS expenditure to different diseases.[149] Programme budgeting data from 2012/13 categorised total primary care trust (PCT) expenditure in England into 56 disease categories and 12 care settings ranging from accident and emergency through to prevention and health promotion (tables 3.2 and 3.3). Programme budgeting data have the advantage of being freely available and including a wide range of NHS expenditure. Furthermore, this top-down approach ensures that the entire NHS England expenditure for each disease area is allocated and accounted for whereas this may not be the case when using a bottom-up approach. However, the approach used by the UK Health Forum is not applicable to programme budgeting data beyond 2012/13 as PCTs were abolished following the 2012 Health and Social Care Act with clinical commissioning groups (CCGs) subsequently taking responsibility for local NHS England expenditure.[7] This led to a change to NHS England budget allocation, with specialised services and primary care expenditure (except primary care prescribing costs) being organised nationally, and remaining budgets being the responsibility of CCGs. The UK Health Forum estimated individual disease specific costs as the proportion of total costs in a given disease category based on disease incidence or prevalence. This assumes that the per person costs of each disease within the same category are the same, which may not be the most appropriate way to disaggregate disease costs (for example, the category *2b, upper gastrointestinal cancers*, includes cancer of the oesophagus, stomach, liver, gall bladder, and pancreas, each of which may have very different associated healthcare costs). The challenges of estimating disease specific costs from 2013/14 programme budgeting data and some of the limitations of the UK Health Forum methodology are discussed in more detail in section 3.1.2.1.

Finally, there is an on-going debate about whether to include costs and outcomes from diseases unrelated to those studied, but that may arise as a result of prolonging life following a modelled intervention.[151,152] The UK models estimating the cost-effectiveness of interventions affecting multiple diseases discussed in chapter one and shown in table 3.1 do not estimate healthcare costs from unrelated diseases.[31,52,62,85,115–117,144,148] This is despite the fact that individuals or cohorts surviving as a consequence of an intervention may go on to incur health and social care costs that would have otherwise been avoided. The 2014 NICE guidelines manual (updated in 2017) recommends that all relevant costs and benefits should be included in public health economic analyses, although changes to costs and utilities resulting from these unrelated diseases are not explicitly mentioned.[13] Conversely, the 2013 NICE guide to the methods of technology appraisal states that unrelated costs should not be included.[153] Morton et al. suggest that possible reasons for not including such costs include the fact that estimating them involves making assumptions about future care, they can be practically difficult to calculate, and they are immaterial to subsequent decision making; these objections have been challenged by Morton et al. and others as their absence might lead to less "robust and defensible analyses", and "suboptimal decisions".[152,154] As such, unrelated disease costs are estimated and included in PRIMEtime CE. Internationally there are examples of cost-effectiveness analyses of public health interventions including unrelated disease costs such as in Australia[60,61] and New Zealand.[114] For example, Blakely et al. estimated the additional average health expenditure for a modelled cohort due to increased longevity following tobacco tax increases in order to more accurately represent their long term cost impact.[114]

With regard to non-health sector costs, programme budgeting data include NHS spending on social and community care but not social care expenditure outside of the NHS. Furthermore, aside from intervention costs, the majority of the studies in table 3.1 only estimated healthcare expenditure rather than costs from the public perspective more broadly, as is recommended by NICE.[13,52,62,85,116,117,144] The only study that aimed to quantify some of the wider societal costs associated with an intervention is Bhimjiyani et al. from the

UK Health Forum who used premature mortality and morbidity data to estimate productivity losses based on time away from work.[148] They did this using a tool published by the Department of Health, which aims to help decision makers estimate wider societal costs as a function of age, quality of life, and ICD-10 (International Statistical Classification of Diseases and Related Health Problems, 10th revision) code (published as an appendix to a paper in *Health Economics*).[150] The tool defines wider societal costs as non-health costs arising as a result of an intervention and it breaks these down into 10 areas, including productivity, social care, domestic work, and volunteering (see section 3.1.2.3 for more detail).

This chapter aims to identify a consistent approach to estimating comparable disease specific costs and unrelated disease costs per person from an NHS England perspective and a social care perspective. Decisions regarding data sources are driven by this need for comparability. The analysis has three principal components: estimating disease specific NHS England expenditure, estimating per person NHS England expenditure for unrelated diseases, and estimating social care costs.

3.1.2 Methods

3.1.2.1 Estimating disease specific NHS England expenditure in PRIMEtime CE

This section describes how disease specific healthcare costs were calculated, first listing the data sources before outlining the methods. Total healthcare costs in PRIMEtime CE are the sum of annual excess costs from diseases related to diet and physical activity and annual baseline costs per person by age and sex from unrelated diseases.

Data sources - NHS England programme budgeting data

Health sector costs are derived from 2013/14 programme budgeting data, which reports individual CCGs expenditure in England.[155] As well as being based on CCG expenditure rather than PCT expenditure, further programme budgeting data differences between 2013/14 and 2012/13 include how costs are divided by disease category and care setting (tables 3.2 and 3.3).

Table 3.2 Programme categories reported by NHS England programme budgeting in 2012/13 and 2013/14[149,155]

Programme categories 2012/13	**Programme categories 2013/14**
01.Infectious diseases	01.Infectious diseases
01a.HIV and AIDS	01a.HIV and AIDS
01x.Infectious diseases (Other)	01x.Infectious diseases (Other)
02.Cancers and tumours	02.Cancers and tumours
02a.Head or neck cancers	02a.Head or neck cancers
02b.Upper gastro intestinal cancers	02b.Upper gastro intestinal cancers
02c.Lower gastro intestinal cancers	02c.Lower gastro intestinal cancers
02d.Lung cancers	02d.Lung cancers
02e.Skin cancers	02e.Skin cancers
02f.Breast cancers	02f.Breast cancers
02g.Gynaecological cancers	02g.Gynaecological cancers
02h.Urological cancers	02h.Urological cancers
02i.Haematological cancers	02i.Haematological cancers
02x.Cancers and tumours (Other)	02x.Cancers and tumours (Other)
03.Disorders of blood	03.Disorders of blood
04.Endocrine, nutritional and metabolic problems	04.Endocrine, nutritional and metabolic problems
04a.Diabetes	04a.Diabetes
04b.Endocrine	04b.Endocrine
04x.Endocrine, nutritional and metabolic problems (Other)	04x.Endocrine, nutritional and metabolic problems (Other)
05.Mental health disorders	05.Mental health disorders
5a.Substance Misuse	
5b.Organic Mental Disorders	
5c.Psychotic Disorders	
5d.Child and Adolescent Mental Health Disorders	
5x.Other Mental Health Disorders	
06.Problems of learning disability	06.Problems of learning disability
07.Neurological	07.Neurological
07a.Chronic pain	07a.Chronic pain
07x.Neurological (Other)	07x.Neurological (Other)
08.Problems of vision	08.Problems of vision

09.Problems of hearing	09.Problems of hearing
10.Problems of circulation	10.Problems of circulation
10a.Coronary heart disease	10a.Coronary heart disease
10b.Cerebrovascular disease	10b.Cerebrovascular disease
10c.Problems of rhythm	10c.Problems of rhythm
10x.Problems of circulation (Other)	10x.Problems of circulation (Other)
11.Problems of the respiratory system	11.Problems of the respiratory system
11a.Obstructive airways disease	11a.Obstructive airways disease
11b.Asthma	11b.Asthma
11x.Problems of the respiratory system (Other)	11x.Problems of the respiratory system (Other)
12.Dental problems	12.Dental problems
13.Problems of the gastro intestinal system	13.Problems of the gastro intestinal system
13a.Upper gastro intestinal system problems	13a.Upper gastro intestinal system problems
13b.Lower gastro intestinal system problems	13b.Lower gastro intestinal system problems
13c.Hepatobiliary problems	13c.Hepatobiliary problems
13x.Problems of the gastro intestinal system (Other)	13x.Problems of the gastro intestinal system (Other)
14.Problems of the skin	14.Problems of the skin
14a.Burns	14a.Burns
14x.Problems of the skin (Other)	14x.Problems of the skin (Other)
15.Problems of the musculoskeletal system	15.Problems of the musculoskeletal system
16.Problems due to trauma and injuries	16.Problems due to trauma and injuries
17.Problems of the genito urinary system	17.Problems of the genito urinary system
17a.Genital tract problems	17a.Genital tract problems
17b.Renal problems	17b.Renal problems
17c.Sexually transmitted infections	17c.Sexually transmitted infections
17x.Problems of genito urinary system (Other)	17x.Problems of genito urinary system (Other)
18.Maternity and reproductive health	18.Maternity and reproductive health
19.Conditions of neonates	19.Conditions of neonates
20.Adverse effects and poisoning	20.Adverse effects and poisoning

20a.Unintended consequences of treatment	20a.Unintended consequences of treatment
20b.Poisoning	20b.Poisoning
20c.Violence	20c.Violence
20x.Adverse effects and poisoning (Other)	20x.Adverse effects and poisoning (Other)
21.Healthy individuals	21.Healthy individuals
22.Social care needs	22.Social care needs
23.Other	23.Other
23a.Other - GMS/PMS expenditure	*23a.Primary care*
23x.Other - miscellaneous	*23d.Condition not known*
	23e.Multiple conditions
	23f.Condition data not recorded/reported
	23g.Pass through Payments
	23x.Miscellaneous Other

Italics indicate where categories differ between years; GMS, general medical services; PMS, personal medical services

Table 3.3 Care settings reported by programme budgeting in 2012/13 and 2013/14[149,155]

Care settings 2012/13	**Care settings 2013/14**
Prevention & health promotion	Primary Care Prescribing
Primary care	Unscheduled Care: Non-elective admission (PBR)
Primary prescribing	Unscheduled Care: A&E
Inpatient: Elective & Daycase	Unscheduled Care: Emergency Transport
Inpatient: Non-elective	Unscheduled Care: Other Urgent Care
Outpatient	Scheduled Care: Daycase and elective (PBR)
Other secondary care	Scheduled Care: Outpatient - (PBR & Non-PBR)
Ambulance	Unbundled Diagnostic Imaging
A&E	Unbundled/high cost: Critical Care
Community care	Unbundled/high cost: Drugs & devices
Care provided in other setting	Unbundled/high cost: Other
Non health / social care	Direct Access Diagnostic Imaging
	Community and Integrated care
	End of Life Care
	Running Costs

PBR, payment by results

In 2013/14, expenditure allocated to CCGs in England was £63.4bn out of a total NHS England budget of £95.6bn, meaning 2013/14 programme budgeting data included 66% of NHS England expenditure.[156] NHS England expenditure not included in 2013/14 programme budgeting data were:

- £25.4bn allocated to specialised services, primary care, and military/offender care;
- £1.3bn for CCG running costs;
- £0.9bn for local authorities for services benefiting health and social care;
- £1.8bn for Public Health England (PHE) for NHS England's public health responsibilities (immunisations, screening, and health visiting);
- £1.2bn surplus from PCTs and Strategic Health Authorities for future investment;
- £1bn for NHS England central health programmes, including clinical excellence awards and private finance initiative schemes; and
- £0.7bn for technical accounting adjustments.

Of these expenses, PRIMEtime CE includes specialised services and primary care costs (of note, primary care prescribing is included by disease category in 2013/14 programme budgeting data).

Data sources - specialised services

In 2013/14, specialised services accounted for £13.4bn, 14% of the total NHS England budget.[157] Specialised services are healthcare services usually located in specialised hospital trusts and are defined based on four criteria:

- the number of individuals requiring the service;
- the cost of providing the service;
- the number of potential providers;
- the financial implications of CCGs if they were required to provide the service.

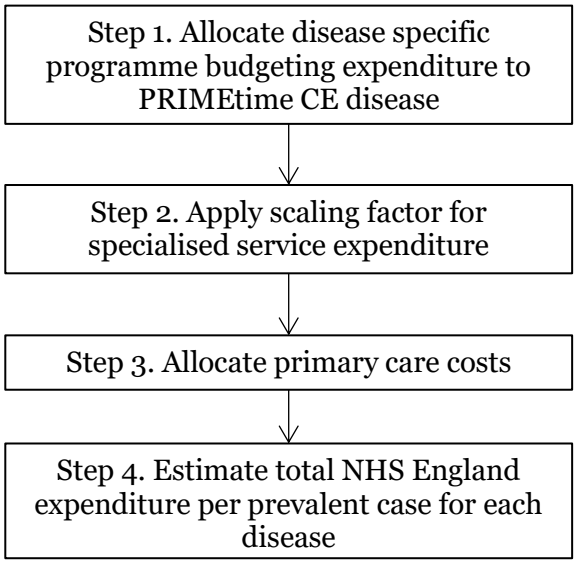
There are 143 nationally prescribed specialised services, grouped into six National Programmes of Care: Blood and Infection; Cancer; Internal medicine; Mental health; Trauma; and Women and Children.[158]

Some healthcare conditions derive a greater proportion of their spending from specialised services than others. For example, all cancer chemotherapy and radiotherapy is paid for through specialised services compared to very little treatment relating to diabetes care which is instead provided through CCGs.

Method for allocating NHS expenditure to specific diseases in PRIMEtime CE

Figure 3.1 shows the steps taken to allocate NHS England expenditure to different diseases.

Figure 3.1 Flow chart describing the steps taken to assign disease specific expenditures



Step 1. Allocate disease specific programme budgeting expenditure

Directly allocating NHS England programme budgeting expenditure to diseases included in PRIMEtime CE is not possible because programme budgeting categories do not directly match modelled diseases. PRIMEtime CE has 11 disease outcomes and the ICD-10 codes for each disease outcome are a subset of ICD-10 codes included in one or more of the 56 programme budgeting categories (table 3.4).[159]

Table 3.4 How ICD-10 codes of diseases included in PRIMEtime CE match with ICD-10 codes of programme budgeting categories

PRIMEtime CE disease outcome	**PRIMEtime CE ICD-10 codes**	**Relevant programme budgeting category**	**Programme budgeting ICD-10 codes** (ICD-10 codes that match to PRIMEtime CE in bold)
Ischaemic heart disease	I20-I25	10a Coronary Heart Disease	**I20-I25**, Z03.4, Z50.0, Z82.4, Z95.5
Stroke	I60-I69	10b Cerebrovascular disease	G46.0-G46.2, G46.8, **I60-I69**, Z82.3
Type two diabetes	E11, E14	04a Diabetes	E10, **E11**, E12, E13, **E14**, M14.2, Z13.1, Z83.3
Breast cancer	C50	02f Cancer, breast	**C50**, D05, Z12.3, Z80.3, Z85.3
Colon cancer	C18-C20	02c Cancer, lower GI	C17, **C18-C20**, C21, C26, Z12.1
Lung cancer	C34	02d Cancer, lung	C33, **C34**, C37-39, C45, Z80.1, Z80.2, Z85.1, Z85.2
Stomach cancer	C16	02b Cancer, upper GI	C15, **C16**, C22-C25, Z12.0
Liver cancer	C22	02b Cancer, upper GI	C15, C16, **C22**, C23, C24, C25, Z12.0
Kidney cancer	C64	02h Cancer, urological	C60-C63, **C64**, C65-C68, Z12.5
Pancreatic cancer	C25	02b Cancer, upper GI	C15, C16, C22, C23, C24, **C25**, Z12.0
Liver disease	K70, K74	13c Hepatobiliary	A06.4, A27.0, A95, B15-B19, B25.2, B25.2, B26.3, B58.1, B67.0, B67.5, I85, I86.4, I98.2, **K70**, K71, K72, K73, **K74**, K75, K76, K77, K80-K83, K85-K87, Q44, Q45, R16, R17, R82.2, R93.2, R94.5, Z22.5, Z52.6, Z94.4

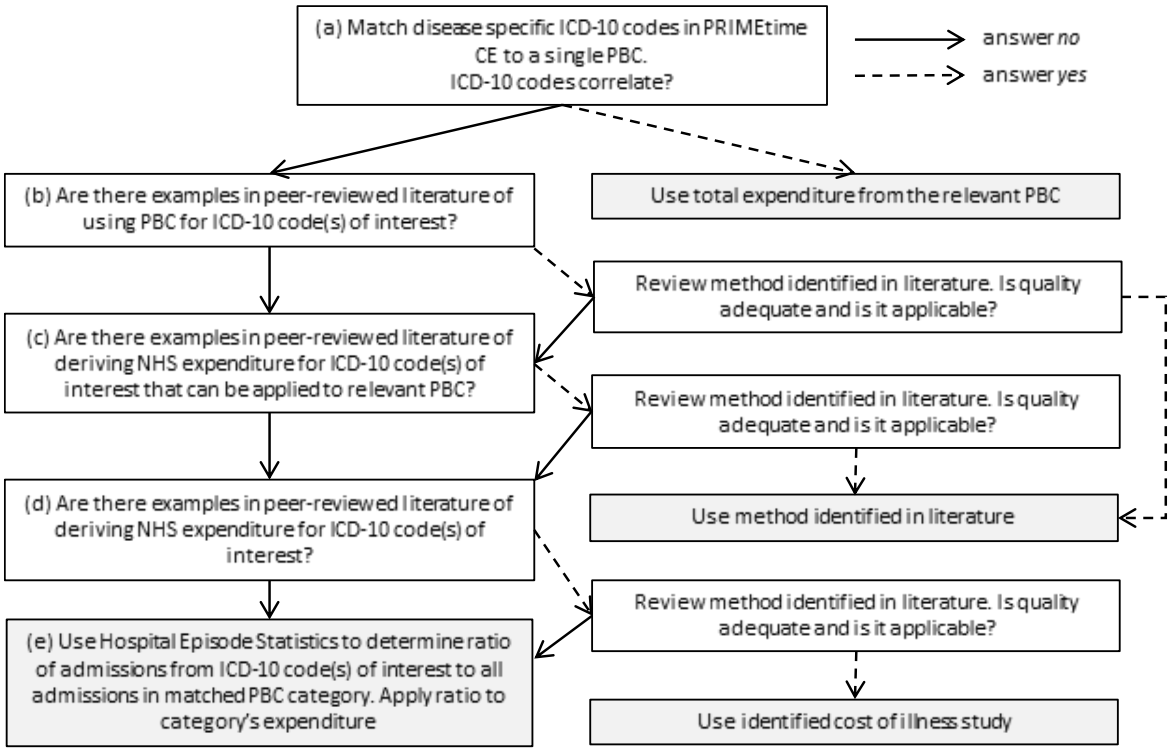
IHD, ischaemic heart disease; GI, gastrointestinal

A review of the literature using the term "programme budgeting" in MEDLINE (all years, no limits) did not identify any peer-reviewed publications that have previously divided programme budgeting category expenditure into their component diseases. The UK Health Forum divided expenditure within different programme budgeting categories in its reports

for Public Health England and Cancer Research UK using different approaches for different diseases.[115,148] Where cancer subtypes modelled by the UK Health Forum did not exactly match one of the 56 programme budgeting disease categories, the programme budgeting category which contained the modelled cancer subtype was first identified. Following this, the ratio of the incidence of the modelled cancer to total incidence of all cancers in the programme budgeting category was multiplied by the expenditure in the category (assuming average cost per patient in the programme budgeting disease category was the same across all cancers). Although the details are not provided, the UK Health Forum state that incidence or prevalence ratios were also used to estimate costs for other modelled diseases except for diabetes, where expert opinion was used as a basis for assuming that 90% of total diabetes prevalence and costs are associated with type two diabetes rather than type one diabetes.

Based on the methodology devised by the UK Health Forum and in the absence of any additional peer-reviewed literature, the flow chart in figure 3.2 was developed to govern the assignment of expenditure from programme budgeting categories to individual diseases. Costs allocated to programme budgeting categories *21, healthy individuals; 22, social care needs;* and *23, other* were not included in this analysis. These categories were responsible for 22% of 2013/14 programme budgeting expenditure but it is not possible to know to which specific diseases (if any) these costs should be allocated.

Figure 3.2 Flow chart for how to match expenditure associated with a programme budgeting category to ICD-10 disease codes found in PRIMEtime CE



PBC, programme budgeting category

Part (a)

Part (a) asks whether the ICD-10 codes for a disease in PRIMEtime CE match with a programme budgeting category. If so, the total expenditure in that programme budgeting category was used.

Parts (b), (c) and (d)

Both parts (b) and (c) identify published literature that directly apportions expenditure for diseases within a programme budgeting category of interest by ICD-10 code thereby allowing the proportion of expenditure from the relevant PRIMEtime CE disease to be identified. In the absence of any papers being identified in parts (b) and (c), part (d) searches for papers containing an overall estimate of NHS expenditure for relevant ICD-10 codes which can then be used in place of programme budgeting expenditure data.

The following search strategy was used for parts (b), (c), and (d) for each PRIMEtime CE disease (using liver disease as an example):

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1946 to Present

Search terms:

1. exp Liver Diseases/
2. exp Great Britain/
3. ("National Health Service" OR NHS).mp.
4. exp "Costs and Cost Analysis"/
5. 1 and (2 or 3) and 4

Retrieved papers were then appraised for whether they were relevant, applicable, and of adequate quality. There are no published critical appraisal guidelines specifically applicable to appraising health economic papers for this purpose, but related guidelines are available. Luengo-Fernandez developed a quality checklist for costing studies as part of his 2009 DPhil thesis (appendix eight of reference [160]). This provided a framework for generating a similar checklist to appraise study quality. To ensure that all relevant criteria were considered, the checklist for assessing economic evaluations in chapter three of Drummond et al.'s Methods for the Economic Evaluation of Health Care Programmes, The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement, ISPOR-SMDM Good Research Practices for Measuring Drug Costs, and the Critical Appraisal Skills Programme (CASP) appraisal tool for economic evaluations were reviewed.[28,139,161,162] The finalised quality criteria checklist used in this thesis is in appendix six.

All studies identified in parts (b), (c), and (d) were assessed against the quality criteria in appendix six. In the case of more than one relevant study being identified, the study with the highest quality score was used.

Part (e)

For PRIMEtime CE diseases yet to have their total expenditure estimated, part (e) disaggregates expenditure within programme budgeting categories based on 2013/14 admissions using hospital episode statistics (HES) data.[163] The ratio of admissions from each PRIMEtime CE disease ICD-10 code to admissions from all ICD-10 codes in the relevant programme budgeting category was first calculated. Using stomach cancer as an example, this is the number of admissions from PRIMEtime CE ICD-10 code C16 divided by total number of admissions from programme budgeting category *02b, Cancer, upper GI* ICD-10 codes C15, C16, C22-C25, and Z12.0. This was then multiplied by the programme budgeting expenditure for the entirety of category 02b.

Step 2. Specialised services expenditure

Specialised services are described in the NHS England Manual for Prescribed Specialised Services as:

"those provided in relatively few hospitals, to catchment populations of more than one million people. The number of patients accessing these services is small, and a critical mass of patients is needed in each treatment centre in order to achieve the best outcomes and maintain the clinical competence of NHS staff."[158]

Such services include cancer chemotherapy and radiotherapy, and cardiac surgery. Following discussion with NHS England, aggregated expenditure data were provided for specialised services divided into 65 disease categories for 2013/14.[157] However, there is insufficient information to directly allocate these to PRIMEtime CE disease categories in table 3.4 and data quality concerns did not permit greater disaggregation of results by NHS England.

Therefore 2012/13 programme budgeting data were used to estimate disease specific specialised services expenditure when PCTs were responsible for their commissioning. Specialised services expenditure in 2012/13 programme budgeting data are mapped to each disease under the programme budgeting care setting 'Other Secondary Care'.[164] As well as specialised services, this care setting includes expenditure on maternity and mental health

services, and COPD discharge care, allocated under their respective maternity, mental health, and respiratory disease programme budgeting categories. None of these conditions are included in PRIMEtime CE (table 3.2). Finally, patient transport services are also recorded as part of 'Other Secondary Care' and are allocated under either a disease specific programme budgeting category if known, or under category *23x, other – miscellaneous*. It was assumed that these did not contribute to a significant proportion of total disease costs.

In order to estimate 2013/14 programme budgeting disease specific expenditure on specialised services for each PRIMEtime CE disease, the ratio of 2012/13 programme budgeting expenditure on 'Other Secondary Care' to total 2012/13 programme budgeting expenditure for each relevant programme budgeting disease category was calculated. Expenditure on prevention and health promotion, and on primary care is included in 2012/13 programme budgeting data but not 2013/14, therefore these care settings were not included in the total programme budgeting expenditure for this calculation. The ratio was then multiplied by the 2013/14 expenditure calculated in step one to estimate the specialised services expenditure by disease for 2013/14, assuming that the proportion spent on specialised services for each disease was unchanged between 2012/13 and 2013/14.

For cancers, the majority of 'Other Secondary Care' expenditure for chemotherapy and radiotherapy was allocated to category *02x, cancers and tumours, other* rather than to the specific cancer subtype.[164] Calculating the ratio of 'Other Secondary Care' expenditure to total 2012/13 programme budgeting expenditure for each specific cancer would significantly underestimate specialised services cancer expenditure. Instead, the ratio used for all PRIMEtime CE cancer subtypes was calculated based on the whole of category *02, cancers and tumours* rather than the cancer subtype programme budgeting category.

Step 3. Allocating primary care costs

Expenditure by NHS England on primary care in 2013/14 was £11.3bn.[157] Total primary care expenditure is not included in 2013/14 programme budgeting data and is not available by disease in 2012/13. However, primary care prescribing expenditure data by disease is

reported and this was used to estimate disease specific total primary care spend in 2013/14. For each programme budgeting category relevant to PRIMEtime CE in table 3.4, the total primary care prescribing expenditure was first multiplied by the proportion of expenditure related to the PRIMEtime CE disease within the category calculated in step one. The result was then divided by total primary care prescribing expenditure across all diseases and multiplied by £11.3bn, the total 2013/14 primary care expenditure. This figure was added to total costs from steps one and two to give the final disease specific expenditure for NHS England in 2013/14.

Step 4. Estimating expenditure per prevalent case

For each disease in PRIMEtime CE, total costs estimated following step three were divided by the 2014 disease prevalence (as estimated by PRIMEtime CE, see appendix four for data sources) to estimate the 2013/14 cost per prevalent case in England. In reality, an individual's disease costs will vary depending on whether someone has been recently diagnosed with a condition, is a long-term patient, or is coming to the end of their life. Top-down expenditure estimates from programme-budgeting data estimate the total population expenditure for each disease, including diagnostic, emergency, and end of life care. Therefore the estimated cost per prevalent case is an average across all individuals with the disease, irrespective of time since diagnosis. The cost per case is used in PRIMEtime CE for all those with the disease to estimate total disease costs as the proportion of the population with the disease changes over time following an intervention.

3.1.2.2 Estimating NHS England expenditure per person for unrelated diseases

Expenditure per person for all ICD-10 codes other than those included in PRIMEtime CE (unrelated disease costs) were estimated using NHS England cost curves.[165] NHS England cost curves describe the relative health expenditure by age group and by sex for four care categories – general and acute care, mental health, prescribing, and primary care. The calculation of age-cost curves are a by-product of NHS England's resource allocation formulae for CCGs, and because of this specialised services and maternity expenditure are

not used in their calculations. For acute care, mental health, and prescribing, ratios of expenditure compared to the most expensive age and sex group were directly calculated from the published curves, whereas this was not possible for primary care where cost-curves report the average time in minutes per year spent with a GP compared to a 0-4 year old male (set to 0 minutes). To estimate primary care ratios, the baseline length of time spent in primary care for a male aged 0-5 years was taken from the NHS England primary medical care allocation formula and the proportional difference in minutes spent with a GP was calculated for each age and sex group.[166]

In order to estimate total age and sex specific NHS England expenditure on unrelated diseases, the total NHS England expenditure on non-PRIMEtime CE ICD-10 codes was first calculated by subtracting the total expenditure on PRIMEtime CE ICD-10 codes estimated in step four from NHS England's budget for clinical services. The remaining NHS England expenditure, except for maternity services, was then divided into the four care categories for which cost-curves are available. Allocation of remaining programme budgeting expenditure (£46.6bn), specialised services expenditure (£13.0bn), and primary care costs (£9.8bn) were made separately. Reported NHS England costs that are not disease specific such as running costs, surplus, and PHE costs, as well as costs from programme budgeting categories 21 (healthy individuals), 22 (social care needs), and 23 (other) were not included.

To allocate programme budgeting expenditure on unrelated diseases to the four cost curves, programme budgeting categories *5, mental health disorders* and *6, problems of learning disability* were allocated to mental health, and all prescribing costs (programme budgeting care settings, *primary care prescribing* and *unbundled/high cost: drugs and devices*) were allocated to prescribing. Remaining programme budgeting expenditure, except for category *18, maternity and reproductive health*, was allocated to general and acute care. Specialised services expenditure was allocated to general and acute care except for the proportion spent on the mental health national programme of care or explicitly spent on prescribing and maternal health services using the specialised services expenditure data provided by NHS

England.[157] The £9.8bn spent on primary care was allocated to primary care except for a proportion allocated to mental health, calculated using the method in step three, section 3.1.2.1. Finally, programme budgeting expenditure on category *18, maternity and reproductive health* and maternal health specialised services expenditure was allocated based on the proportion of total 2014 live births by mothers' age in England and Wales.[167]

Total unrelated NHS England expenditure by age and sex was the sum of the costs allocated according to each cost curve described above (including maternity services) for each age and sex group. The effect on the model's results of not including NHS England unrelated disease costs are explored in a sensitivity analysis (see appendix five).

3.1.2.3 Estimating societal costs

Social care costs were estimated using the wider societal costs tool developed by the Department of Health.[150] The tool estimates the age and sex specific effect on production (paid and unpaid) and consumption following a change in quality of life (quantified using utility values measured by the EuroQol five dimensions questionnaire [EQ-5D], see section 3.2) and ICD-10 code. Consumption includes social care (described as formal care by the tool), informal care (care provided by family and friends), private paid (goods and services purchased for consumption), private unpaid (benefit from goods and services not paid for, such as domestic work), and government (services provided by the government not included in other categories). The tool uses different methods and data sources for each category of production and consumption. Consistent with stakeholder responses reported in chapter two, primary analyses include estimates of the effect of each intervention compared to baseline on social care consumption. Productivity gains and wider societal costs are included as sensitivity analyses.

Social care costs are estimated as a function of age and quality of life, in PRIMEtime CE they assume an average monthly cost of £4,826 per person (the 2013/14 monthly local authority residential care cost[168]).[150] This reflects the total societal costs of adult social care rather than the direct costs to local authorities because many people are required to fund a

proportion of their care from personal savings.[169] The 2011 Adult Social Care Survey of 65,600 adult social care users' experience of social care and its impact on their quality of life,[170] and the 2011 GP Patient Survey of 530,174 individuals' experiences of using primary care and their quality of life,[171] were used to conduct a regression analysis estimating the probability of needing social care by age and quality of life. The average cost of local authority residential care was then applied to this probability with the added demands of dementia and stroke patients reflected by the use of cost multipliers derived from a literature review of how social care use varies by disease (a multiplier of 8.4 is used for dementia and 5.9 for stroke).

The Department of Health tool also includes paid production which quantifies productivity losses and gains resulting from a change in health status. As described by Bhimjiyani et al. who used the tool,[148] it uses a human capital approach which assumes full employment and that a person will not be replaced when they stop working due to illness. The tool estimates the effect of ill health on employee productivity using the average time spent working given an individual's age and their quality of life. The analysis is based on Wave 1 of the Understanding Society dataset (years 2009 and 2010), a survey of 50,994 adults that includes questions on whether a respondent worked in the past week and a measure of quality of life.[172] Responses to these parts of the survey are then used to estimate whether a person worked in the past week based on their quality of life. The results of the regression are combined with estimates of the average number of hours worked each week and the average hourly wage by age and sex to calculate the effect of quality of life on total wages. Unpaid production is estimated based on the economic gain of time spent on providing sickness care, childcare, and doing other general unpaid work such as domestic chores. Finally, informal care, private paid and unpaid consumption, and consumption of Government services are included to estimate the cost of using goods and services either purchased or provided by the government and informal carers.

Before using the Department of Health tool, input data were updated to 2014 where possible including using the 2014 Annual Survey of Hours and Earnings to estimate mean hourly wage rates and sickness rates,[173] and obtaining 2013/14 monthly local authority residential care costs from the 2015 PSSRU unit costs for health and social care manual.[174] In PRIMEtime CE, both the change in social care costs arising as a result of changes to modelled diseases and from unrelated diseases due to increasing longevity are included. Age and sex specific utility values for quality of life at baseline and for each disease were calculated using the methods described in section 3.2.

3.1.2.4 Interventions

Intervention costs are included in all analyses and are discussed in chapter four.

3.1.2.5 Uncertainty

As healthcare costs are disaggregated from total NHS England expenditure there are no empirical estimates of their uncertainty. The Department of Health wider societal costs tool also does not include uncertainty estimates. Therefore, based on Blakely et al.'s method of quantifying uncertainty,[114] healthcare and social care cost estimates are assumed to be "moderately uncertain" and therefore each vary according to a gamma distribution based on a normal distribution with a standard deviation of +/- 10% of the mean. A gamma distribution is used to ensure that costs do not become negative.

3.1.2.6 Discounting

In line with NICE guidance, all costs and health outcomes will be discounted at 1.5% with 3.5% discounting included as a sensitivity analysis.[13]

3.1.3 Results

3.1.3.1 Disease specific NHS England expenditure

Total NHS England expenditure by PRIMEtime CE disease is shown in table 3.5, with the largest categories of expenditure being type two diabetes (£2,057m) and IHD (£1,481m). Using PRIMEtime CE disease prevalence rates, the annual excess cost to NHS England per

prevalent case varies between £3,074 for pancreatic cancer and £314 for liver disease (table 3.5 and figure 3.3).

In order to allocate programme budgeting expenditure to diseases included in PRIMEtime CE (figure 3.2), part (a) was used for IHD (ICD-10 I20-I25) and stroke (I60-I69), where ICD-10 codes matched programme budgeting categories *10a, coronary heart disease* and *10b, cerebrovascular disease* respectively. There were some additional ICD-10 codes included in the programme budgeting categories 10a and 10b not included in PRIMEtime CE (table 3.4), however these were responsible for 0.01% and 0.02% of all HES admissions within categories 10a and 10b respectively in 2013/14 and were therefore disregarded.

There were no relevant papers identified for any disease of interest following part (b), whereas at part (c) a relevant paper of adequate quality was retrieved for type two diabetes. In this paper, Hex and colleagues estimated the 2010/11 total direct NHS cost of type one and type two diabetes including screening, testing, primary care, and secondary care (ICD-10 codes E10, E11, E14, and Z13.1).[175] Together these codes were responsible for 99.2% of 2013/14 HES admissions for programme category *04a, diabetes*. The ratio of direct NHS expenditure on type one diabetes to type two diabetes calculated by Hex et al. was 0.10. This ratio was applied to the 2013/14 expenditure in programme budgeting category 04a to estimate type two diabetes costs. No other relevant publications of adequate quality were identified for any other diseases of interest in either part (b) or (c).

Breast cancer was the only condition for which a relevant study of adequate quality was identified in part (d). Luengo-Fernandez and colleagues estimated the 2009 healthcare costs, informal care costs, and productivity losses resulting from breast, lung, colorectal, and prostate cancers for different countries in Europe, including the UK.[176] For breast cancer, the ICD-10 codes matched those used in PRIMEtime CE (ICD-10 C50), this was not the case for other cancers studied. Healthcare service costs included by Luengo-Fernandez et al. were those incurred from primary care, emergency care, outpatient care, hospital inpatient care, and drugs. National datasets and previous analyses were used to estimate

healthcare usage by breast cancer patients before unit costs were applied to estimate total NHS expenditure.[176] Total 2009 UK healthcare costs from Luengo-Fernandez were converted to 2013/14 English healthcare costs by initially assuming that breast cancer costs per person in the UK are the same as the cost per person in England.[177] The population adjusted costs were then converted from Euros to pounds sterling using 2009 exchange rates[178] and scaled to 2013/14 costs using the hospital and community health services (HCHS) index giving a total cost of £472m.[179]

For the remaining PRIMEtime CE diseases (colon, lung, stomach, liver, kidney, and pancreatic cancer, hypertension, liver disease, and renal disease), expenditure within the related programme budgeting categories were derived using part (e), based on 2013/14 HES admissions data.[163] The ratio of admissions from each PRIMEtime CE disease ICD-10 code to admissions from all ICD-10 codes in the relevant programme budgeting category was first calculated before being multiplied by programme budgeting expenditure on the entire category.

3.1.3.2 NHS England expenditure per person from unrelated diseases

Expenditure on unrelated diseases (those not included in PRIMEtime CE) is shown in table 3.6 by care category and category of expenditure. Annual unrelated disease costs increase with age, except for a secondary peak at 30-34 years for women reflecting maternity resource use (figure 3.4).

3.1.3.3 Societal costs

Baseline social care costs increase with age for both men and women, reaching £5,474 for women and £5,455 for men aged 100. The effect on results of including paid production (productivity) along with social care costs, and all wider societal costs as quantified by the Department of Health tool are both tested using sensitivity analyses. Net productivity for social care costs and productivity varied by age and sex with the maximum net productivity occurring at age 42 for men (£35,304) and age 39 for women (£22,201). When all wider societal costs are quantified, net productivity for men peaks at age 48 (£20,228) and women

aged 46 (£9,204). Figure 3.5 illustrates net productivity by age and sex using colorectal cancer to demonstrate how societal costs change with disease; figure 3.5a shows social care costs only, figure 3.5b shows social care costs and paid production, and figure 3.5.c shows all wider societal costs.

Table 3.5 Total 2013/14 expenditure by disease included in PRIMEtime CE (£000s except cost per prevalent case)

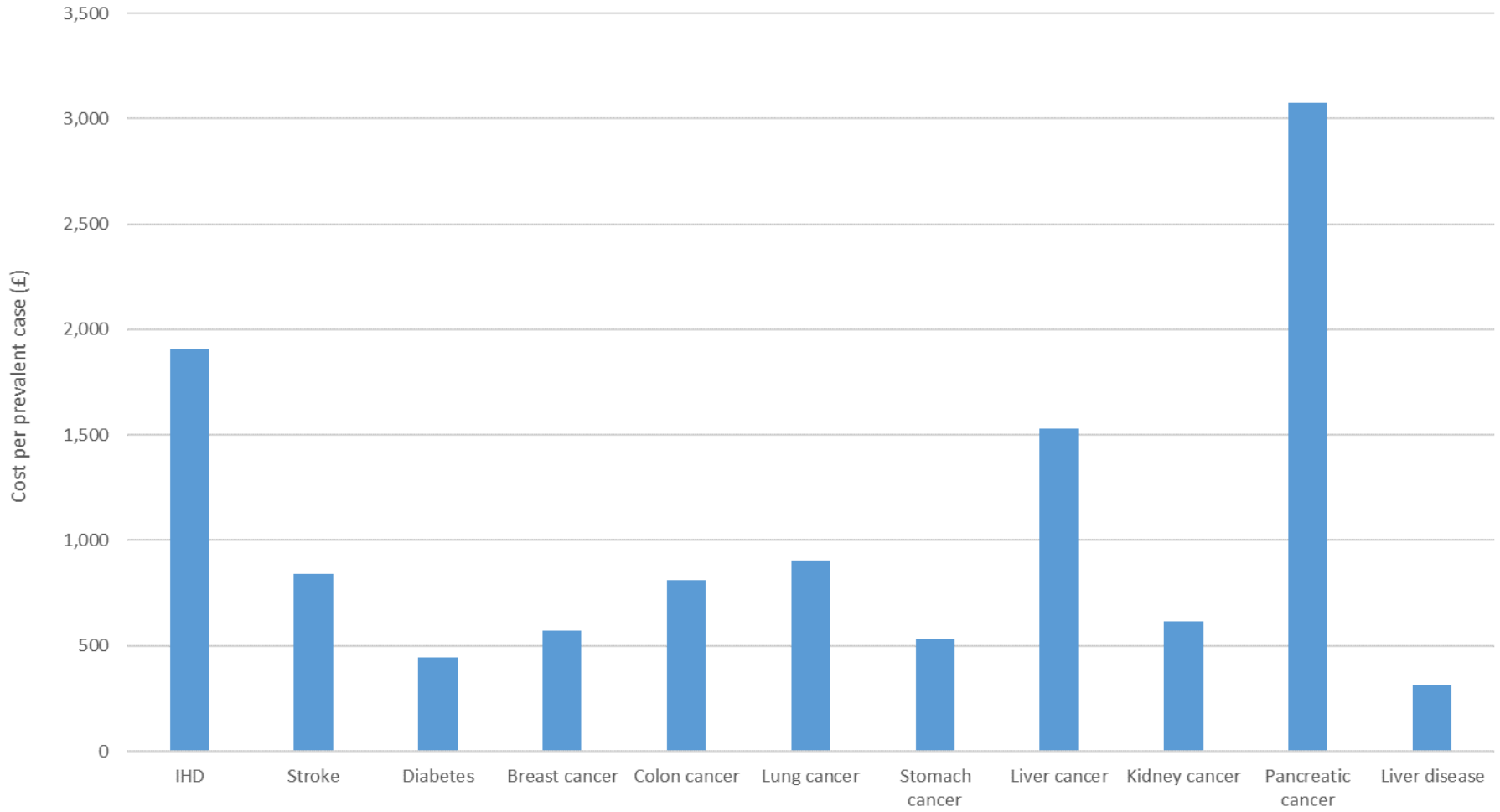
PRIMEtime CE disease	Programme budgeting category	Programme budgeting expenditure	Specialised services expenditure	Primary care expenditure	Total NHS England disease costs	PRIMEtime CE annual cost per prevalent case (£)
		Step 1	Step 2	Step 3		
Ischaemic heart disease	10a Coronary Heart Disease	953,743	41,818	485,056	1,480,617	1,905
Stroke	10b Cerebrovascular disease	689,876	55,443	29,475	774,794	843
Type two diabetes	04a Diabetes	1,071,537	25,577	959,716	2,056,831	444
Breast cancer	02f Cancer, breast	472,194*	N/A	N/A	472,192	573
Colon cancer	02c Cancer, lower GI	248,315	84,919	20	333,253	810
Lung cancer	02d Cancer, lung	98,250	33,599	0†	131,849	904
Stomach cancer	02b Cancer, upper GI	32,794	11,215	0†	44,008	535
Liver cancer	02b Cancer, upper GI	16,990	5,810	0†	22,801	1,532
Kidney cancer	02h Cancer, urological	25,145	8,599	7,833	41,577	618
Pancreatic cancer	02b Cancer, upper GI	42,133	14,409	0†	56,542	3,074
Liver disease	13c Hepatobiliary	59,702	4,543	2,963	67,209	314

*Breast cancer costs not estimated from programme budgeting expenditure but directly from Luengo-Fernandez et al.[176]. GI, gastrointestinal; N/A, not applicable; †primary care costs estimated to be negligible

Table 3.6 NHS England expenditure on unrelated diseases by care category and by category of expenditure (£000s)

	General and acute	Mental health	Prescribing	Primary care	Maternity	Total
Programme budgeting	27,583,622	8,699,759	7,647,576	0	2,660,078	46,591,034
Specialised services	10,581,726	1,764,000	629,900	0	12,700	12,988,326
Primary care	0	931,601	0	8,818,666	0	9,750,267
Total	38,165,348	11,395,360	8,277,476	8,818,666	2,672,778	69,329,627

Figure 3.3 Annual cost per prevalent case in 2013/14 for diseases included in PRIMEtime CE (£)



IHD, ischaemic heart disease

Figure 3.4 NHS England annual per person expenditure on PRIMEtime CE unrelated diseases by age and sex

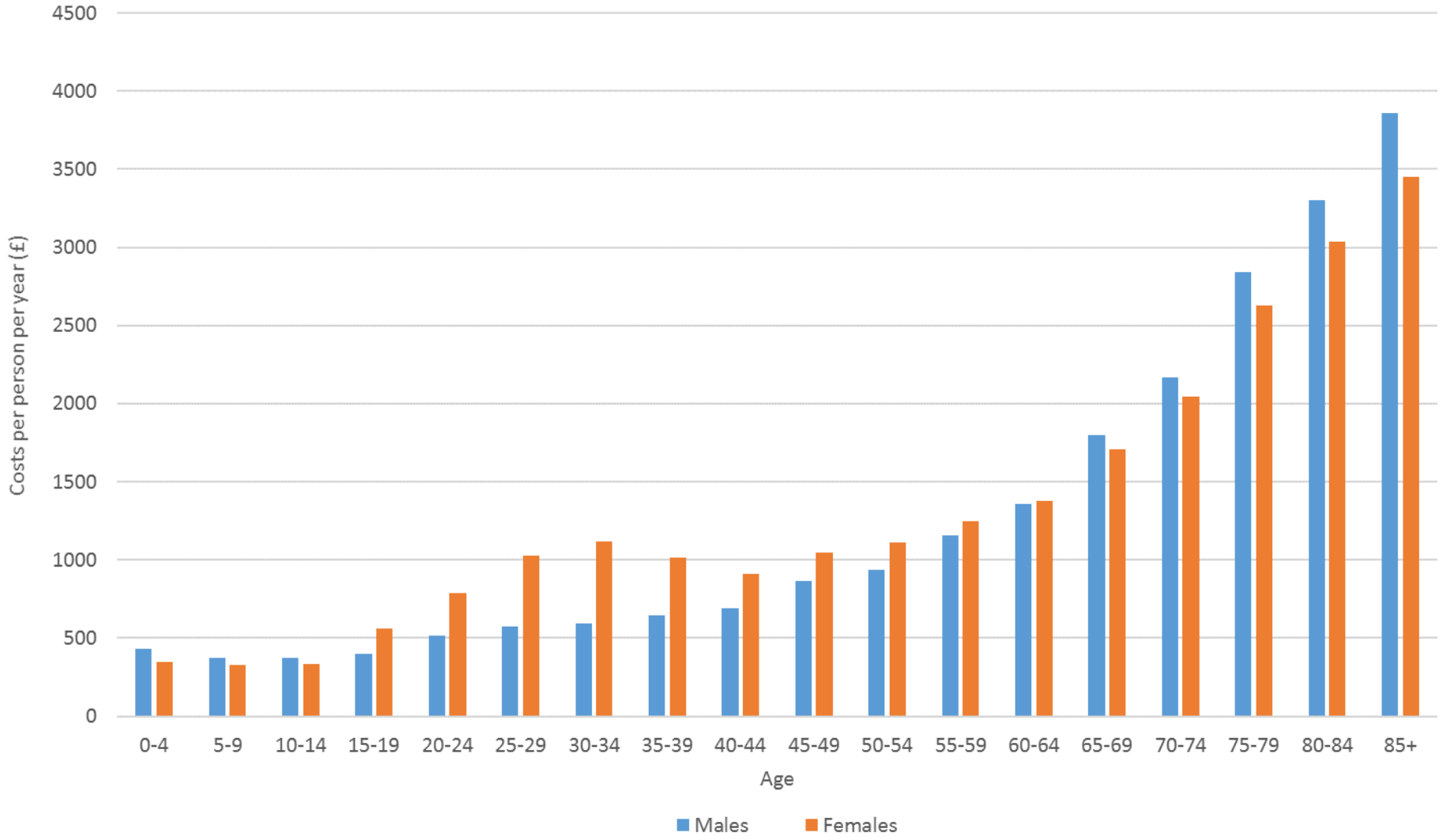


Figure 3.5a Annual social care costs by age and sex with and without a diagnosis of colorectal cancer

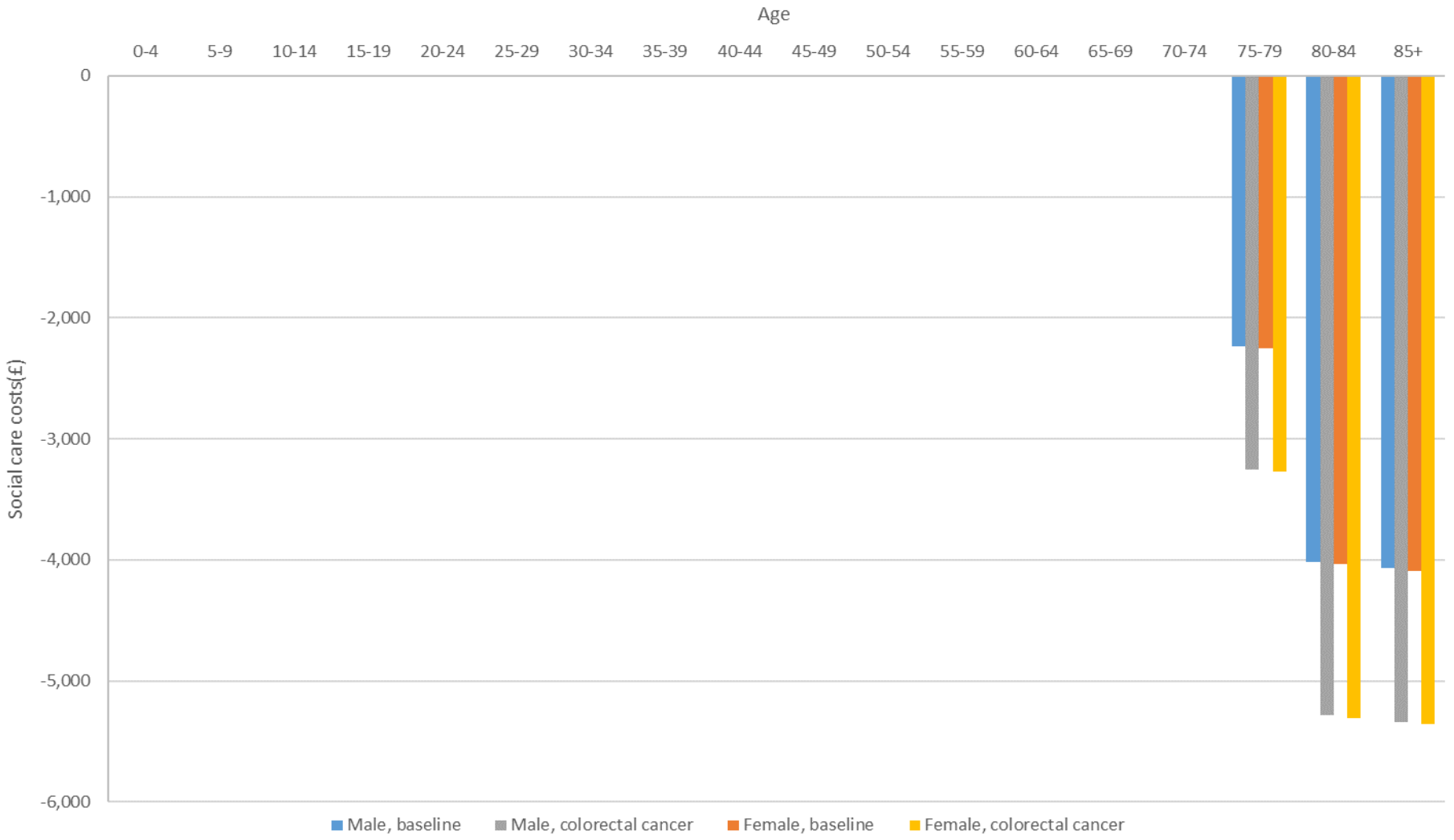


Figure 3.5b Annual net productivity from paid production and social care costs by age and sex with and without colorectal cancer

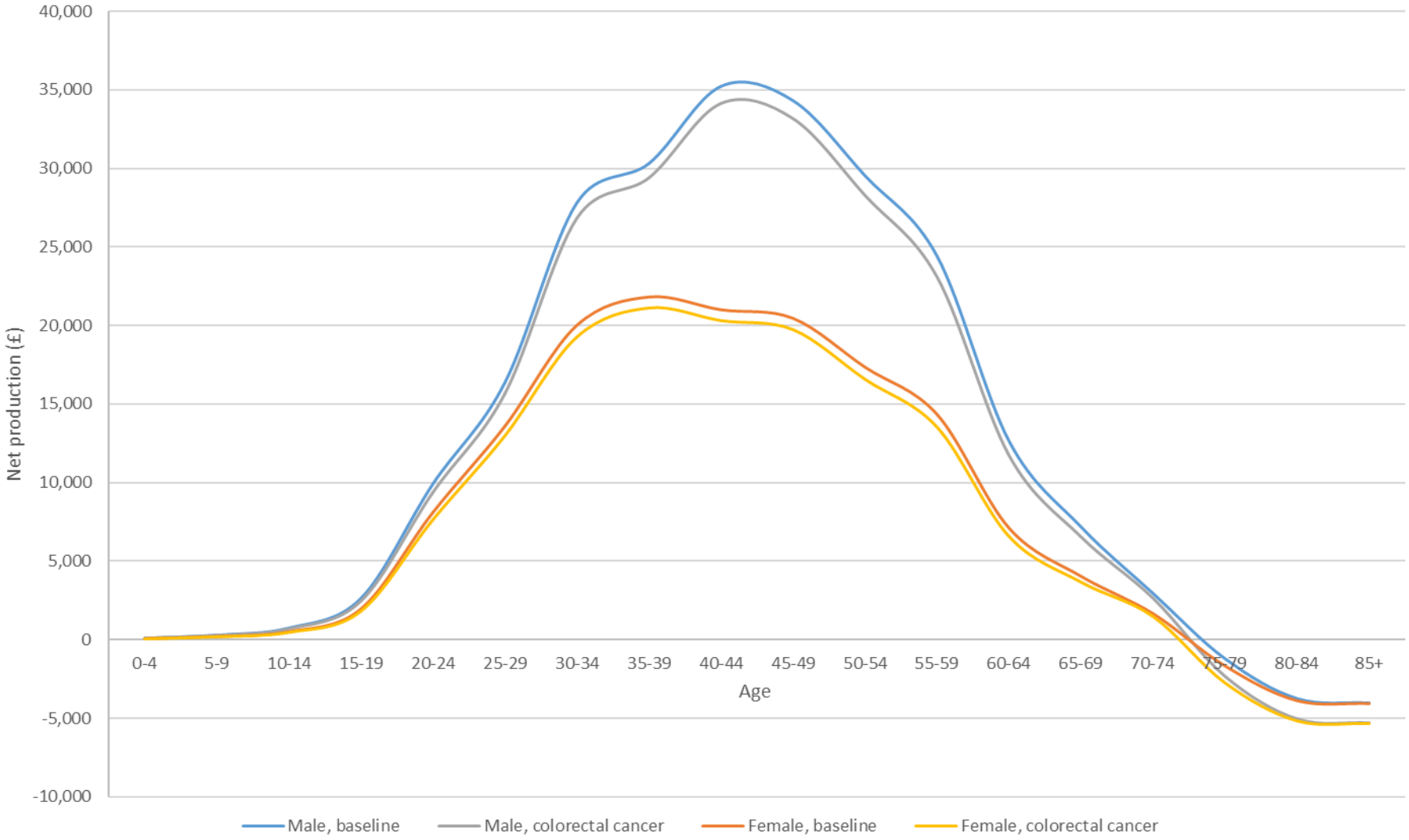
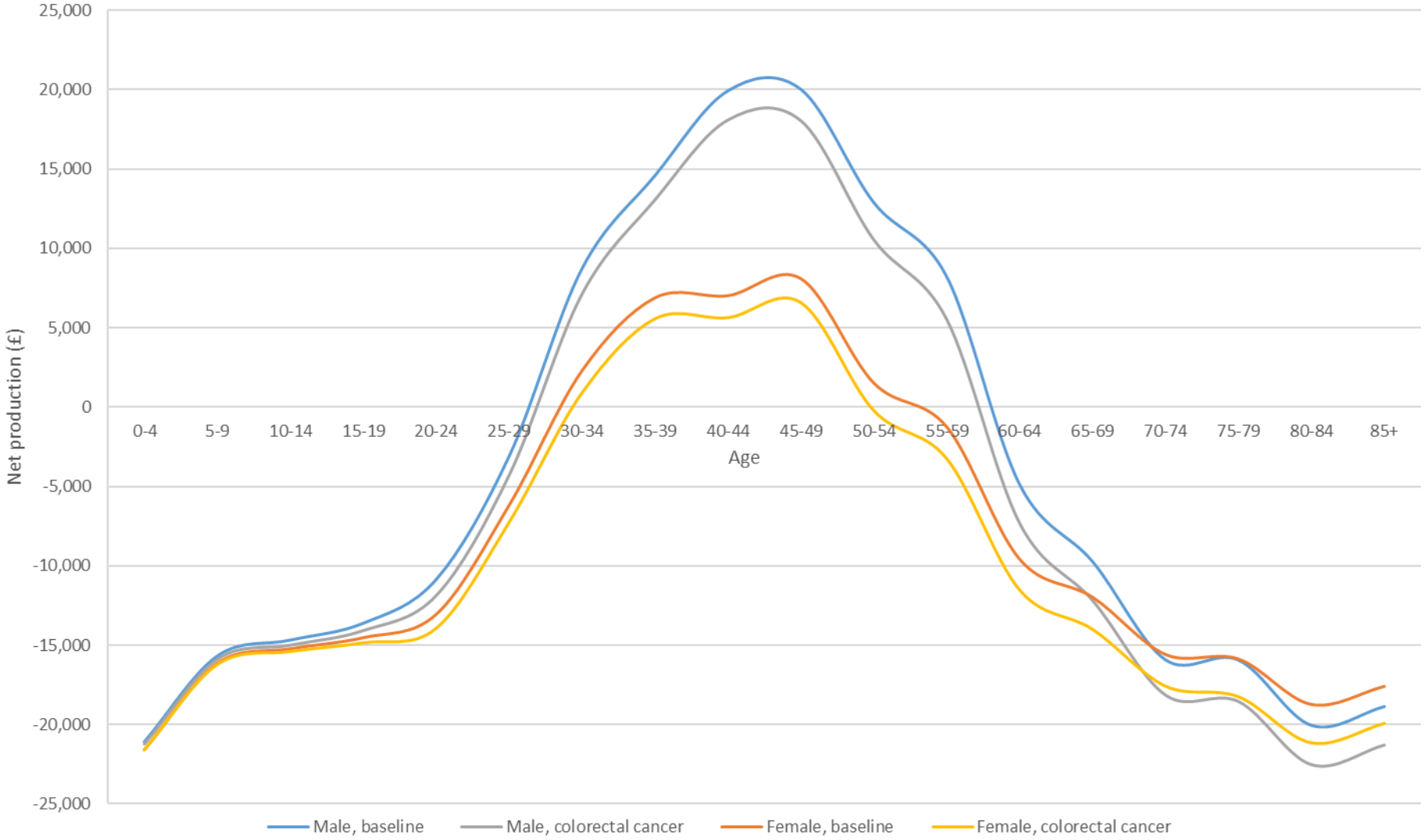


Figure 3.5c Annual net productivity from all wider societal costs by age and sex with and without colorectal cancer



3.1.4 Discussion

This chapter describes how PRIMEtime CE costs are identified, and shows that both health and social care costs vary by age, sex, and diagnosis.

3.1.4.1 Strengths

Strengths are firstly that a consistent, transparent, and comparable approach was used to estimate both healthcare costs and societal costs. As a result, any limitations are shared by each disease included in the model and should not bias outcomes in one particular direction. Secondly, using programme budgeting data and estimating disease specific expenditure from specialised services and primary care budgets means that 78% (£75.0bn) of all NHS England expenditure is accounted for. Thirdly, baseline health expenditure from unrelated diseases was estimated using NHS England cost curves, a novel approach that allows age and sex specific costs to be applied to everyone in the population cohort before PRIMEtime CE disease costs are added. This is essential for accurately estimating the true healthcare costs and savings following an intervention reflecting that individuals may live longer and incur costs arising from other age related diseases. Fourthly, costs due to residential social care are included both for PRIMEtime CE diseases and unrelated diseases. Fifthly, healthcare costs are based on 2013/14 NHS England programme budgeting data following the 2012 Health and Social Care Act and the formation of CCGs, and therefore the methodology described will remain applicable to future years' programme budgeting data as long as the structure of NHS financing and budget holding remains unchanged.

3.1.4.2 Assumptions and limitations

Estimating disease specific NHS England expenditure

The method used to disaggregate NHS England expenditure to diseases included in PRIMEtime CE relies on CCGs and acute trusts accurately reporting how their expenditure is allocated by care setting and disease area. Expenditure varies by CCG due to local population size, need, and availability of services, and the reliability of programme budgeting data has been challenged by other commentators.[180] In the 2013/14 data, for

example, one CCG reported spending £66 per person on learning disabilities (a total of £22.1m) whereas another reported spending just £0.02 per person (£4,000). Such wide variation in per capita spending is unlikely to be solely due to differences in demand but also due to differences in reporting practices. It is not possible to know if there are any reporting biases however, as discussed by Claxton et al., systematic reporting errors are unlikely.[181] There is little precedent in disaggregating programme budgeting data to estimate disease specific costs. Claxton et al. used programme budgeting data from 2003/4 to 2008/9 as part of their work estimating the NICE cost effectiveness threshold.[181] However, they are not comparable to work here as rather than estimating the cost per prevalent case for individual diseases, the authors derived expenditure and disease outcome elasticities for each of the 23 main programme budgeting categories to estimate the cost of a life-year gained.

It is possible that using a top-down approach to estimating healthcare costs overestimates savings from fixed indirect costs such as administration and overhead costs. If all NHS expenditure is allocated to individual diseases and an intervention leads to fewer people having a disease, PRIMEtime CE assumes that the proportion of all disease specific costs are reduced accordingly, including fixed costs which may in reality remain unchanged following small changes to disease prevalence. The amount that this method may over-estimate savings is limited by CCG running costs not being included in calculations (see section 3.1.2.1), and by other non-commissioning running costs (see list of care settings in table 3.3) accounting for only 2% of programme budgeting expenditure with 18% of these costs also not included as they are in category *23, other*. Using bottom-up costing with NHS reference costs may not solve this problem as these are calculated on a full absorption basis, meaning that along with direct costs, they also include fixed indirect and overhead costs.[182]

Category *23, other* was not included when estimating total disease specific NHS England expenditure because it is not possible to disaggregate these costs to specific diseases. This category includes situations where the condition was not known or reported, particularly

emergency care and transportation. It also includes some community and integrated care services such as hospital at home and rehabilitation. Category 23 was responsible for 20% of all 2013/14 programme budgeting expenditure, more than any other category, and therefore overall expenditure by disease may be underestimated. Although some diseases may be more likely to have associated costs coded under category 23 than others, it is not possible to estimate which diseases are more affected.

For the majority of diseases in PRIMEtime CE, HES admissions data were used to estimate the relative disease burden within a given programme budgeting category (part (e) in figure 3.2). This assumes that disease burden is reflected by the number of secondary care admissions and that the cost per admission is the same for each disease within a programme budgeting category. This method may underestimate total NHS England expenditure on a disease if it is managed more in an outpatient or community setting compared to other diseases in the same category, and vice versa. An example is diabetes (category 04a), where 44% of HES admissions within the programme budgeting category *04a, diabetes* report type two diabetes as the cause of admission with type one diabetes accounting for most other admissions.[163] However, Hex et al. estimate that type two diabetes accounts for 90% of total diabetes costs.[175]

The allocation of specialised services expenditure to PRIMEtime CE diseases assumes that the ratio of expenditure on 'Other secondary care' to the total spend for each disease by PCTs in 2012/13 is equivalent to the relative spend by each disease category on specialised services. This ratio is then assumed to apply to 2013/14 programme budgeting data. This method was used because, following multiple communications with NHS England, it was not possible to obtain data on specialised services expenditure for programme budgeting categories and there is no precedent in the literature. This method might over or underestimate true expenditure on different diseases, although specialised services expenditure accounted for only 18% of total 2013/14 costs thereby limiting the effect of these variations on results. Furthermore, for all cancers included in the model, it was

assumed that the same proportional expenditure from specialised services (based on category *02x, cancers and tumours, other*) whereas in reality this may vary by cancer subtype.

Primary care costs were estimated using primary care prescribing data. Expenditure on primary care was only available from NHS England as a single figure and not by specific disease. Using primary care prescribing data assumes that overall primary care expenditure is proportional to expenditure on medication. The analysis was repeated using the ratio of disease specific HES admissions to total HES admissions and included as a sensitivity analysis (a method similar to part (e) of step 1). Using HES admissions meant that primary care costs from diabetes were found to be much lower compared to using primary care prescribing data, and cancer costs were much higher. It is likely that using HES admissions would underestimate primary care costs for diseases managed primarily in the community such as diabetes, and overestimate costs from diseases managed predominantly by specialist centres such as cancer. This is in keeping with estimates that suggest just 2.5% of breast cancer costs occur outside of secondary care.[183] Another method would be to adapt the method used for specialised services in step two: in 2012/13, programme budgeting data included the care setting, *primary care,* which was not included in 2013/14, and these data could be used to estimate relative disease specific 2013/14 expenditure on primary care. These are the primary care expenditure data reported by Bhatnagar et al. when estimating the 2012/13 cost of CVD in the UK.[184] However, it is not possible to use this approach across all programme budgeting disease categories because 92% of 2012/13 non-dental *primary care* costs are assigned to category *23, other* and only 8% to specific diseases, reflecting the difficulty of coding primary care expenditure by disease.

To understand the implications of some of these assumptions in more detail, removing all cancers from the model and removing all diseases not directly related to the modelled risk factor (with the removed disease costs added to unrelated disease costs) are both included

as sensitivity analyses, as is estimating specialised services costs based on HES data (as in step (e), see appendix five for the full list of sensitivity analyses).

Estimating NHS England expenditure per person from unrelated diseases

Baseline healthcare costs were estimated using NHS cost-curves, however these were derived without using specialised services data.[165] It was therefore assumed that cost-curves represent how specialised services expenditure varies by age and sex. Furthermore, within the 65 categories of specialised services expenditure provided by NHS England, it is not known how much of a given category should be allocated to prescribing.[157] It was therefore assumed that unless the category related explicitly to a drug budget, it fell under general and acute care. This will under-estimate expenditure on drugs from some categories, in particular those relating to cancer care where cancer drugs are included in specialised services costs but are not available as a separate category from all other cancer related costs. This may under-estimate unrelated disease costs among older individuals as the prescribing cost-curve more heavily weights costs towards those aged between 60 and 80 years than the general and acute services cost-curve.

Allocation of expenditure on maternity services by age assumes that the costs per live birth are the same irrespective of the age of the mother and does not take into account miscarriages and stillbirths. These assumptions may underestimate costs from unrelated diseases for older women of childbearing age who may be more likely to suffer miscarriages or may have more complicated first time deliveries. Equally, however, costs may be overestimated for multiparous older women who require lower than average delivery support. Finally, NHS cost-curves are derived using all diseases however they are used to apportion NHS costs only to unrelated diseases, thereby not including the diseases modelled by PRIMEtime CE which are all more common among older age groups. The method used may therefore under-estimate unrelated disease costs for younger individuals and over-estimate them for older individuals.

These assumptions are tested in two sensitivity analyses: removing unrelated disease costs (both from NHS costs and social care costs), and estimating maternity costs based on the general and acute cost curve rather than number of live births by age (appendix five).

Estimating societal costs

Social care costs are estimated based on the 2013/14 local authority costs of residential care, assuming this represents the average monthly cost of all social care in England.[168] Social care costs are estimated as a function of an individual's quality of life and are therefore affected by any limitations associated with how quality of life is estimated. Quality of life is measured using the EQ-5D questionnaire and is discussed in detail in section 3.2 below. There are also limitations to the underlying methods used by the Department of Health tool when estimating societal costs.[150] The tool assumes that quality of life is the key driver of costs irrespective of diagnosis (with the exception of stroke and dementia). This may underestimate the effect of certain diagnoses on the ability of an individual to self-care beyond that which is reflected by their quality of life score.

Social care costs from unrelated diseases, which may accumulate with increased longevity following an intervention, are assumed to be reflected in the baseline age and sex specific social care costs. These costs are calculated based on the average utility value for each age and sex group across all diseases whereas they are applied to the population for unrelated diseases (not including those modelled in PRIMEtime CE). As with NHS cost-curves, this may over-estimate unrelated disease costs, particularly among older-individuals. However, the impact of this on final model results will be limited due to both discounting and because of the relatively smaller proportion of the population living to old age.

3.1.4.3 Comparisons with other studies

Given the possible variation in estimated costs using bottom-up and top-down approaches,[140–142] the top-down methods used in this chapter were compared with bottom-up costs for breast cancer, the only disease for which a comparable example from the peer-reviewed literature of sufficient quality was identified. Using the top-down methods,

2013/14 NHS England costs for breast cancer were estimated to be £469,771,000, almost identical to the £472,192,000 estimated by Luengo-Fernandez et al. using bottom-up methods.[176]

There was wide variation by disease in the estimated annual NHS cost per prevalent case, ranging from £314 for liver disease to £3,074 for pancreatic cancer (table 3.5). These estimates rely on both the estimate of the total cost of a disease and the number of people affected. Costs were estimated using the methods described in this chapter, and disease prevalence is estimated using either routine data or derived from DisMod II, a computer programme previously used by the Global Burden of Disease study for estimating either the incidence, prevalence, remission, case fatality, or mortality of a disease when a complete set of routine epidemiological data are not available (see appendix four).[185] Although the external validity of DisMod II has been questioned,[186] it remains a widely used tool with its successor, DisMod-MR, requiring the use of Python computer packages and not being available in a user-friendly format.

The two UK Health Forum studies do not report disease costs per prevalent case making it not possible to compare them with PRIMEtime CE.[115,148] Of the other studies cited in table 3.1, annual costs per prevalent case were reported by Trueman and Anokye,[31] Trueman et al.,[62] and Frew et al.[144] Each study used different sources for cost estimates ranging from £817 to £1,934 per year for diabetes, £114 to £2,047 for CHD, and £415 to £2,591 for stroke (costs for the first year after the initial episode reported by Frew et al. were not considered for this comparison as PRIMEtime CE estimates average disease prevalence costs; all costs converted to 2014 healthcare costs using the HCHS index[179]). Annual costs for diabetes, IHD, and stroke for use in PRIMEtime CE calculated in this chapter are £444, £1,905, and £843 respectively. Cost estimates per case for CVDs from the other published studies are in line with those estimated in this chapter, however costs per diabetes case are considerably higher. The cost of £1,934 per person per year was used by Trueman and Anokye and is based on a 1994/5 estimate of the excess cost of diabetic patients admitted in South

Glamorgan Health Authority, calculated using diagnosis related group costs.[31,187] Although the total cost may represent the additional burden on acute care among diabetic patients compared with non-diabetic patients, it is based on costs associated with all diabetic patients including those being admitted for co-morbid conditions such as CVD. These co-morbid costs are captured in PRIMEtime CE through the conditions IHD and stroke, thus avoiding double counting. Frew et al. estimated type two diabetes costs to be £1,474 per person per year based on unit costs over a period of 12 months before and following a 2003 trial of self-monitoring blood glucose among type two diabetics.[144,188] Costs included all interactions with the healthcare system over that period and may therefore include consultations and admissions due to, albeit potentially related, co-morbid conditions rather than type two diabetes, and the patients in the trial may not be representative of the wider diabetes population. Trueman et al. estimated per person per year diabetes costs to be £817, taken from a 2005 company submission of cost-effectiveness analysis of the weight loss drug, sibutramine (reported in the 2006 NICE guideline on obesity prevention).[62,118] Diabetes costs in this report are taken from a variety of sources and it is not stated how exactly they are calculated.

Frew et al. also reported costs per prevalent case for breast cancer and colorectal cancer.[144] They cite a study by Dolan et al. that used bottom-up costs to estimate the annual cost per case of breast cancer in the five years post diagnosis to be £9,832 per year (2014 prices),[183] and estimate the cost of colorectal cancer to be £10,814 per case per year based on a report by the York Health Economics Consortium using bottom-up costs to estimate that the total annual cost of colorectal cancer in England is £1,326m (2014 prices).[189] This compares to £333m estimated in table 3.5 and £487m for colorectal and anal cancer estimated by Luengo-Fernandez et al. (2014 prices).[176] Trueman et al. estimated the annual cost per case of colorectal cancer to be £9,154 from a 2004 Australian study modelling the cost-effectiveness of sigmoidoscopy screening over a 10 year timescale, which uses bottom-up costs from a variety of sources.[62,190] These reported cancer costs are all significantly higher than results in table 3.5. In part this may be due to reported estimates not including low

long term costs of those who are cured of their disease yet still remain a prevalent case and die of an unrelated cause (remission rates are not included in PRIMEtime CE). And as with diabetes cost estimates, costs attributed to the cancer diagnosis in these reports may be attributed to co-morbid conditions in PRIMEtime CE. Furthermore, in PRIMEtime CE, as with Blakely et al.,[114] age and sex specific expenditure from unrelated diseases is added to total healthcare costs (figure 3.4) thereby increasing total per person costs to be much closer to those estimated using bottom-up cost estimates. This means that were an individual not to die of colorectal cancer at age 80, for example, then they would still likely have relatively high healthcare costs due to other co-morbid disease. Finally, the higher annual cost of pancreatic cancer compared with other cancers in PRIMEtime CE is consistent with other studies suggesting that pancreatic cancer is a very resource intensive diagnosis.[114,191]

Given that estimates of the excess cost per person of a cancer diagnosis are generally lower in table 3.5 than in other published studies, a sensitivity analysis explores the impact of removing cancer from PRIMEtime CE.

When estimating disease costs per prevalent case, Blakely et al. and Frew et al. both used higher costs for the first year post diagnosis than subsequent years due to a greater need for diagnostics and treatment in the acute phase of the disease; and Blakely et al. used higher costs for the final year of life resulting from the additional costs of palliative care.[114,144] Using programme budgeting costs means that in PRIMEtime CE all healthcare costs, irrespective of time since diagnosis, are averaged across all prevalent cases in the cohort. Therefore if an intervention has the effect of changing the average time between contracting a modelled disease and dying from the disease, modelled costs may overestimate (if the time with a disease lengthens) or underestimate (if the time with a disease shortens) actual healthcare expenditure.

Following the review of modelling studies in chapter one, there were no UK studies identified that estimated unrelated healthcare costs by age and sex. Blakely et al. estimated healthcare expenditure for diseases unrelated to tobacco consumption for the New Zealand

population by age and sex.[114] They explicitly modelled COPD, CHD, stroke, and 12 different cancers, with costs from unrelated diseases quantified using bottom-up unit costs of patient care from a linked national healthcare database. Costs for male individuals by age are reported by Blakely et al. and although the included diseases differ a little from PRIMEtime CE (PRIMEtime CE does not include COPD and some cancers but unlike Blakely et al., does include liver disease), the magnitude of expenditure pattern by age is similar to that shown in figure 3.3 except for those aged under four years. Blakely estimates annual expenditure on males aged under one year to be 5,228 New Zealand Dollars (NZD) in 2011 (approximately £2,744 in 2014 adjusted for healthcare inflation) and NZD1,215 (£638) for age 1-4 years before falling to NZD621 (£326) for children aged 5-9 years. The spike in healthcare costs for those aged under one year is not included in PRIMEtime CE but this will not affect results as only adults are modelled.

3.1.4.4 Conclusion

In summary, section 3.1 describes a novel approach to estimating NHS England costs for PRIMEtime CE diseases and for unrelated diseases. Furthermore, it outlines how social care costs are quantified. Despite the limitations associated with these methods, their key strength is that they use a consistent approach to estimating costs across multiple diseases meaning that the cost implications of different interventions affecting these diseases can be directly compared.

3.2 Baseline disease utilities

3.2.1 Background

Health state utility values are used to estimate the impact of a disease on health related quality of life (HRQoL) often on a scale between zero (equivalent to death, although health can be valued as worse than death) and one (perfect health). Assigning utilities to a range of diseases allows for comparisons between interventions affecting more than one disease, and are used by NICE to calculate QALYs.[192] As discussed in section 1.3.4, a QALY is a

measure of both quality and quantity of life, and cost-effectiveness analyses reported by NICE present results in terms of costs per QALY to enable different interventions, technologies, and diseases to be compared.

There are different ways to estimate HRQoL.[193] Guidelines from NICE state that for economic modelling, the EQ-5D should be used by patients to describe their health state.[194] The EQ-5D questionnaire is a standardised tool for measuring health across five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.[195] In order to generate utility values from EQ-5D results, a variety of methods can be used. These include a visual analogue scale (VAS), standard gamble (SG, for example, at what risk of death would someone gamble on being cured from a disease), and time trade off (TTO, the length of time in poor health someone would be willing to give up in order to live for a shorter amount of time in perfect health).[193] In order to value the HRQoL assessed by EQ-5D, NICE prefers TTO completed by a representative sample of the UK general population.[194]

This section describes how comparable disease specific EQ-5D derived utility values for use in PRIMEtime CE were identified and validated.

3.2.2 Methods

3.2.2.1 Source of utility values

If not collected prospectively, utility values can be identified through systematic literature searches, however this may result in finding non-comparable utility values for different diseases resulting from differences between the underlying patient groups and study methodology. For example, for baseline EQ-5D derived utility scores valued using TTO by age, sex, and BMI, Trueman et al. used an analysis of the 1996 Health Survey for England (HSE).[62,196] Multipliers were then used to adjust the baseline utilities following a diagnosis of colon cancer, CVD, or diabetes, based on a cost-effectiveness analysis of investigating rectal bleeding among 35 year old adults[197] and an economic evaluation of sibutramine in weight loss (used in the 2006 NICE guideline on obesity prevention).[118] Therefore the utility

value for colon cancer among men ranged between 0.78 and 0.84 and between 0.71 and 0.76 for diabetes depending on BMI. This is compared to the utilities used by Frew et al.[144] of 0.68 (year one post diagnosis) and 0.71 (year two onwards) for colon cancer derived from US cancer survivors answering the Health and Activities Limitation index (HALex),[198] and 0.80 for diabetes taken from the UK Prospective Diabetes Study using EQ-5D scores with TTO valuation.[199] Both Trueman et al. and Frew et al. used different valuation methods and different population groups at different ages and disease stage to estimate HRQoL. This means that both between the studies and within each study, utility scores for different diseases are unlikely to be comparable with each other and may not be representative of the wider population.

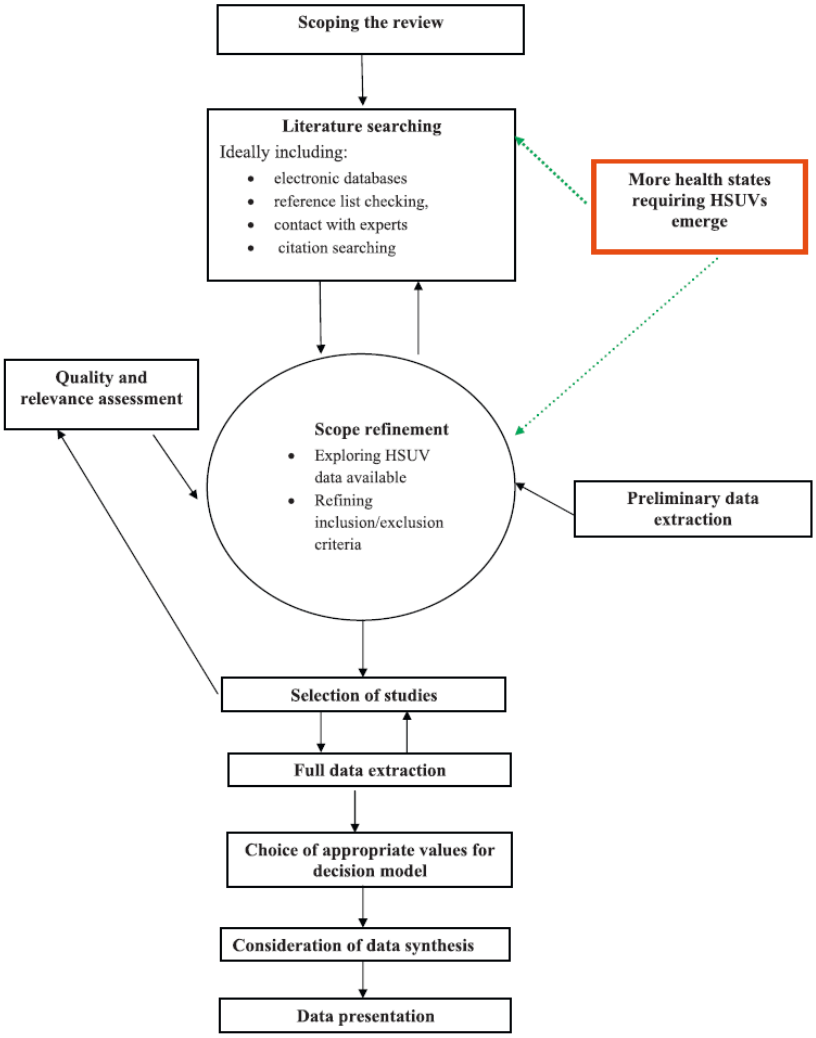
PRIMEtime CE uses Sullivan et al.'s catalogue of EQ-5D utility values derived from the US Medical Expenditure Panel Survey using UK general population based TTO scores.[200,201] The EQ-5D results were collected from annual US medical expenditure surveys between 2000 and 2003, and included 79,522 unique responses with linked clinical data for the preceding year covering 135 ICD-9 codes; an equivalent representative UK source of EQ-5D results does not exist. The results were then valued using TTO by a representative UK population sample. As well as providing baseline utility values by age and sex, the catalogue includes both unadjusted results and the results of a regression analysis reporting the marginal disutility (utility decrement) of a given disease adjusting for age, comorbidity, gender, race, ethnicity, income, and education.

3.2.2.2 Validation of Sullivan et al. utility values

Potential limitations of Sullivan et al. are that the EQ-5D scores are sampled from a non-English patient population and there is likely to be underrepresentation of those at very early and late stages of disease as these population groups may be too unwell or unwilling to participate. In order to identify whether utility values published in the Sullivan et al. catalogue were similar to those identified in other populations using different methods, EQ-5D derived breast cancer utility values were systematically identified from the literature

using a pre-defined protocol following NICE guidance (shown in figure 3.6, for full protocol see appendix seven).[194] Results were then compared to utility values published by Sullivan et al.[200]

Figure 3.6 Reproduction of figure one from Papaioannou et al., 2010 on how to identify utility values from the literature (reproduced with permission)[194]



HSUV, health state utility value

3.2.2.3 Applying utility values from Sullivan et al. to PRIMEtime CE

Matching the ICD-9 disease codes published by Sullivan et al. to ICD-10 codes used by PRIMEtime CE was done with the website www.icd10data.com.[202] The calculation of baseline utility values by age and sex, disease specific utility decrements, and adjustment for number of chronic conditions followed guidance published by Sullivan et al. in their UK

and US papers.[200,203] To use these utility values to estimate the change in QALYs in PRIMEtime CE following an intervention, the age and sex specific utility values change with age based on number of years that the model has been running for, and on changes to the proportion of the population in different disease states. Age, sex, and disease specific utility values are also used to inform social care costs and productivity losses as discussed in section 3.1.

Where Sullivan et al. did not report a utility decrement for a disease included in PRIMEtime CE, the utility decrement for the most closely related disease was used. Where more than one utility decrement was reported for a single PRIMEtime CE disease, utility decrements were weighted based on the number of respondents under each ICD-9 code in the MEPS survey. Furthermore, where ICD-9 codes with reported utility decrements corresponded to different stages following diagnosis (for example, ICD-9 410 corresponds to acute MI and ICD-9 412 to old MI), separate utility decrements were calculated for incident cases in the first year post diagnosis and prevalent cases from two years post diagnosis onwards. For all other diseases the same utility decrement was used irrespective of time since diagnosis.

Standard errors are taken directly from Sullivan et al. Where utility decrements are a weighted average from more than one disease, standard errors were converted to standard deviations and combined using methods described in the Cochrane Handbook for Systematic Reviews of Interventions before being converted back to standard errors.[204]

3.2.3 Results

3.2.3.1 Validation of Sullivan et al. utility values

Using the protocol in appendix seven, the breast cancer search identified 196 studies for full text review from which 23 studies were included for data-extraction. Four studies received the joint highest study *quality and applicability score* and the extracted utility values from these four studies are shown in table 3.7. The utility values from each of these four studies overlap with the mean utility value among breast cancer patients reported by Sullivan et al.[200]

Table 3.7 Comparison of breast cancer utility values identified through the systematic review and by Sullivan et al.[200]

Study	**Patient population**	**Range of published utilities** (SE/95% confidence intervals)
Sullivan et al., 2011[200]	Catalogue of EQ-5D scores for the United Kingdom (EQ-5D results from US population, valued by a UK population). Patients with ICD-9 code 174, malignant neoplasm female breast	0.75 (0.72-0.78)
Prescott et al., 2007[205]	English and Scottish patients with surgically treated node negative breast cancer up to 15 months post treatment with or without radiotherapy	0.72 (0.68-0.76) to 0.78 (0.74-0.81)
Hall et al., 2015[206]	English breast cancer survivors up to 15 months post diagnosis	0.74 (0.71-0.78) to 0.83 (0.76-0.89)
Haines et al., 2010[207]	Australian newly diagnosed patient undergoing adjuvant chemotherapy with or without a home based exercise programme	0.78 (0.19) to 0.85 (0.19)
Vrettos et al., 2012[208]	Greek breast cancer patients presenting to an oncology day clinic	0.69 (0.29)

3.2.3.2 Applying Sullivan et al. utility values to PRIMEtime CE

Baseline utility values for use in PRIMEtime CE are shown in table 3.8.

Table 3.8 Baseline EQ-5D utility values by age and sex for use in PRIMEtime CE

Age	**Male**	**Female**	**Age**	**Male**	**Female**
0	1.000	1.000	**51**	0.800	0.799
1	1.000	1.000	**52**	0.800	0.799
2	1.000	1.000	**53**	0.799	0.798
3	1.000	1.000	**54**	0.799	0.798
4	1.000	1.000	**55**	0.799	0.798
5	1.000	1.000	**56**	0.799	0.798
6	1.000	1.000	**57**	0.798	0.797
7	1.000	1.000	**58**	0.798	0.797

8	1.000	1.000	**59**	0.798	0.797
9	1.000	1.000	**60**	0.776	0.775
10	0.916	0.916	**61**	0.776	0.775
11	0.916	0.916	**62**	0.776	0.775
12	0.916	0.916	**63**	0.775	0.774
13	0.916	0.916	**64**	0.775	0.774
14	0.915	0.915	**65**	0.775	0.774
15	0.915	0.915	**66**	0.774	0.773
16	0.915	0.915	**67**	0.774	0.773
17	0.914	0.914	**68**	0.774	0.773
18	0.914	0.914	**69**	0.774	0.773
19	0.914	0.914	**70**	0.725	0.724
20	0.907	0.906	**71**	0.725	0.724
21	0.907	0.906	**72**	0.725	0.724
22	0.907	0.906	**73**	0.724	0.723
23	0.906	0.905	**74**	0.724	0.723
24	0.906	0.905	**75**	0.724	0.723
25	0.906	0.905	**76**	0.724	0.723
26	0.906	0.905	**77**	0.723	0.722
27	0.905	0.904	**78**	0.723	0.722
28	0.905	0.904	**79**	0.723	0.722
29	0.905	0.904	**80**	0.659	0.658
30	0.881	0.880	**81**	0.659	0.658
31	0.881	0.880	**82**	0.658	0.657
32	0.881	0.880	**83**	0.658	0.657
33	0.880	0.879	**84**	0.658	0.657
34	0.880	0.879	**85**	0.658	0.657
35	0.880	0.879	**86**	0.657	0.656
36	0.880	0.879	**87**	0.657	0.656
37	0.879	0.878	**88**	0.657	0.656
38	0.879	0.878	**89**	0.656	0.655
39	0.879	0.878	**90**	0.656	0.655
40	0.839	0.838	**91**	0.656	0.655
41	0.839	0.838	**92**	0.656	0.655
42	0.839	0.838	**93**	0.655	0.654
43	0.838	0.837	**94**	0.655	0.654

44	0.838	0.837	**95**	0.655	0.654
45	0.838	0.837	**96**	0.655	0.654
46	0.838	0.837	**97**	0.654	0.653
47	0.837	0.836	**98**	0.654	0.653
48	0.837	0.836	**99**	0.654	0.653
49	0.837	0.836	**100**	0.653	0.652
50	0.800	0.799			

Disease specific utility decrements are shown in table 3.9. For pancreatic cancer (ICD-9 157) and kidney cancer (ICD-9 189.0) there was no specific utility decrement reported by Sullivan et al. Instead, for pancreatic cancer the decrement for *ICD-9 195, other malignant neoplasm* was used, and the kidney cancer decrement in PRIMEtime CE was assumed to be the same as for *ICD-9 189, urinary cancer*.

For IHD and stroke, more than one utility decrement was reported by Sullivan et al. for the same PRIMEtime CE disease (table 3.9). Utility decrements for IHD ICD-9 codes 410-414 were all provided, including acute MI, old MI, and angina pectoris. A utility decrement was therefore calculated for both the year of IHD diagnosis (incident case) and for IHD year two post incidence onwards (prevalent case). Acute IHD ICD-9 codes *410, acute myocardial infarction* and *411, other acute ischaemic heart disease* were only used for incident IHD and not prevalent IHD, and ICD-9 code 410 was assumed to contribute 8/52nds of its decrement score as acute MI specifically refers to the first eight weeks post MI. By the same logic, ICD-9 code *412, old myocardial infarction* was assumed to contribute 44/52nds of its decrement score towards incident IHD. The other ICD-9 codes contributed to both incident and prevalent decrements. For stroke, utility decrements for five relevant ICD-9 codes were available. As with IHD, acute ICD-9 codes (*cerebrovascular accident, 436* and *precerebral occlusion, 433*) did not contribute to the utility decrement for prevalent stroke.

Table 3.9 Utility values used in PRIMEtime CE

PRIMEtime CE disease outcome	**PRIMEtime CE ICD-10 codes**	**Equivalent ICD-9 codes**	**Available utility decrement from Sullivan et al.[200]***	**Utility value used in PRIMEtime CE** (SE)
Ischaemic heart disease	I20-I25	410-414	410: -0.063 411: -0.087 412: -0.037 413: -0.085 414: -0.063	Incident case: -0.071 (0.024) Prevalent: -0.070 (0.015)
Stroke	I60-I69	430-438	433: -0.035 435: -0.033 436: -0.117 437: -0.031 438: -0.073	Incident: -0.094 (0.019) Prevalent: -0.046 (0.031)
Type two diabetes	E11, E14	250.x0	250: -0.071	-0.071 (0.005)
Breast cancer	C50	174, 175	174: -0.019	-0.019 (0.014)
Colon cancer	C18-C20	153, 154.0, 154.1	153: -0.067	-0.067 (0.017)
Lung cancer	C34	162.2-162.9	162: -0.119	-0.119 (0.043)
Stomach cancer	C16	151	151: -0.071	-0.071 (0.105)
Liver cancer	C22	155	155: -0.093	-0.093 (0.044)
Kidney cancer	C64	189.0	189: -0.048	-0.048 (0.041)
Pancreatic cancer	C25	157	195: -0.086	-0.086 (0.027)
Liver disease	K70, K74	571.0-571.3, 571.5, 571.6, 571.9	571: -0.083	-0.083 (0.031)

*reported utility decrement controlled for age, comorbidity, gender, race, ethnicity, income, and education; SE, standard error

3.2.4 Discussion

3.2.4.1 Strengths and limitations of using Sullivan et al.

The major strength of using Sullivan et al. to provide utility values for PRIMEtime CE is that the same dataset is used for all diseases.[200] Therefore the utilities are broadly comparable across diseases and any limitations and biases are unlikely to systematically

affect one disease more than any other. Furthermore, given that the EQ-5D scores are taken from a cross-section of the population, they should capture both the clinical and demographic heterogeneity associated with the disease population.

The systematic review of breast cancer utilities from the literature helps to validate Sullivan et al.'s catalogue, especially given the limitation that data are from 2000, 2001, 2002, and 2003 US population samples rather than a more recently sampled UK population (although the valuation method does use UK specific TTO scores). A limitation is that individuals at very early or late stages of their disease are likely to be under-represented due to being too unwell or unwilling to participate. Therefore disease specific utility decrements might under-estimate the true average reduction in quality of life. A more robust method might be to conduct a systematic review and meta-analysis of utility values for each disease included in PRIMEtime CE, but this would be considerably more resource intensive. Finally, as with estimating unrelated disease costs (see section 3.1.4.2), baseline age and sex utility values are derived based on all diseases in the population rather than only unrelated diseases (those not explicitly modelled by PRIMEtime CE). Therefore, utility decrements that accrue due to increased longevity following an intervention may over-estimate the true utility decrement. However, again the effect of this on model results is limited due to discounting and the relatively smaller proportion of the modelled cohort living to old age.

3.2.4.2 Comparisons with other studies

The UK Health Forum and Gulliford et al. also used Sullivan et al. to estimate HRQoL for diseases in their models.[85,115,148,200] Gulliford et al. used a similar method to that described in this chapter however rather than using utility decrements by ICD-9 code, they used results reported by clinical classification category (CCC), another form of disease classification. Furthermore, where more than one utility was reported that was relevant to a disease, a single CCC was chosen rather than any weighting used. The UK Health Forum used Sullivan et al. for some disease outcomes, and other sources for utility values relating to stroke, CHD, and pancreatic cancer. Utility values estimated using EQ-5D scores for CHD

and stroke were taken from a variety of sources reported in a NICE economic modelling paper on childhood obesity.[209] The EQ-5D derived utility value for pancreatic cancer was taken from a US randomised controlled trial of chemotherapy.[210] Therefore the utility values for stroke, CHD, and pancreatic cancer are not directly comparable to those collated by Sullivan et al. as they are from different population groups at different stages of disease and are not adjusted for confounding variables such as income, gender, and comorbidity. Furthermore, it is unclear from some of the sources used by the UK Health Forum whether EQ-5D scores were valued using TTO or another method. Of the other studies in table 3.1, baseline EQ-5D utilities by age and sex were either not used[31,52,116,117,144] or in the case of Trueman et al., not valued using TTO.[62]

Finally, Ara and Brazier published HRQoL scores for the English population based on the EQ-5D questionnaire and TTO valuation.[211] They used data from 41,174 respondents to the HSE between 2003 and 2006 to estimate how utility scores vary by age in the general population and by disease. General population utility values reported by Ara and Brazier are approximately 0.03 higher than those reported by Sullivan et al. across all ages except for those aged over 85 years where utility values are similar.[200] The difference between the disease specific utility decrements calculated by Brazier and Ara and those reported by Sullivan et al. varies by disease. For example, utility decrements are similar for hypertension and cancer, but much larger for heart attack and stroke. These differences may be explained by some key differences between the two analyses. Sullivan et al. adjusted for multiple confounders, including co-morbidity, and will likely more accurately reflect the utility decrement associated with having a disease independent of all other conditions compared to the Ara and Brazier results which are unadjusted. Furthermore, the HSE questions used by Ara and Brazier to identify those with disease was: "Do you have a long-standing illness, disability, or infirmity? By long-standing I mean anything that has troubled you over a period of time, or that is likely to affect you over a long period of time?"[211] whereas Sullivan et al. is based on any condition ever recorded for an individual, irrespective of its current impact on quality of life.[203] As such, Sullivan et al.'s analysis will

include individuals who have diagnoses that are no longer significantly impacting on their quality of life, reflecting the fact that PRIMEtime CE does not include remission rates. Furthermore, Ara and Brazier's results would not be as useful as Sullivan et al.'s data for informing PRIMEtime CE as only a limited number of disease specific utility scores are reported, for example all cancers are grouped together into a single category.

3.2.4.3 Conclusion

Utility decrements from Sullivan et al. are ideal for use in PRIMEtime CE because, as with health and social care costs described in section 3.a, they are comparable across diseases. The impact of the choice of costs and utilities on results is discussed in more detail in chapter six where PRIMEtime CE is compared with other cost-effectiveness models.

Chapter 4. Effect of interventions

Stakeholders including charities, government officials, and patients were asked to select one physical activity intervention and one dietary intervention for modelling in PRIMEtime CE from a list of eight options (section 2.1.1). The two selected interventions were reformulating food to be healthier and increasing access to leisure centres. This chapter details how the two interventions were parameterised and outlines the major strengths and limitations of the methods used.

4.1 Dietary intervention

4.1.1 Background and methods

The diet intervention models the health impact of reformulating food in England to have less salt. This was chosen because at the time of deciding on an intervention in 2015 there were food specific government targets for salt reduction but not for other nutrients such as fats or sugar.[212] Furthermore, changing salt consumption represents a robust test of the PRIMEtime CE model as both proximal and intermediate risk factors are affected (salt and hypertension), along with multiple disease outcomes.

In 2003, the UK Scientific Advisory Committee on Nutrition (SACN) recommended that average salt consumption should be no more than 6g (2,400mg sodium) per day for adults.[213] The Food Standards Agency (FSA) began a voluntary salt reduction campaign in 2004, working alongside industry to reduce salt in food.[214,215] The programme was widely interpreted as a success, and between 2005 and 2014, salt consumption declined by 11% to an average of 8.0g per day.[216] However, this is still substantially greater than the 6g per day advised by SACN. In 2010, responsibility for the salt reduction programme transferred to the Department of Health in England. The Department of Health set industry targets for the amount of salt in food and both 2012 and 2017 salt targets were established and managed by the Responsibility Deal, a government led forum for businesses and other organisations to make voluntary pledges to take action to improve public health.[217,212] The

2017 targets are different for food consumed at home and food consumed out of home, with at home targets listed by food category (table 4.1) and out of home targets focusing on meals.

4.1.1.1 Estimating the change in salt consumption

This thesis aims to estimate the cost-effectiveness of achieving the at home 2017 salt targets (table 4.1) assuming a baseline year of 2014. Salt consumption in England was taken from the combined results of four years of the National Diet and Nutrition Survey (NDNS) from between 2008 and 2012 and assumed to be unchanged between then and 2014 (more recent NDNS results from 2013 and 2014 were not available at the individual level at the time of writing).[218] The NDNS is an annual cross-sectional survey of a representative sample of around 1,000 people in the UK. The survey asks individuals (or their parent/carer if under 12 years old) aged over 1.5 years from across the UK to complete four day food diaries recording everything eaten or drunk, of which 4,156 diaries with at least three complete days were available for years 2008-2012 (3,441 diaries from people living in England). For individuals within the NDNS sample, data were available for approximately 11,000 different foods, aggregated into 61 food groups and 156 subsidiary food groups.

Table 4.1 The Department of Health 2017 salt targets for food purchased for consumption at home.[212] Targets are average salt levels unless stated

Main Product Category	**Sub categories** (where relevant)	**Target for 2017** (g salt or mg sodium per 100g)
1. Meat Products	1.1 Bacon	2.88g salt or 1150mg sodium
	1.2 Ham/other cured meats	1.63g salt or 650mg sodium
	1.3 Sausages	
	1.3.1 Sausages	1.13g salt or 450mg sodium
	1.3.2 Cooked sausages and sausage meat products	1.38g salt or 550mg sodium

	1.4 Meat Pies	
	1.4.1 Delicatessen, pork pies and sausage rolls	0.98g salt or 390mg sodium
	1.4.2 Cornish and meat-based pasties	0.9g salt or 360mg sodium
	1.4.3 Other meat-based pastry products including pies and slices, canned and frozen products	0.68g salt or 270mg sodium
	1.5 Cooked uncured meat Includes all roast meat, sliced meat etc. Excludes ham (see category 1.2)	
	1.5.1 Whole muscle.	0.68g salt or 270mg sodium (max)
	1.5.2 Reformed whole muscle	0.9g salt or 360mg sodium (max)
	1.5.3 Comminuted or chopped reformed meat	1.35g salt or 540mg sodium (max)
	1.6 Burgers and Grill Steaks	0.75g salt or 300mg sodium
	1.7 Frankfurters, hotdogs, and burgers	
	1.7.1 Canned frankfurters, canned hotdogs and canned burgers only	1.38g salt or 550mg sodium
	1.7.2 Fresh chilled frankfurters	1.5g salt or 600mg sodium
2. Bread	2.1 Bread and rolls	0.9g salt or 360mg sodium
	2.2 Bread and rolls with additions	1g salt or 400mg sodium
	2.3 Morning goods - yeast raised	0.73g salt or 290mg sodium

	2.4 Morning goods - powder raised	1.13g salt or 450mg sodium
3. Breakfast Cereals	3.1 Breakfast cereals	0.59g salt or 235mg sodium
4.Cheese	4.1 Cheddar and other similar "hard pressed" cheeses	1.75g salt or 700mg sodium
	4.2 "Fresh" cheeses 4.2.1 Soft white cheese	0.5g salt or 200mg sodium
	4.2.2 Cottage cheese - plain and flavoured	0.5g salt or 200mg sodium
	4.3 Mozzarella	1.35g salt or 540mg sodium
	4.4 Blue cheese	2.0g salt or 800mg sodium
	4.5 Processed Cheese 4.5.1 Cheese spreads	1.63g salt or 650mg sodium
	4.5.2 Other processed cheese	1.7g salt or 680 mg sodium
5. Butter	5.1 Salted butters and buttery spreads	1.48g salt or 590mg sodium
	5.2 Lightly salted butter	1.13g salt or 450mg sodium
6. Fat spreads	6.1 Margarines/other spreads	1.06g salt or 425mg sodium
7. Baked Beans	7.1 Baked beans in tomato sauce without accompaniments	0.56g salt or 225mg sodium (max)
	7.2 Baked beans and canned pasta with accompaniments	0.68g salt or 270mg sodium
8. Ready meals and meal centres	8.1 Ready Meals and Meal Centres	0.63g salt or 250mg sodium
9. Soups	9.1 Soups (as consumed)	0.53g salt or 210mg sodium
10. Pizzas	10.1 All Pizzas (as consumed)	1.0g salt or 400mg sodium

11.Crisps and snacks	11.1 Standard potato crisps	1.31g salt or 525mg sodium
	11.2 Extruded and sheeted snacks	1.7g salt or 680mg sodium
	11.3 Pelleted snacks	2.13g salt or 850mg sodium
	11.4 Salt and Vinegar products	1.88g salt or 750mg sodium
12. Cakes, pastries, fruit pies and other pastry-based desserts.	12.1 Cakes	0.43g salt or 170mg sodium
	12.2 Pastries	0.35g salt or 140mg sodium
	12.3 Sweet Pies and other shortcrust or choux pastry based desserts	0.25g salt or 100mg sodium
13. Bought Sandwiches	13.1 Sandwiches with high salt fillings	0.9g salt or 360mg sodium
	13.2 Sandwiches without high salt fillings	0.68g salt or 270mg sodium
14. Table Sauces	14.1 Tomato ketchup	1.7g salt or 680mg sodium (max)
	14.2 Brown sauce	1.2g salt or 480mg sodium (max)
	14.3 Salad cream	1.58g salt or 630mg sodium (max)
	14.4.1 Mayonnaise (not reduced fat/calorie)	1.25g salt or 500mg sodium (max)

	14.4.2 Mayonnaise (reduced fat/calorie only)	1.7g salt or 680mg sodium (max)
	14.5 Salad dressing	1.5g salt or 600mg sodium (max)
15. Cook-in and Pasta Sauces, thick sauces and pastes	15.1 All cook in and pasta sauces (except Pesto and other thick sauces and pastes)	0.75g salt or 300mg sodium
	15.2 Pesto and other thick sauces	1.38g salt or 550mg sodium
	15.3 Thick pastes	3.25g salt or 1300mg sodium
16. Biscuits	16.1 Sweet Biscuits	0.55g salt or 220mg sodium
	16.2 Savoury biscuits	1.3g salt or 520mg sodium
17. Pasta	17.1 Pasta and noodles, plain and flavoured	0.5g salt or 200mg sodium
18. Rice	18.1 Rice (unflavoured), as consumed	0.18g salt or 70mg sodium (max)
	18.2 Flavoured rice, as consumed	0.45g salt or 180mg sodium
19. Other cereals	19.1 Other cereals	0.55g salt or 220mg sodium
20. Processed puddings	20.1 Dessert mixes, as consumed	0.45g salt or 180mg sodium (max)
	20.2 Cheesecake	0.28g salt or 110mg sodium
	20.3 Sponge-based processed puddings	0.43g salt or 170mg sodium
	20.4 All other processed puddings	0.18g salt or 70mg sodium
21. Quiche	21.1 Quiches	0.55g salt or 220mg sodium
22. Scotch Eggs	22.1 Scotch eggs	0.78g salt or 310mg sodium (max)

23. Canned Fish	23.1 Canned tuna	0.9g salt or 360mg sodium
	23.2 Canned salmon	0.8g salt or 320mg sodium
	23.3 Other canned fish	0.85g salt or 340mg sodium
24. Canned vegetables	24.1 Canned and bottled vegetables	0.13g salt or 50mg sodium (max)
	24.2 Canned processed, marrowfat and mushy peas	0.45g salt or 180mg sodium (max)
25. Meat alternatives	25.1 Plain meat alternatives	0.63g salt or 250mg sodium (max)
	25.2 Meat free products	0.9g salt or 360mg sodium
	25.3 Meat-free bacon	1.88g salt or 750mg sodium (max)
26. Other processed potatoes	26.1 Dehydrated instant mashed potato, as consumed	0.15g salt or 60mg sodium (max)
	26.2 Other processed potato products	0.46g salt or 185mg sodium
27. Beverages	27.1 Dried Beverages, as consumed	0.15g salt or 60mg sodium (max)
28. Stocks and gravies	28.1 Stocks, as consumed	0.75g salt or 300mg sodium
	28.2 Gravy, as consumed	0.95g salt or 380mg sodium

Individual level NDNS data for 3,441 individuals living in England who completed either three or four days of their food diary were used to calculate the average salt level consumption by age group (0-3, 4-12, 13-17, 18-34, 35-54, and over 55 years) and by sex. To do this, food diary data by age and sex were extracted and weighted within each age and sex group using the relevant NDNS weighting variable to adjust for non-response by adults and children, thereby ensuring the sample was representative of the UK population. The average

salt content per 100g of the 156 subsidiary food groups was calculated separately for each age and sex group. This was then multiplied by the weight of each food group consumed at the individual level for each age and sex group to calculate the average daily salt consumption per person for all food recorded in the food diaries.

Salt consumption following the intervention was calculated by changing the average salt content of each of the 156 subsidiary food groups to the 'at home' average 2017 salt target before again multiplying by the weight of each subsidiary food group consumed at the individual level for each age and sex group. The maximum salt target was used where no average target is specified (see table 4.1), and no change to the salt content of a subsidiary food group was made following the intervention where the 2017 target was greater than the NDNS recorded salt content.

This method assumes that nobody changes the food they eat following the intervention and that all out of home consumption recorded by NDNS food diaries achieves the 'at home' salt targets. For 19 of the NDNS subsidiary food groups, the included products were a composite of different foods included in the 2017 targets and in these circumstances the average salt target was used. It was not possible to weight the salt targets for these 19 food groups by the mass of each food-type consumed because this would require applying the relevant target to each of 11,000 individual foods included in NDNS, which was beyond the scope of this thesis. For the NDNS subsidiary food group, *1R other cereals*, couscous was not included under a salt target but all other foods were. In this case the salt target was applied to the entire category, assuming that couscous was a minor contributor to the category.

The standard error of the difference in salt consumption between baseline and following the intervention was calculated from NDNS data for each age and sex group and used to estimate uncertainty in the intervention's effect size, assuming a normal distribution.

4.1.1.2 Estimating intervention costs

Both industry and government costs were estimated. Industry costs were taken from Collins et al. who estimated the cost-effectiveness of different policies aimed at reducing salt

consumption and CHD in England.[219] Collins et al. used salt reformulation costs of either £0 or £25,000 per product (£29,953 at 2014 prices). The lower estimate is based on food products being reformulated as part of the natural product cycle and therefore absorbed within the food industry's running costs, and the higher estimate is from the FSA's Impact Assessment 2009, which also suggests that up to 20,000 products would need to be reformulated.[220] The effect on results of not including industry costs is explored as a sensitivity analysis (appendix five).

Government costs were split into monitoring costs and administrative costs. Monitoring costs were estimated based on the current costs to PHE of urinary sodium surveys taking place every other year to monitor salt consumption and Kantar Worldpanel purchasing data for foods to monitor their salt content. In 2014, the urinary sodium survey in England cost £327,289, and the annual cost of the Kantar dataset in 2016 was £94,100 (£91,588 in 2014 prices).[221]

Government administrative costs were calculated using the WHO costing tool for prevention and control of NCDs (WHO NCD costing tool) following methods used by Webb et al.[66,222] The WHO NCD costing tool uses country specific resource unit costs taken from the WHO-CHOICE database[146] and estimates the country specific inputs required for a range of national level interventions tackling tobacco consumption, alcohol misuse, poor diets, and physical inactivity. The amounts and types of resources required for each intervention vary across four phases of implementation: planning (year one), development (year two), partial implementation (years three to five), and full implementation (years six to 10). For PRIMEtime CE, the annual resources required for administering the salt intervention were based on programme management and meeting resources during the full implementation phase used by Webb et al. when they estimated the cost of the UK government's sodium reduction strategy (see eTable 2 from Webb et al. supplemental materials for the full list of resources[66]). The administrative inputs used were multiplied by the UK specific unit costs from the WHO NCD costing tool, with costs inflated from 2008

to 2014 costs using the World Bank GDP deflator figures, and adjusted for England by scaling results by the ratio of the English population to the UK population (assuming that administrative costs are proportional to the size of the population).

There are no empirical estimates of uncertainty for industry or government costs and, as with healthcare and social care costs (see section 3.1.2.5), they are assumed to be "moderately uncertain" as defined by Blakely et al.[114] Therefore industry costs and government costs are separately allowed to vary in PRIMEtime CE according to a gamma distribution based on a normal distribution with a standard deviation of 10%.

4.1.1.3 Other intervention inputs

In PRIMEtime CE, salt consumption is estimated to affect IHD (ICD-10 I20-I25) and stroke (I60-I69) incidence via changing blood pressure. The parameter describing the relationship between salt and blood pressure is taken from a meta-analysis of randomised controlled trials and estimates a 5.8mmHg fall in systolic blood pressure for a 100mmol reduction 24hr urinary sodium excretion (equivalent to 5.9g salt [2,360mg sodium] consumption) after adjusting for age and baseline blood pressure.[223] For estimating the effect of a change in blood pressure on CVD incidence, age-specific parameters were taken from the prospective studies collaboration.[224] Uncertainty in all these parameters, as reported by the source articles, is included in PRIMEtime CE.

The intervention was assumed to be introduced evenly over the three years between 2014 and 2017, with a third of the salt reduction and the industry costs occurring in each year. Full implementation in year one is explored as a sensitivity analysis.

Age and sex specific dietary data were extracted from NDNS using R x64 v3.2.0,[225] and were analysed using Microsoft Excel 2013.[135]

4.1.2 Results

By 2017, the intervention was estimated to reduce salt consumption in the population by an average of 1.0g salt (398mg sodium) per day, with males having a larger reduction in

consumption than females (1.1g [457mg] and 0.9g [340mg] respectively). For the modelled adult population aged 15 years and above, the average fall in salt consumption was 1.0g salt (409mg sodium) per day (1.2g [477mg] among men, 0.9g [344mg] among women). These declines in salt consumption would lead to an average reduction in systolic blood pressure of 1.2mmHg among men and 0.8mmHg among women. For both men and women, the largest falls in salt consumption and blood pressure were among 18-34 year olds. Baseline and post-intervention salt consumption by age and sex is shown in table 4.2, and figures 4.1a and 4.1b.

Table 4.2 Sodium and salt consumption as measured by the National Diet and Nutrition Survey, four year rolling programme pre- and post-intervention

Age and sex group	**Baseline** (SE)		**Intervention** (SE)		**Difference** (SE)	
	Sodium (mg)	Salt (g)	Sodium (mg)	Salt (g)	Sodium (mg)	Salt (g)
0-3 male	1,314 (70)	3.3 (0.2)	1,078 (55)	2.7 (0.1)	237 (0.5)	0.6 (0.04)
4-12 male	1,982 (55)	5.0 (0.1)	1,596 (43)	4.0 (0.1)	387 (0.6)	1.0 (0.03)
13-17 male	2,604 (109)	6.5 (0.3)	2,078 (86)	5.2 (0.2)	526 (0.9)	1.3 (0.06)
18-34 male	2,768 (145)	6.9 (0.4)	2,201 (115)	5.5 (0.3)	566 (1.1)	1.4 (0.08)
35-54 male	2,528 (79)	6.3 (0.2)	2,073 (63)	5.2 (0.2)	455 (0.8)	1.1 (0.04)
55+ male	2,354 (74)	5.9 (0.2)	1,931 (60)	4.8 (0.1)	423 (0.7)	1.1 (0.04)
0-3 female	1,271 (73)	3.2(0.2)	1,047 (58)	2.6 (0.1)	224 (0.5)	0.6 (0.04)
4-12 female	1,790 (48)	4.5 (0.1)	1,446 (38)	3.6 (0.1)	343 (0.5)	0.9 (0.03)

13-17 female	1,931 (74)	4.8(0.2)	1,551 (58)	3.9 (0.1)	380 (0.7)	0.9 (0.04)
18-34 female	2,076 (93)	5.2 (0.2)	1,663 (74)	4.2 (0.2)	413 (0.8)	1.0 (0.05)
35-54 female	1,938 (50)	4.8 (0.1)	1,615 (41)	4.0 (0.1)	323 (0.5)	0.8 (0.03)
55+ female	1,869 (58)	4.7 (0.1)	1,556 (47)	3.9 (0.1)	313 (0.5)	0.8 (0.03)

SE, standard error

Figure 4.1a Salt consumption at baseline and following the intervention, males

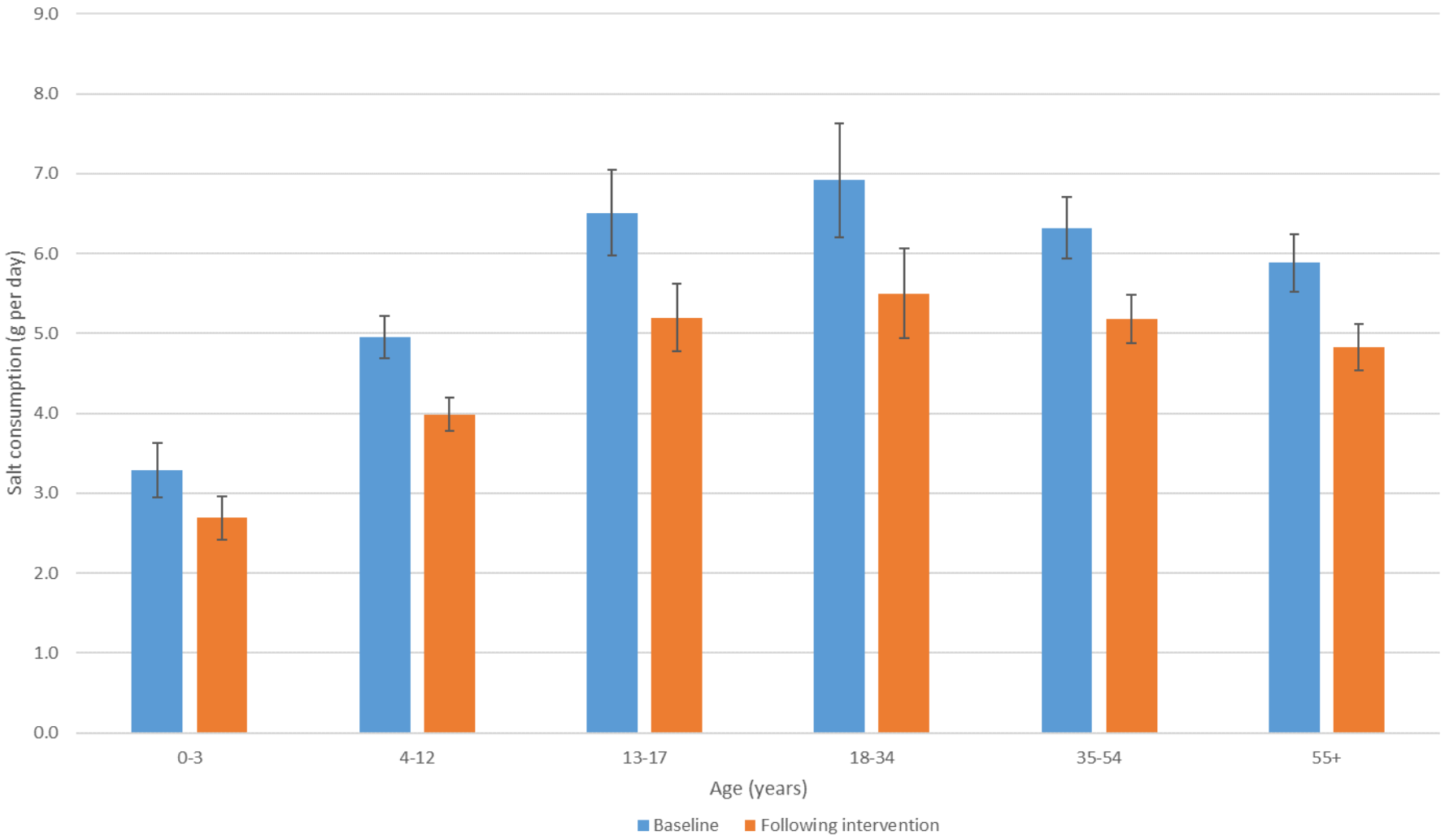
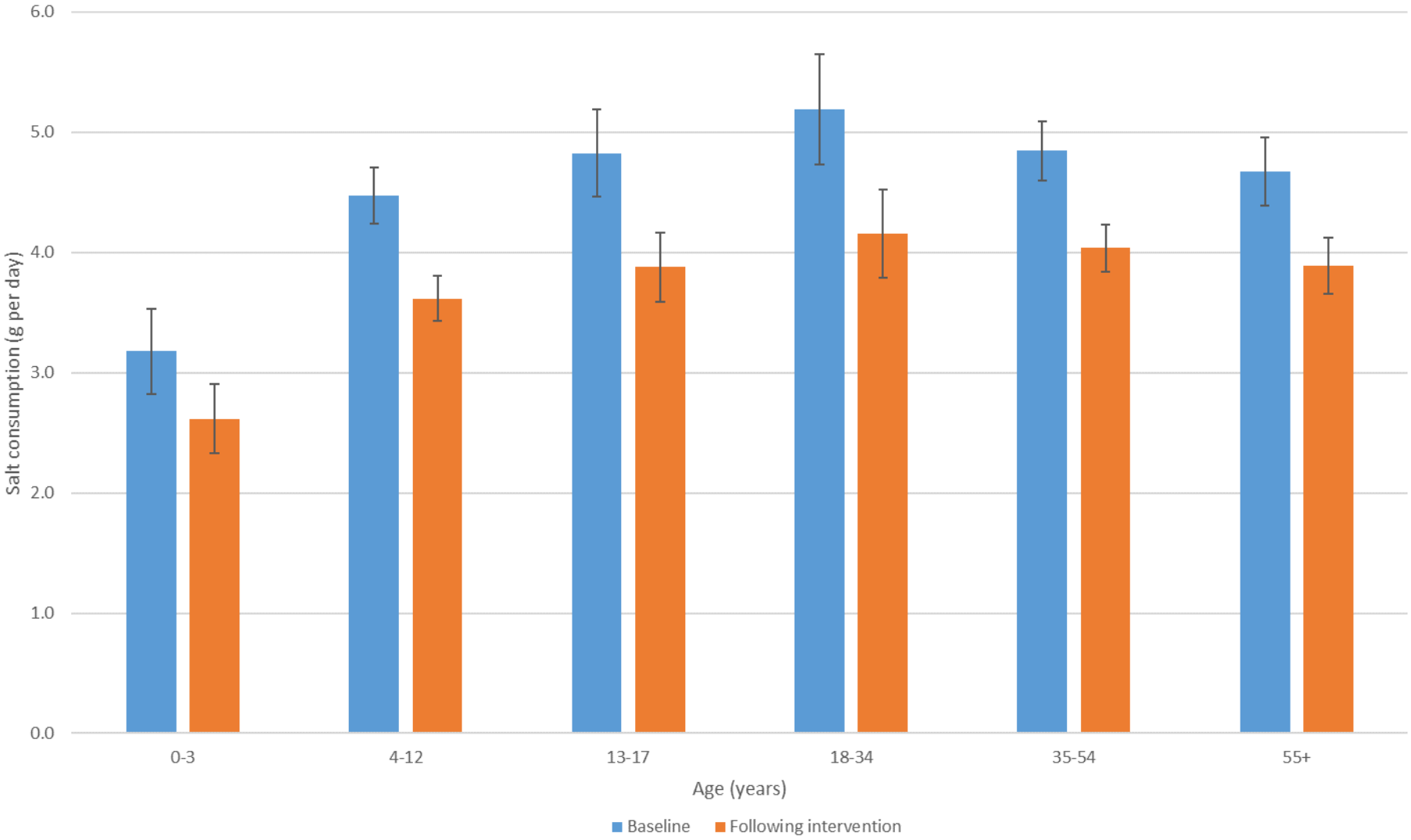


Figure 4.1b Salt consumption at baseline and following the intervention, females



Total industry costs included in PRIMEtime CE were £599m (20,000 products reformulated at £29,953 per product). Annual government monitoring costs were estimated to be £255,233 based on annual Kantar data for monitoring and for a biannual sodium survey; annual government administration costs were £570,892 per year.

4.1.3 Discussion

The diet intervention is estimated to result in an average reduction in salt consumption of 1.0g salt (409mg sodium) per person per day among adults, with greater declines among men rather than women. This is equivalent to an 18% fall in salt consumption as recorded by NDNS food diaries and would, on average, reduce systolic blood pressure by 1.0mmHg.

4.1.3.1 Strengths and limitations

The main strengths in this method are that it uses the NDNS four year rolling programme, a representative sample of dietary habits in England and that both government monitoring and administration costs of the intervention are estimated alongside industry costs.

There are some limitations, however. The effect size of the intervention is based on salt consumption recorded in four day food diaries reported between 2008 and 2012. These data suggest that average salt consumption among adults is 5.6g per day (6.4g among men and 4.9g among women). By contrast, the 2014 NDNS urinary sodium excretion survey reports that daily salt consumption among adults was 8.0g (9.1g salt consumed per day by men and 6.8g by women).[216] It is widely recognised that NDNS food diaries both under report total food consumption and also do not record salt added either at the table or during cooking.[226] If NDNS food diaries under report the amount of food (and therefore salt) consumed, then modelled results may underestimate the true effect on health and cost savings. The intervention assumes that the amount of salt added at the table or during cooking does not change and the effect on health is estimated in PRIMEtime CE based on the absolute change in salt consumption rather than the relative change. In reality, the amount of salt added at home could change in either direction – reducing the amount of salt in food may lead people to add more at the table and during cooking to make up the

difference, or conversely, as added salt in food declines, people's palates may adjust meaning even less salt is subsequently added to food at home (as was the original premise of the UK salt reduction strategy[215]). Furthermore, the 2008-2012 food diary data used to estimate salt consumption by age and sex are assumed to apply to the 2014 population, assuming that salt consumption did not change during this time. This is likely to overestimate the effect of the intervention as it fails to account for declining trends in salt consumption.[216]

The intervention also assumes that at home salt targets were applied to all food consumed, including food eaten out. This is because it was not possible to distinguish between at home and out of home food consumption by subsidiary food group category, and NDNS data was not analysed at the individual food level. In 2014 it was estimated that 11% of total energy consumed was from eating out[227] and out of home food generally contains more salt (plus the out of home salt targets are higher) meaning that modelled results may over-estimate health benefits.

Regarding how the salt targets were applied to NDNS data, 19 of the 66 salt targets are maximum thresholds for a given food group rather than an average. Where only maximum thresholds were available, they were applied as an average salt target to the relevant NDNS subsidiary food groups meaning that final results may underestimate the reduction in salt consumption from these food groups. Similarly, for the 19 NDNS subsidiary groups where more than one salt target applied, an average of the salt targets was used which may either under or overestimate salt reduction in those categories.

Finally, the intervention does not account for any of the potential administrative and trade complexities surrounding food legislation. For example, costs to industry of reformulation may be passed on to the consumer in the form of price rises. This could lead to further reductions in salt consumption if people purchase fewer foods with higher prices, and to increases in purchases of other substituting foods. Similarly it may be difficult to both

measure and enforce salt reductions among imported foods, leading to a reduction in the possible effect size of the intervention and higher administration costs.

4.1.3.2 Comparisons with other studies

Webb et al. also estimated the impact on health of a national government policy to reduce salt consumption, including industry agreements, monitoring, and a media campaign.[66] When estimating costs to government of the diet intervention modelled in this thesis, government administrative costs were based on Webb et al.'s unit resources required for programme management and meetings. However, resources costed by Webb et al. for monitoring of the intervention and for a public health media campaign were not included because the monitoring costs obtained from PHE are assumed to be more accurate, and the media campaign is outside the scope of the modelled intervention in this thesis. Rather than using resource requirements reported by the WHO NCD costing tool at the planning stage or early implementation phase for estimating government administrative costs,[222] resources needed at the full implementation phase were used because the Responsibility Deal was established in 2011 and was assumed to be fully implemented by 2014, when the diet intervention is assumed to start. Furthermore, there was little difference in costs during the full implementation phase (£570,892) and during the planning phase, where costs were highest (£581,345). Not including government administration costs, along with removing industry costs, is included as a sensitivity analysis (similar to Collins et al. who only included government monitoring costs in some of their analyses).[219]

Other studies to have investigated the health and economic impact of salt reformulation strategies in the UK include a report for Public Health England produced by the UK Health Forum,[115] and a modelling study by Collins et al. using the IMPACT CHD model.[219] Both are discussed in more detail in section 6.4.2 as part of PRIMEtime CE's validation.

Both PRIMEtime CE and the three studies mentioned above use He et al.'s meta-analysis of randomised controlled trials to estimate the impact of changing salt consumption on blood pressure,[223] and Lewington et al. to model how a change in blood pressure affects

CVD.[224] He et al. included 34 trials with a total of 3,230 participants and found that reducing salt consumption leads to falls in blood pressure (with reductions still seen at just 3g salt per day). They also found that falls in blood pressure were larger for those with hypertension than those without, and increased with age.[223] Lewington et al. analysed prospective cohort data on 958,074 individuals and reported a log-linear relationship between vascular mortality and increasing systolic blood pressure from a baseline of 115mmHg, with greater relative benefits occurring in younger adults. However, recent observational data conflict with these findings and suggest that there may be a U-shaped relationship between salt consumption and CVD.[228] Mente et al. pooled data on 133,118 individuals from four cohort studies to estimate the relationship between 24 hour urinary sodium excretion and a composite outcome of death and major cardiovascular events. The authors found that among individuals without hypertension, when compared to those consuming between 4g and 5g of sodium per day (equivalent to 10g-12.5g salt), higher sodium consumption was not associated with increased risk of death and major cardiovascular events whereas consumption of less than 3g sodium per day (7.5g salt) was associated with increased risk of disease (hazard ratio [HR] 1.26, 95% confidence interval 1.10 to 1.45). For those with hypertension, as well as significant risk of death and major cardiovascular events if consuming less than 3g sodium (HR 1.34, 1.23 to 1.47), there was increased risk if consuming more than 7g sodium (17.5g salt, HR 1.23, 1.11 to 1.37). If the relationship between sodium consumption and CVD is U-shaped, as suggested by Mente et al., then the health benefits and savings of the diet intervention simulated in this thesis would be much smaller and potentially harmful. However, several studies and commentaries suggest that the U-shaped relationship may in fact be a consequence of measurement bias due to the lack of repeated urinary sodium measurements among study participants, and other limitations such as residual confounding and reverse causality (for example, resulting from participants with low sodium consumption suffering from chronic renal failure).[229–231] Furthermore, UK and other international dietary recommendations

based on systematic reviews of the literature still suggest that salt consumption should be 6g per day or lower.[213,232,233]

4.1.3.3 Conclusion

Reformulating food such that it meets the 2017 salt targets has the potential to significantly reduce salt consumption in England. Results of the cost-effectiveness of this intervention are presented in chapter five, along with results of the sensitivity analyses.

4.2 Physical activity intervention

4.2.1 Background and methods

In 2009, Birmingham City Council and local Birmingham PCTs introduced 'Be Active', a city-wide scheme giving over one million Birmingham City residents free access to local leisure centres at certain times of the day.[144,234] Initially started as a pilot in the Ladywood Constituency of Birmingham in 2008, Be Active offers free swimming, fitness classes, and gym sessions for those who pay council tax to Birmingham City Council.

This thesis estimates the economic and health consequences of introducing the Be Active scheme across England, assuming that no other parts of England already have a similar scheme in place. Data from a health-economic evaluation of Be Active by Frew et al. was used to estimate its impact on population physical activity levels and intervention costs.[144] Frew et al. estimate the proportion of residents aged over 16 years participating in Be Active and their pre- and post-intervention levels of physical activity (table 4.3). Physical activity levels were measured among a sample of individuals signing up to the scheme over an eight-week period between 21 June, 2010 and 13 August, 2010. All individuals enrolling during this time across 19 participating leisure centres (purposively selected to try and ensure that the participating sample was representative of the population) were invited to complete the Godin Leisure Time Exercise Questionnaire[235] at enrolment to measure pre-intervention physical activity levels, and again three to four months later for post-intervention levels (this assumes that leisure centre attendance patterns are fixed by three months). In total,

2,163 (85% of participants) completed the baseline questionnaire, 1,156 (49%) completed the follow-up, and 797 (31%) completed both. The Godin Leisure Time Exercise Questionnaire asks how many times on average in the past week someone spent doing 30 minutes of either mild, moderate, or strenuous exercise, and provides examples of exercises that fall within each category. These data were then used to calculate the total number of metabolic equivalent of task (MET) minutes per week for each participant. A MET is defined as consuming 3.5mm of oxygen per kg of body weight, the resting metabolic rate. Levels of physical exertion can be measured in relation to this resting metabolic rate, for example running at seven miles per hour is equivalent to 11 METs.[236] Physical activity levels for the 797 participants who completed both questionnaires were calculated by Frew et al. by assuming that strenuous activity is equivalent to 6-10 METs, moderate activity 3-5.9 METs, and mild activity <3 METs. From the resulting number of MET minutes per week, the participating population was categorised into three groups: inactive (from here onward labelled as under active in order to distinguish them from the sedentary population), active, and recommended, and assumed to be representative of the total population enrolled in the scheme (table 4.3). Under active corresponds to less than 60 minutes of moderate physical activity per week, active is equivalent to between 60 and 150 minutes per week, and recommended levels are based on the UK Chief Medical Officers' Guidelines for adults of 150 minutes or more per week, assuming moderate physical activity corresponds to a minimum of three METs.[237]

Table 4.3 Reported physical activity levels among those participating in Be Active pre and post-intervention [144]

Physical activity level	MET mins/week	MET hours/week	Baseline % of Be Active population	Post intervention % of Be Active population
Under active	</=180	</=3	19%	10%
Active	181-449	3-7.5	21%	15%
Recommended	>/=450	>/=7.5	60%	76%

Numbers may not sum to 100% due to rounding. MET, metabolic equivalent of task

A draft evaluation of the Be Active programme by the University of Birmingham, provided by Birmingham City Council (personal communication), reported the proportion of the population of Birmingham by age and sex registered in the scheme (taken from the Be Active database in 2010).[238] These data were used to estimate the proportion of the English population by five year age and sex group that enrolled in the intervention. The ratio of men to women reported by Be Active is assumed to apply equally across all age groups. Using the 2014 English population, 3,764,078 adults were estimated to have enrolled in the intervention.[239]

Pre-intervention (baseline) physical activity levels by age and sex for England were taken from the Active People Survey, 2010-2011 (APS).[240] The APS collects information on sport participation rates from a representative English population sample and includes the degree of participation in a variety of activities such as walking and cycling, mirroring the questions asked in the Be Active economic evaluation. Physical activity data were available for 166,275 adults aged 16 years and over after outliers reporting over 200 MET hours/week were removed. The APS from 2010-2011 was used because the raw data for this year were available and the computer code to calculate physical activity rates in MET hours/week for each individual was already written based on 2010-2011 APS variables (calculated by applying MET values provided by APS to the activities reported). Adults reporting zero MET hours/week were classed as sedentary. The percentage of the population in each physical activity category (sedentary, under active, active, and recommended) was then derived for each age and sex group.

Following the intervention, those enrolled in the programme changed physical activity category as shown in table 4.3. The percentage of each age and sex group in each physical activity category following the intervention was calculated as follows. Firstly, the percentage of each age and sex category that was classed as sedentary by APS was assumed not to change post-intervention. For the remaining active population, the percentage of those enrolled in the intervention (calculated using data from Birmingham City Council[238] and

described above) that fell into each physical activity category pre- and post-intervention was taken from table 4.3, with the same percentages used for each age and sex group. The percentage of the total active population in England in each physical activity category following the intervention was then calculated by adding the difference in the number of people in a given physical activity category pre- and post-intervention to the baseline number of people in that category. This was repeated for each age and sex group.

As an example, it was estimated from APS that 20% of males aged 50-54 (374,720 men) are sedentary, and that of the remaining active population, 55% (810,950) achieve recommended levels of physical activity. Using data from Frew et al., it was assumed that 8% (159,930) of this age and sex group enrolled in the intervention of which 60% (88,920) achieved recommended levels of activity at enrolment (pre-intervention), rising to 76% (112,880) following the intervention. Therefore, following the intervention, there were an additional 23,960 men aged 50-54 years achieving recommended levels of physical activity, and adding this to the baseline 810,950 men, the final number of men aged 50-54 years achieving recommended levels of physical activity was 834,910, or 45% of the age and sex group. For full results, see table 4.4.

The relationships between physical activity and IHD, stroke, type two diabetes, breast cancer, and colorectal cancer were added to PRIMEtime CE. These diseases were included because they were prioritised by stakeholders (see appendix three), they were already in PRIMEtime, and there are meta-analyses of prospective observational studies describing their relationship with physical activity.[241–243] The same risk factor-disease pairs are used by numerous other studies modelling the health outcomes of physical activity interventions.[60,73,85,144,244,245] In PRIMEtime CE, the beta-coefficient describing the dose-response relationship between physical activity and CVD, diabetes, breast cancer, and colorectal cancer was derived using two recent meta-analyses of observational studies, both conducted by the same research group using the same methodology.[242,243] The parameters were unadjusted for obesity as this is assumed to act on the causal pathway. The mean and

standard error MET hours/week was calculated from APS for each age and sex group for individuals in the under active, active, and recommended physical activity categories. These were applied to results from the meta-analyses to estimate the relative risk of disease for each physical activity category compared to being sedentary. These relative risks were then used in PRIMEtime CE to predict the change in disease burden following the intervention. Furthermore, in PRIMEtime CE decreasing the prevalence of type two diabetes reduces the risk of the modelled cohort developing IHD and stroke. Using the relationships describing the effect of physical activity on IHD, stroke, and diabetes derived using Wahid et al. would over-estimate the effect of physical activity on IHD and stroke as falls in diabetes would result in additional CVD reductions beyond those estimated in the meta-analyses. Therefore, the relative risk of physical activity on IHD and stroke from Wahid et al. was adjusted downwards using results from Cobiac and Scarborough such that the overall effect of physical activity on CVD is the sum of the direct effect plus that mediated via diabetes.[137]

Uncertainty in the size of the relationship between physical activity and disease (the beta-coefficient) was estimated from 95% confidence intervals reported in the meta-analyses. Both this uncertainty and uncertainty in the mean physical activity level for each physical activity category based on the standard error reported in APS were included in PRIMEtime CE. Finally, uncertainty in the number participating in each age and sex group was estimated using the same approach used for health and societal costs (see section 3.1.2.5) and taken from Blakely et al., assuming that the number participating is a "moderately uncertain" variable with a gamma distribution based on a normal distribution with a standard deviation of 10%.[114]

Age and sex specific population physical activity levels were analysed using STATA 14,[246] pre- and post-intervention physical activity distributions were calculated using Microsoft Excel 2013.[135]

Intervention costs were estimated by Frew et al. to be £47 (£54 in 2014) per participant in year one and £26 (£30 in 2014) per participant in year two onwards, each was assumed to

have a triangular distribution.[144] Intervention costs were calculated by dividing the total council spend on the scheme (including income replacement for the gyms, refurbishment costs, marketing, monitoring, and project management) by the number of participants.

4.2.2 Results

The percentage of men achieving recommended levels of physical activity following the intervention was estimated to increase from 46.4% to 47.7% alongside a decrease among under active men from 17.4% to 16.6%. Among women, there was a slightly higher absolute and relative increase in the percentage achieving recommended levels of physical activity (38.6% to 40.0%) with an associated decrease in under active women from 19.4% to 18.6%.

Younger adults had lower percentages of sedentary individuals and higher percentages of individuals achieving recommended levels of physical activity than older adults, with the same pattern seen for men compared with women. As a result of higher percentages of younger people and women enrolling in the intervention, the absolute percentage gains in those achieving recommended physical activity levels were higher in these population groups.

The percentage of the population that was sedentary, under active, active, and achieving recommended levels of physical activity by age and sex pre- and post-intervention is shown in table 4.4. Mean physical activity levels among the active population shown in table 4.4 are for illustrative purposes only and are not used by PRIMEtime CE. They are calculated assuming that physical activity levels follow a log-normal distribution (the distribution with the best fit compared to gamma, Weibull, normal, uniform, exponential, and logistic when tested using the Bayesian Information Criterion). They estimate that on average following the intervention, non-sedentary men would increase their physical activity by 1.7 MET hours/week (1.3 MET hours/week across all men), and non-sedentary women would increase their physical activity levels by 0.5 MET hours/week (0.3 MET hours/week across all women).

Intervention costs per participant in 2014 after adjusting for inflation were £53.80 (triangular 43.50, 53.80, 85.90) in year one, and £29.80 (24.00, 29.80, 48.10) in year two onwards. Total intervention costs were therefore estimated to be £202.6m (163.8, 202.6, and 323.3) in year one and £112.1m (90.5, 112.1, 181.1) in year two onwards. Per participant costs were included in PRIMEtime CE with triangular distributions used to quantify uncertainty in these estimates, as per Frew et al.[144]

Table 4.4 Percentage of the English population with different levels of physical activity by age and sex pre- and post-intervention

Age and sex group	Pre-intervention					Post-intervention				
	Sedentary	Under active	Active	Recommended	Mean*	Sedentary	Under active	Active	Recommended	Mean*
	0 MET hrs/wk	<3MET hrs/wk	3-7.5 MET hrs/wk	>7.5 MET hrs/wk	MET hrs/wk	0 MET hrs/wk	<3MET hrs/wk	3-7.5 MET hrs/wk	>7.5 MET hrs/wk	MET hrs/wk
M16-19	6.2%	9.0%	9.9%	74.9%	38.8	6.2%	7.9%	9.1%	76.8%	45.1
M20-24	9.9%	10.7%	10.7%	68.7%	30.3	9.9%	9.6%	9.9%	70.6%	35.4
M25-29	12.0%	13.9%	11.5%	62.6%	23.6	12.0%	12.7%	10.8%	64.6%	27.6
M30-34:	12.7%	15.7%	13.5%	58.1%	16.7	12.7%	14.5%	12.8%	60.0%	16.8
M35-39	14.4%	17.6%	14.9%	53.1%	12.9	14.4%	16.6%	14.1%	54.9%	14.5
M40-44	15.2%	18.5%	15.2%	51.1%	11.9	15.2%	17.4%	14.4%	53.0%	12.9
M45-49	16.8%	19.3%	14.8%	49.1%	11.4	16.8%	18.5%	14.3%	50.4%	12.2
M50-54	20.4%	20.3%	15.3%	44.1%	9.4	20.4%	19.5%	14.7%	45.4%	8.0
M55-59	24.4%	21.1%	14.0%	40.5%	8.8	24.4%	20.5%	13.6%	41.5%	8.8
M60-64	26.4%	20.6%	13.1%	39.8%	9.1	26.4%	20.0%	12.7%	40.9%	9.8

M65-69	27.9%	18.8%	14.0%	39.3%	9.1	27.9%	18.5%	13.9%	39.7%	9.3
M70-74	32.4%	18.0%	12.6%	37.1%	9.4	32.4%	17.7%	12.4%	37.5%	9.6
M75-79	40.9%	16.0%	10.9%	32.2%	9.2	40.9%	15.7%	10.8%	32.6%	9.6
M80-84	51.7%	14.4%	9.6%	24.3%	7.6	51.7%	14.1%	9.5%	24.7%	7.9
M85+	48.6%	13.8%	9.9%	27.7%	8.9	48.6%	13.5%	9.8%	28.1%	9.2
Overall	23.0%	17.4%	13.2%	46.4%	16.8	23.0%	16.6%	12.6%	47.7%	18.5
F16-19	13.8%	17.4%	16.0%	52.8%	12.1	13.8%	16.2%	15.1%	54.9%	13.7
F20-24	16.9%	16.9%	16.7%	49.5%	10.9	16.9%	15.7%	15.8%	51.6%	11.6
F25-29	14.4%	19.5%	17.9%	48.2%	9.6	14.4%	18.3%	17.1%	50.2%	9.6
F30-34	14.0%	19.8%	17.9%	48.3%	9.6	14.0%	18.6%	17.0%	50.3%	9.7
F35-39	14.7%	21.6%	18.0%	45.7%	8.6	14.7%	20.5%	17.3%	47.6%	9.5
F40-44	14.7%	20.9%	17.8%	46.6%	9.0	14.7%	19.7%	17.1%	48.5%	10.0
F45-49	16.9%	20.5%	16.3%	46.3%	9.6	16.9%	19.7%	15.8%	47.6%	9.7
F50-54	19.9%	20.7%	17.2%	42.2%	8.3	19.9%	19.9%	16.7%	43.5%	8.8
F55-59	21.8%	22.0%	17.8%	38.4%	7.2	21.8%	21.4%	17.4%	39.4%	7.6

F60-64	24.1%	21.1%	18.1%	36.7%	7.1	24.1%	20.4%	17.7%	37.8%	7.4
F65-69	28.2%	20.5%	16.4%	34.9%	7.1	28.2%	20.3%	16.2%	35.3%	7.3
F70-74	36.3%	19.3%	15.2%	29.1%	6.4	36.3%	19.1%	15.1%	29.5%	6.5
F75-79	46.1%	17.6%	13.4%	22.8%	5.7	46.1%	17.4%	13.3%	23.2%	5.8
F80-84	62.2%	9.5%	11.3%	17.0%	6.5	62.2%	9.3%	11.2%	17.3%	6.7
F85+	52.2%	14.1%	11.4%	22.2%	6.6	52.2%	13.8%	11.3%	22.6%	6.8
Overall	25.6%	19.4%	16.4%	38.6%	9.1	25.6%	18.6%	15.8%	40.0%	9.6

*This is the mean MET hours per week among the non-sedentary population, these are for illustrative purposes and are not used by PRIMEtime CE

MET hrs/wk, metabolic equivalent of task hours per week

4.2.3 Discussion

The physical activity intervention is estimated to lead to absolute percentage increases of 1.3% among men and 1.4% among women in the number of people achieving recommended levels of physical activity. These increases, alongside corresponding decreases in the percentage of people who are in under active and active physical activity categories, result in an average population increase of approximately 1.3 MET hours/week among men, and 0.3 MET hours/week among women, equivalent to 11 minutes and 3 minutes of jogging per week respectively.[236] Although absolute percentage increases in those achieving recommended levels of physical activity are greater among women than men, average baseline MET hours/week are smaller among women (particularly among those achieving recommended levels), meaning that the overall average increase in MET hours/week is greater among men than women.

4.2.3.1 Strengths and limitations

The strengths of using the described method to estimate the effect of expanding the Be Active scheme across England include using an empirical evaluation of the scheme to parameterise the change in physical activity levels and uptake rates by age and sex. Baseline data on physical activity levels in England are taken from the APS. This is the largest representative survey of adult physical activity in England meaning that physical activity levels can be estimated by five-year age group and by sex. The HSE also reports self-assessed physical activity levels however their sample size is smaller (8,291 adults included compared to 166,275 in APS) and the questionnaire used by APS is more representative of the physical activity data collected in the Be Active evaluation.[247] Other strengths of the methods used in this thesis include adjusting the relationships between physical activity and CVD to avoid over-estimating effects from diabetes, and quantifying the uncertainty associated with input parameters including baseline physical activity levels, the relationships between physical activity and disease, the number participating in the intervention, and intervention costs.

There are also some important methodological limitations. The comparator for the intervention is 'no Be Active scheme', therefore assuming that similar schemes are not in place elsewhere in England. If other schemes are in place, then the overall effect size may be overestimated. Furthermore, the intervention assumes that the scheme can be replicated elsewhere in England with the same age and sex specific uptake rates. Be Active is based in an urban environment with a high density of local authority leisure facilities. More rural settings without such well-established networks of leisure centres may not have the infrastructure to either implement the scheme or support similar uptake rates, meaning overall effect sizes may be overestimated. Similarly, the cost of the intervention to local authorities may vary by region and rather than using one intervention cost estimate from Frew et al.,[144] a bottom up costing exercise across multiple and varied English local authorities may provide a better estimate. Reported Be Active uptake rates also varied by socio-economic group and by ethnicity.[238] At the time of the intervention, 19.8% of the population in Birmingham were from the most deprived quintile and 20.6% from the least deprived, however 23.2% in the most deprived quintile compared to 17.9% in the least deprived registered for the scheme. Therefore Be Active has the potential to reduce socioeconomic inequalities. Uptake by ethnicity suggested that non-white individuals were more likely to participate than white. However, results of the intervention were not adjusted for SES or ethnicity because participation rates for these subgroups were not available by age or by sex. Also, uncertainty in the percentage of individuals in each physical activity category pre- or post-intervention was not explicitly included in the model. This uncertainty is in part captured by allowing the mean physical activity level in each category to vary for each age and sex group according to its standard error. However, overall uncertainty in the model outcome may be underestimated.

Physical activity levels recorded in both the APS and the Be Active evaluation are self-reported and may not accurately represent pre- and post-intervention physical activity rates. Given that the direction of bias for the self-reported physical activity levels is unknown, reported results may under- or over-estimate the true effect of the

intervention.[248] The intervention also assumes that there is no change in the sedentary population following the intervention. If some of those engaging in the intervention are sedentary at baseline (classified within the inactive physical activity category by Frew et al.) then the effect size of the intervention may be under-estimated. Furthermore, the Be Active evaluation assumes that post-intervention physical activity rates do not change from three months post intervention, however this may overestimate the true effect size if, as with some other physical activity interventions, the effect size attenuates beyond three months post intervention.[249] This is explored in a sensitivity analysis replicating one of Frew et al.'s sensitivity analyses: half of those who move from active to recommended levels of physical activity after the intervention return to the active category after one year.[144]

In their meta-analyses of the effect of physical activity on health, Wahid et al. included domains of physical activity such as occupational activity and household activity that are not captured in the Be Active evaluation or the APS.[144,240,242,243] As such, the disease relative risks used by PRIMEtime CE are based on higher levels of physical activity than those reported by Be Active and APS where only leisure time physical activity (LTPA) is recorded. Therefore, the overall health benefit modelled by PRIMEtime CE may be an over-estimate because the effect of physical activity on health is non-linear with greater health benefits accruing from gains at lower baseline physical activity levels. Baseline physical activity data was taken from the 2010-2011 APS. According to the APS, rates of participation in active recreation three or more times a week have increased from 21.8% in 2010/11 to 23.8% in 2015/16, although any sports participation has decreased from 48.5% to 46.6% over the same period. Data on physical activity rates are only available from 2012 to 2014 and over these three years, the percentage of people inactive decreased from 29.4% to 27.7% and the percentage active increase from 54.9% to 57.0%.[250] If physical activity levels did increase between 2010/11 and 2014 (the start of the intervention), then the intervention may again over-estimate health benefits because there are greater health gains for increases in physical activity levels from a lower baseline.

Finally, PRIMEtime CE only models the relationships between physical activity and IHD, stroke, type two diabetes, colorectal cancer, and stroke. Other studies have modelled other relationships including dementia[244] and depression,[85] and there is increasing evidence of physical activity reducing the risk of other cancer subtypes.[251] Therefore, the overall health benefits of the intervention may be underestimated.

4.2.3.2 Comparisons with other studies

Various studies have investigated the relationship between physical activity and ill health. To parameterise their Markov model, Frew et al. performed systematic reviews and meta-analyses for each disease outcome and each physical activity category included in their model (under active, active, and recommended, see table 4.3), and explicitly sought studies reporting the relationship between LTPA and health.[144] This is because the Be Active evaluation is based on LTPA rather than other domains of physical activity such occupational or domestic. The meta-analyses conducted by Frew et al. were not used in PRIMEtime CE because they do not report a dose-response relationship between physical activity levels and disease. This means they are unable to estimate the effect of the intervention across different age and sex groups where average physical activity levels within each category differ. Wahid et al. do report a dose-response in effect size however, the meta-analyses also included studies measuring additional physical activity domains alongside LTPA and therefore may not reflect the health gains from LTPA alone.[242,243] More recently, Kyu et al. published a meta-analysis estimating the dose-response relationship of increasing physical activity on IHD, stroke, diabetes, colorectal cancer, and breast cancer.[241] As with Wahid et al., Kyu et al. included all domains of physical activity and found greater health gains from increasing activity at low levels. Rather than identify a distribution to describe the overall pattern of effect of changing physical activity on health (as done by Wahid et al.), the relative risk of moving away from being sedentary was reported for different categories of MET minutes/week. Kyu et al. also included significantly more studies than Wahid et al. (for example 69 studies related to CVD and 55 related to diabetes, compared to just 33 and 3 included by Wahid et al.), this was primarily because for studies

to be included by Wahid et al., at least two domains of physical activity had to be measured and reported results had to be adjusted for body weight whereas these restrictions were not used by Kyu et al.

The relative risk of disease resulting from a change in MET hours/week equivalent to moving from sedentary or under active to recommended levels was compared between Frew et al.,[144] Kyu et al.,[241] and PRIMEtime CE (using Wahid et al.[242,243]). The results are in table 4.5, alongside the change in MET hours/week used for each comparison. The relative risks reported by PRIMEtime CE are calculated before IHD and stroke relative risks are adjusted for diabetes, and are based on the average MET hours/week of a 30 year old male in each physical activity category: zero if sedentary, 1.5 if under active, and 28.1 if meeting recommended levels. For each disease, Frew et al. estimated lower relative risks of disease as a result of moving from being under active (<3 MET hours/week) to achieving recommended levels of physical activity (>/=7.5 MET hours/week) than PRIMEtime CE. In part this difference is explained by those who are sedentary being included in Frew et al.'s under active disease category but not in PRIMEtime CE, as shown in table 4.5 where the PRIMEtime CE relative risks of disease due to moving from sedentary to meeting recommendations are closer to those reported by Frew et al. Furthermore, because Wahid et al. includes non-LTPA, the total amount of physical activity (LTPA and non-LTPA) being done by someone classified as achieving recommended physical activity levels in Frew et al.'s meta-analyses will be higher than if classified by Wahid et al. Therefore the absolute disease risk will be lower and if this difference disproportionately affects those doing more physical activity then the relative risk of moving between categories will also be lower. Estimates from Kyu et al. of the relative risk of moving from sedentary to 30 MET hours/week (approximately equivalent to the 28.1 MET hours/week achieved by a 30 year old male) are not dissimilar to results using PRIMEtime CE, with confidence intervals overlapping (table 4.5). In general, slightly larger relative risk reductions are estimated using PRIMEtime CE compared with Kyu et al. (except for colorectal cancer). Therefore,

PRIMEtime CE may under-estimate the impact of the intervention on health compared with Frew et al., and over-estimate effects if Kyu et al. were used instead of Wahid et al.

Table 4.5 Relative risk of disease from moving from under active or sedentary to recommended levels of physical activity used by Frew et al., Kyu et al., and PRIMEtime CE (95% uncertainty intervals)

Source of relative risk	**Frew et al.**	**Kyu et al.**	**PRIMEtime CE**	**PRIMEtime CE**
Change in physical activity	<3 to >/=7.5 MET hrs/wk	0 to 30 MET hrs/wk	0 to >/=7.5 MET hrs/wk*	>0-3 to >/=7.5 METhrs/wk*
Ischaemic heart disease relative risk	0.50 (0.44 to 0.56)	0.78 (0.71 to 0.84)	0.71 (0.64 to 0.79)	0.83 (0.78 to 0.88)
Stroke relative risk	0.73 (0.67 to 0.87)	0.78 (0.70 to 0.89)	0.72 (0.61 to 0.85)	0.83 (0.75 to 0.91)
Type two diabetes relative risk	0.57 (0.55 to 0.59)	0.83 (0.77 to 0.90)	0.66 (0.60 to 0.74)	0.79 (0.60 to 0.85)
Breast cancer relative risk	0.81 (0.76 to 0.81)	0.95 (0.91 to 0.99)	0.92 (0.86 to 0.97)	0.96 (0.86 to 0.98)
Colorectal cancer relative risk	0.75 (0.75 to 0.75)	0.83 (0.76 to 0.91)	0.89 (0.72 to 1.08)	0.94 (0.83 to 1.04)

*Based on average physical activity levels of a 30 year old male, where sedentary is 0 MET hours/week, under active is 1.5 MET hours/week, and recommended is 28.1 MET hours/week

MET hrs/wk, metabolic equivalent of task hours per week

4.2.3.3 Conclusions

Implementing the Be Active scheme across England would lead to small, but potentially important population level gains in physical activity. The cost-effectiveness of the intervention is reported in chapter five, and differences between PRIMEtime CE and Frew et al. are explored in more detail in chapter six.

Chapter 5. Results

This chapter describes the results from the primary analysis and sensitivity analyses for each intervention – reformulating food to have less salt (the diet intervention) and increasing access to leisure centres (the physical activity intervention). The primary analysis reports results with a 10 year time horizon from an NHS and social care perspective but with industry costs included in the diet intervention (see section 2.1.2), discounted at 1.5% per year. Costs from unrelated diseases are included as are all costs of the intervention to industry and government (see table 2.2). The £30,000 cost per QALY threshold has been included in figures to illustrate how results compare to the upper end of the current NICE threshold for determining cost effectiveness.[13] Medians are presented rather than means (usually presented in economic evaluations) to better represent the skewed nature of the underlying parameter distributions. All baseline PRIMEtime CE input parameters can be found in appendix four and the interventions are described in detail in chapter four.

5.1 Reformulating food to have less salt

If the food industry in England were to reformulate food to meet the Department of Health 2017 salt targets for food purchased for consumption at home, average salt consumption would fall by 1.2g (477mg sodium) per day among men and 0.9g (344mg sodium) per day among women, resulting in reductions in systolic blood pressure of 1.2mmHg and 0.8mmHg respectively (see section 4.1.2).

5.1.1 Results of the primary analysis

Over 10 years, the intervention resulted in a median gain of 15,000 QALYS (95% UI 5,930 to 25,520), an NHS cost saving of £141.7m (£58.8m to £235.9m), and a saving from social care expenditure of £724.8m (£306.5m to £1,200m) at a total implementation cost to industry and government of £598.0m (£489.2m to £717.7m). The intervention was therefore dominant, with a median cost per QALY (including NHS costs, social care costs, industry implementation costs, and government implementation costs) of -£17,050 (-

£39,550 to £38,550; mean, -£22,490), and a 97% chance of falling below the £30,000 upper limit cost effectiveness threshold used by NICE.[13] The return on investment was £1.44 (£0.50 to £2.94) for every £1 spent. A summary of results for the primary analysis is in table 5.1, figure 5.1 illustrates the 2,000 iterations of the Monte Carlo simulation on the cost-effectiveness plane, and the cost-effectiveness acceptability curve in figure 5.2 shows the probability that the intervention is cost effective at different cost-per-QALY thresholds.

The intervention led to 24,460 (10,573 to 38,686) fewer cases of stroke and 16,570 (7,170 to 26,220) fewer cases of IHD (figure 5.3). There were small increases in other modelled diseases due to increased life expectancy following the intervention, allowing more time for other diseases to develop.

Results varied by age and sex. Men had greater health benefits per person, and in total, than women, and greater NHS and social care savings per person. Women were responsible for more social care savings due to living longer. The greatest gain in QALYs per person was among men aged 70-74 years at the start of the intervention, with an average of 126 QALYs gained per 100,000 men. Social care savings peaked at £134 per person among men aged 75-79 years, and NHS savings per person peaked at £10 per person for men aged 65-69 years. Figure 5.4 shows the change in QALYs per person and the increase in total QALYs for each age and sex group, and figure 5.5 illustrates the equivalent results for NHS and social care savings. Age specific results were based on point estimates and therefore do not include UIs. The nature of the food reformulation scenario means that the intervention and its costs affect the entire population and cannot be separated into population subgroups. Therefore cost per QALY results are not reported by sex or by age.

The intervention became cost saving nine years following its implementation, after which time NHS and social care savings outweighed industry and government costs.

Table 5.1 Summary of results of the diet intervention primary analysis

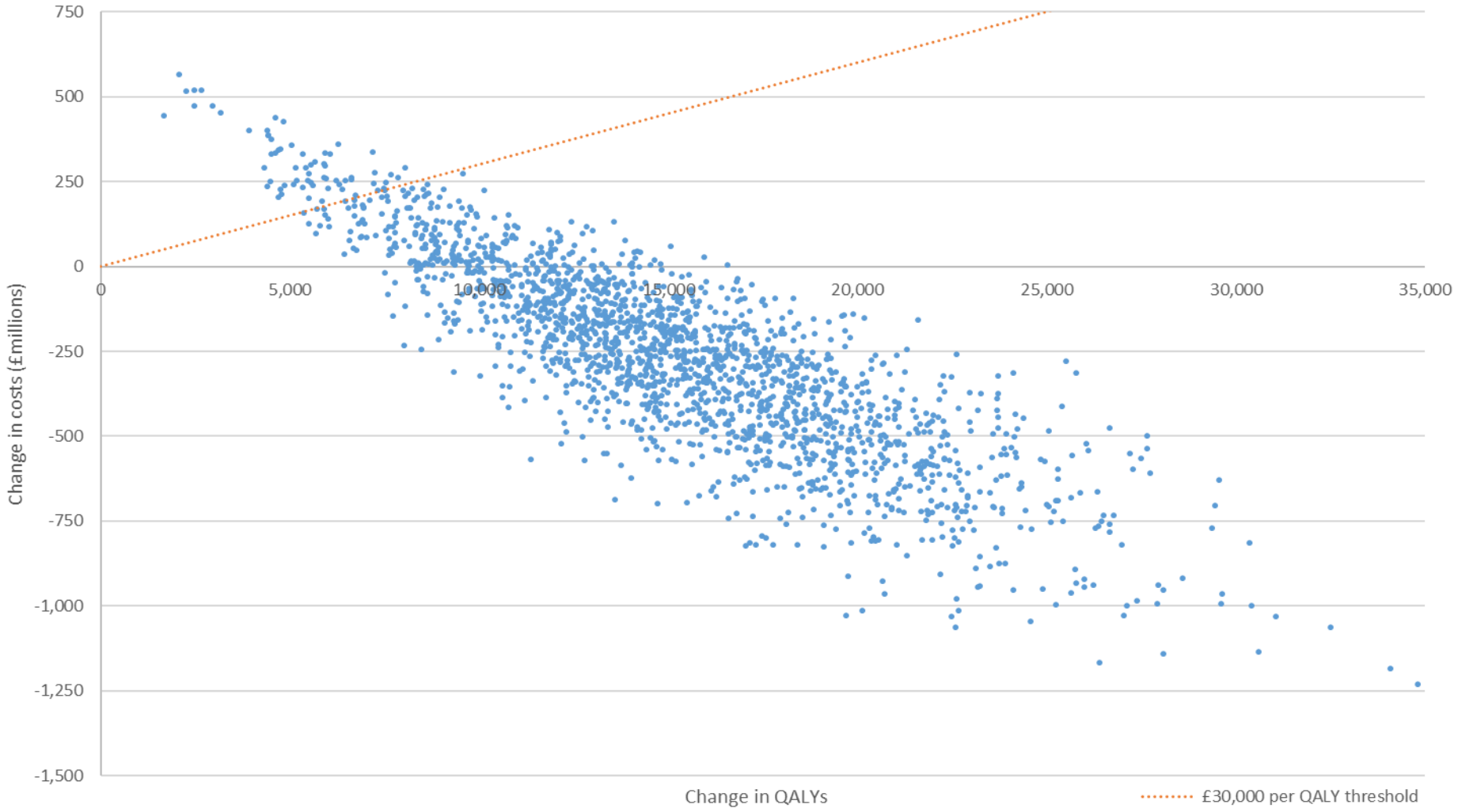
	Males	**Females**	**Total**
Cost per QALY – overall (£)*	-	-	-17,046 (-39,553 to 38,547)
Return on investment (£ for £1 spent)*	-	-	1.44 (0.50 to 2.94)
Change in QALYs	9,045 (3,629 to 15,338)	5,945 (2,320 to 10,186)	15,000 (5,929 to 25,519)
NHS savings (£)	95,623,517 (39,784,418 to 159,346,380)	45,721,114 (18,519,937 to 76,690,196)	141,701,342 (58,825,774 to 235,925,461)
Social care savings (£)	350,748,015 (148,153,324 to 582,543,734)	371,065,063 (154,269,323 to 619,047,856)	724,832,212 (306,525,399 to 1,200,201,282)
Intervention costs (£)*	598,039,731 (489,209,385 to 717,730,482)	598,039,731 (489,209,385 to 717,730,482)	598,039,731 (489,209,385 to 717,730,482)
Cost per QALY - NHS perspective (£)*	-	-	30,547 (13,898 to 90,823)
Cost per QALY - social care perspective (£)*	-	-	-7,824 (-29,778 to 48,329)
Change in IHD cumulative incidence [incidence rate per 100,000]	-11,637 [-53.3] (-18,375 to -5,027)	-4,927 [-21.6] (-7,819 to -2,110)	-16,568 [-37.1] (-26,223 to -7,167)
Change in stroke cumulative incidence [incidence rate per 100,000]	-13,840 [-63.4] (-21,855 to 5,964)	-10,606 [-46.5] (-16,666 to -4,545)	-24,460 [-54.8] (-38,686 to -10,573)

Change in type two diabetes cumulative incidence [incidence rate per 100,000]	54 [0.2] (24 to 86)	47 [0.2] (20 to 73)	101 [0.2] (44 to 158)
Change in breast cancer cumulative incidence [incidence rate per 100,000]	0 [0]	16 [0.1] (7 to 24)	16 [0.0] (7 to 24)
Change in colorectal cancer cumulative incidence [incidence rate per 100,000]	20 [0.1] (8 to 31)	10 [0.0] (4 to 16)	30 [0.1] (13 to 47)
Change in lung cancer cumulative incidence [incidence rate per 100,000]	22 [0.1] (9 to 34)	11 [0.0] (4 to 17)	32 [0.1] (14 to 50)
Change in stomach cancer cumulative incidence [incidence rate per 100,000]	4 [0.0] (2 to 6)	2 [0.0] (1 to 3)	6 [0.0] (2 to 9)
Change in pancreas cancer cumulative incidence [incidence rate per 100,000]	3 [0.0] (1 to 5)	3 [0.0] (1 to 4)	6 [0.0] (3 to 10)
Change in kidney cancer cumulative incidence [incidence rate per 100,000]	4 [0.0] (2 to 5)	2 [0.0] (1 to 2)	5 [0.0] (2 to 8)
Change in liver cancer cumulative incidence [incidence rate per 100,000]	2 [0.0] (1 to 3)	1 [0.0] (0 to 2)	3 [0.0] (1 to 5)
Change in liver cirrhosis cumulative incidence [incidence rate per 100,000]	3 [0.0] (1 to 5)	2 [0.0] (1 to 3)	5 [0.0] (2 to 7)

*The intervention and its costs affect the entire population and cannot be targeted such that only one subgroup is affected. Therefore cost per QALY results are not reported by sex

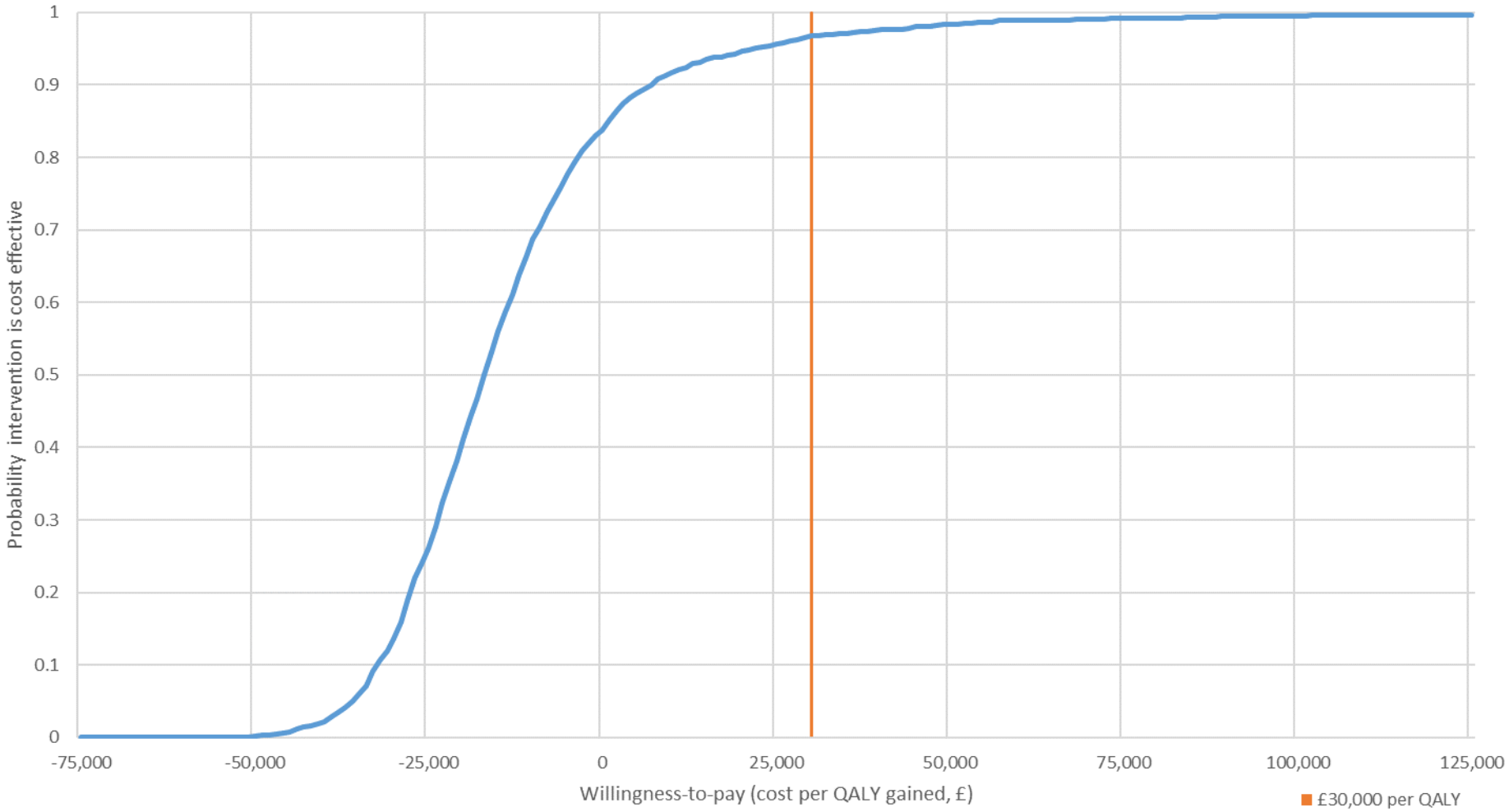
Results presented are the median values from 2,000 iterations of a Monte Carlo simulation, as such, the numbers in the final column may not equal the sum of males and females; 95% uncertainty intervals in parentheses; QALY, quality adjusted life year; IHD, ischaemic heart disease

Figure 5.1 Cost effectiveness plane showing results of the diet intervention primary analysis



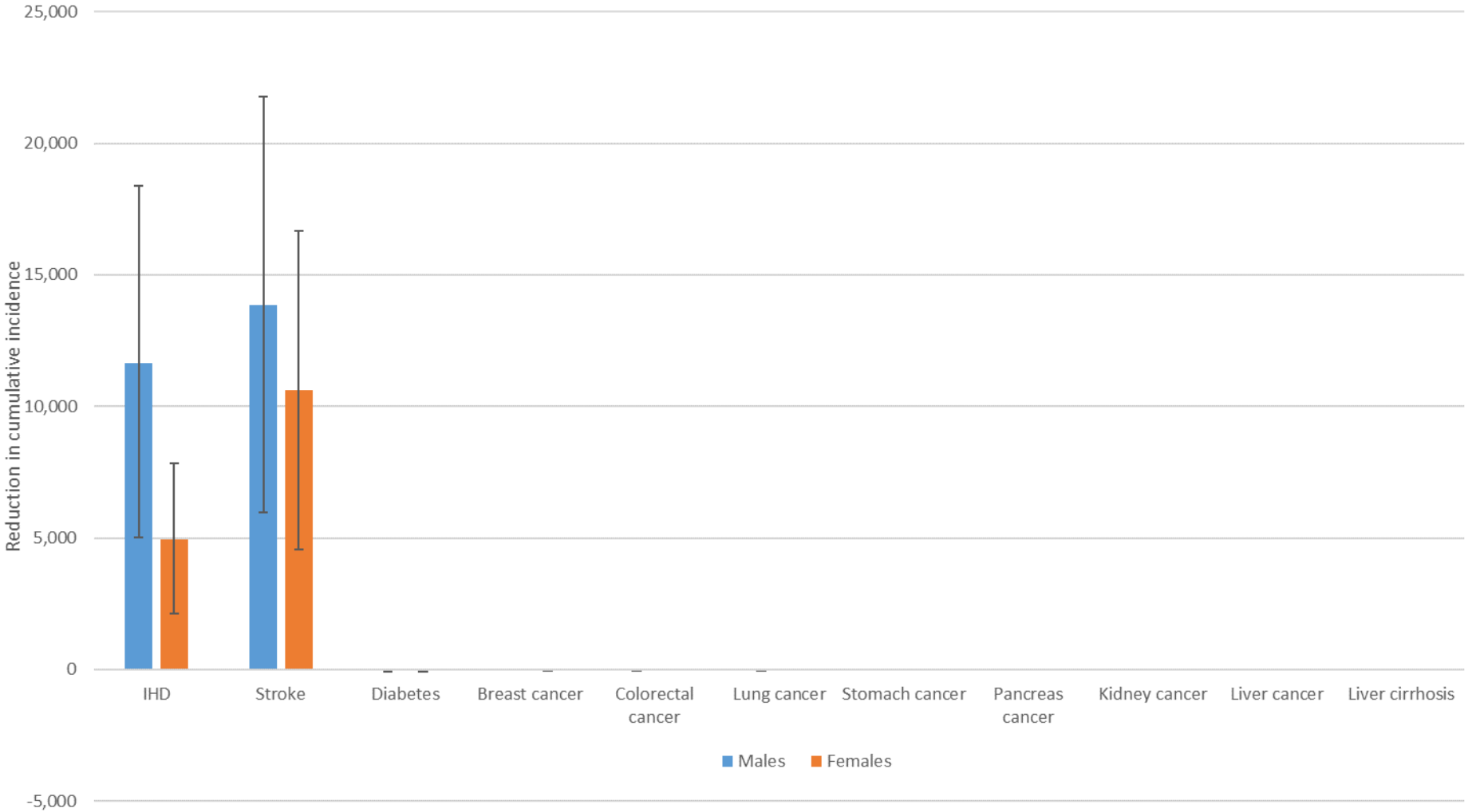
QALY, quality adjusted life year

Figure 5.2 Probability the diet intervention primary analysis is cost-effective



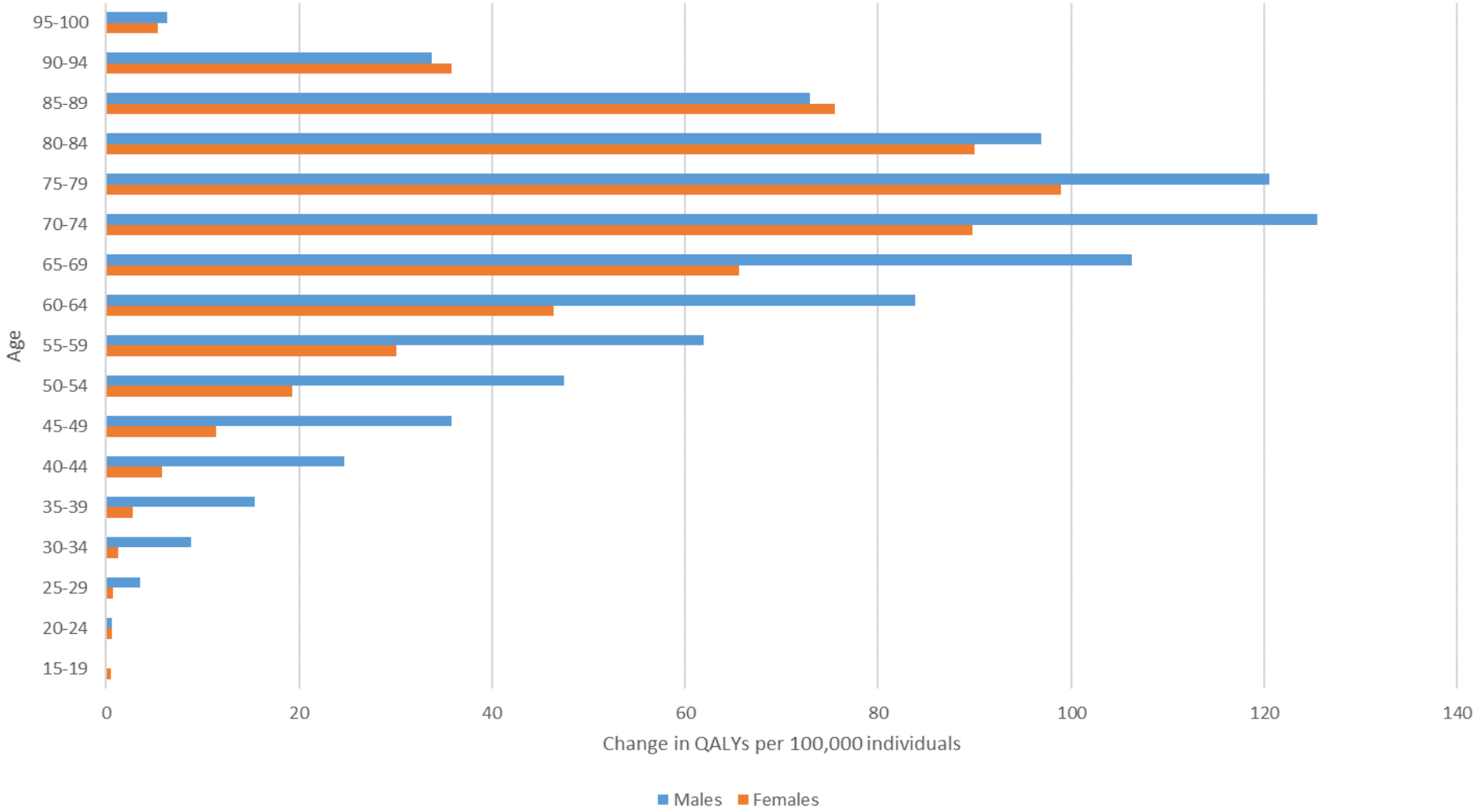
QALY, quality adjusted life year

Figure 5.3 Reduction in cumulative incidence by disease and by sex following the diet intervention primary analysis



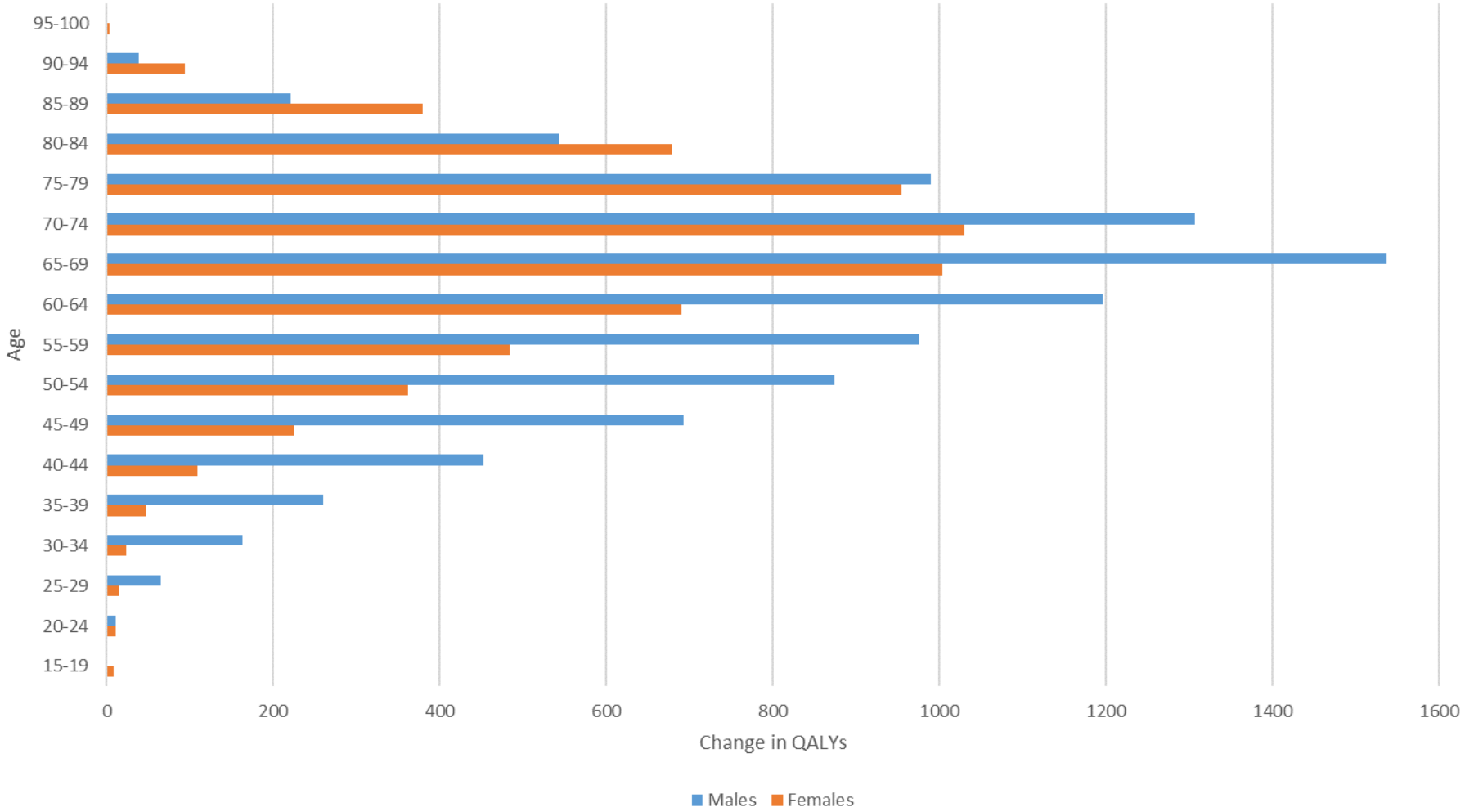
IHD, ischaemic heart disease

Figure 5.4a Change in QALYs per 100,000 individuals by age and sex for the diet intervention primary analysis



QALY, quality adjusted life year

Figure 5.4b Aggregate change in QALYs for each age and sex group for the diet intervention primary analysis



QALY, quality adjusted life year

Figure 5.5a Change in health and social care costs per person by age and sex for the diet intervention primary analysis

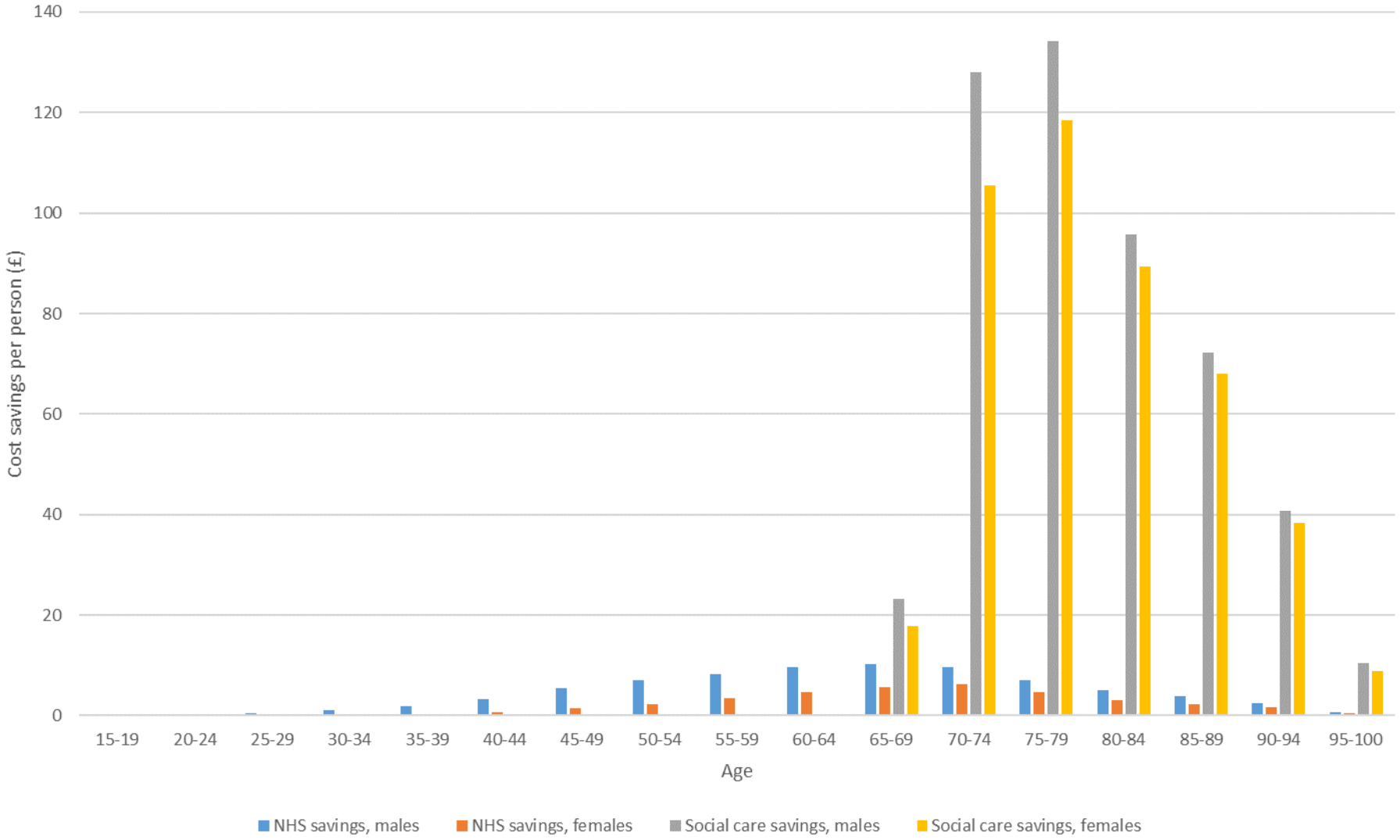
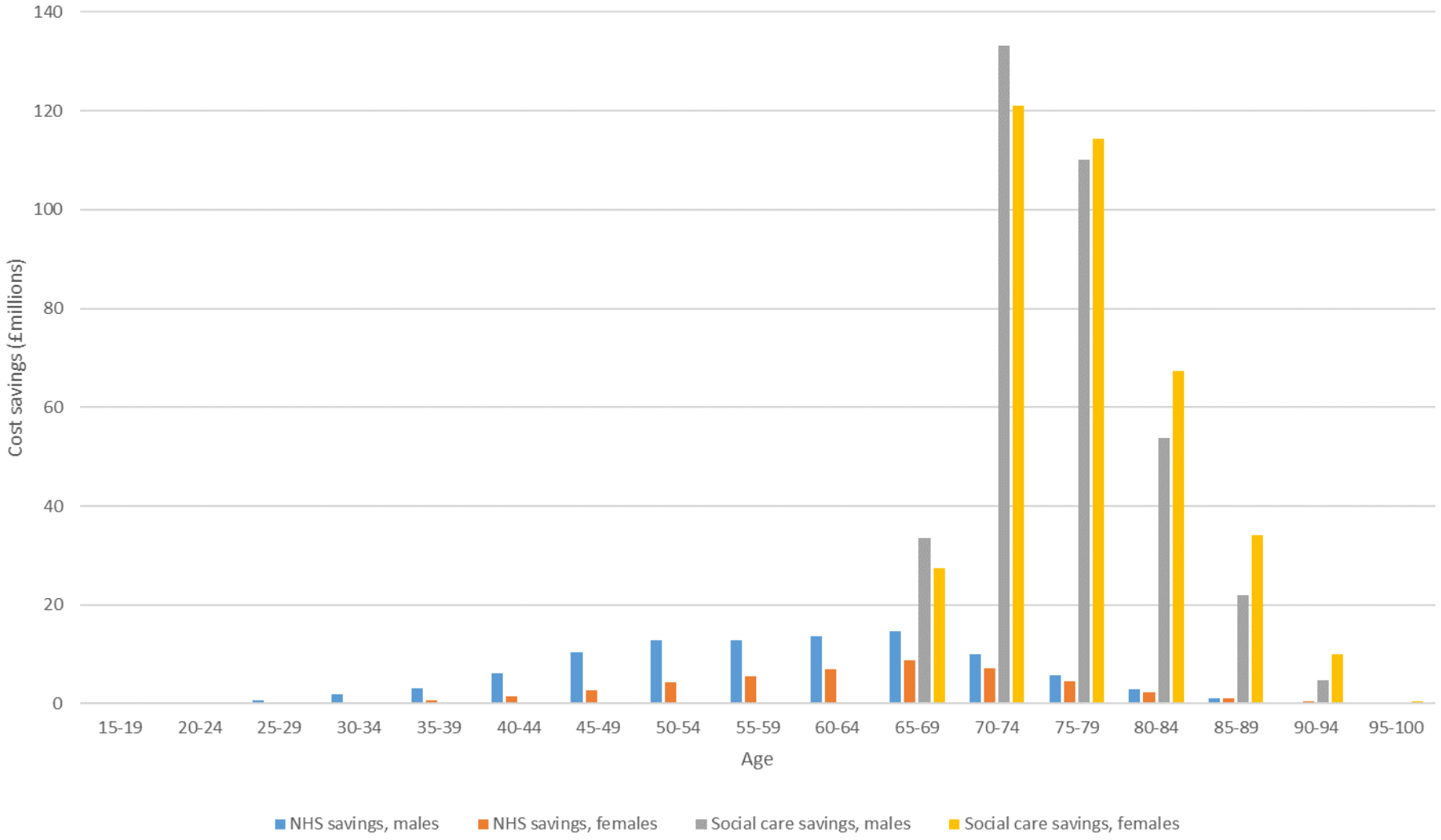


Figure 5.5b Change in aggregate health and social care costs for each age and sex group for the diet intervention primary analysis



5.1.2 Uncertainty and sensitivity analyses

Parametric uncertainty was estimated using 2,000 iterations of a Monte Carlo simulation. This allows all model input parameters to simultaneously vary according to their probability distribution (see appendix four for a full list of PRIMEtime CE input parameters and how their uncertainty is quantified, and section 4.1.1 for how uncertainty in the intervention is parameterised). The uncertainty interval for the cost per QALY in the primary analysis was -£39,553 to £38,547. The contribution of different parameters to this uncertainty is illustrated in the tornado plot in figure 5.6. The tornado plot shows the size of the uncertainty interval when the model only includes uncertainty resulting from either a single model input (such as the effect size of how a change in blood pressure alters stroke risk) or a collection of related model inputs (such as utility decrements). Model inputs leading to the greatest uncertainty are shown at the top of the plot. The largest contributor to the overall parametric uncertainty was the relationship between how a change in salt consumption alters blood pressure. Other input parameters also affected the overall uncertainty interval but to a lesser extent.

Multiple deterministic sensitivity analyses were undertaken testing various assumptions in the model, a complete list of the sensitivity analyses and their rationale is in appendix five. The results of the sensitivity analyses are shown in table 5.2 and illustrated in figures 5.7a (showing all results) and 5.7b (with the x-axis scale amended so as to better visualise results that are similar to the primary analysis). Changing the time horizon had a large influence on results: the cost per QALY was £2.10m (£0.97m to £6.78m) after one year, the intervention was dominant after nine years, and the return on investment was £16.50 (£5.70 to £33.30) for every £1 spent after 100 years (table 5.2, figures 5.7 and 5.8). Removing costs to industry also resulted in a significant increase in return on investment (£110.30 [£37.90 to £223.50] for every £1 spent) and reduced the cost per QALY. Removing government administrative costs made little difference to the cost per QALY but did increase the return on investment to £356.00 (£125.40 to 743.50) as annual intervention costs dropped from £826,125 to £255,233.

As a result of savings from social care being greater than NHS savings, the cost per QALY from an NHS perspective was £30,500 (£13,900 to £90,800) compared with -£7,800 (-£29,800 to £48,300) from a social care perspective. However, in both cases the uncertainty intervals overlap with the result of the primary analysis. Other sensitivity analyses had little effect on the overall cost per QALY compared with the primary analysis.

Table 5.2 Results of sensitivity analyses for the diet intervention

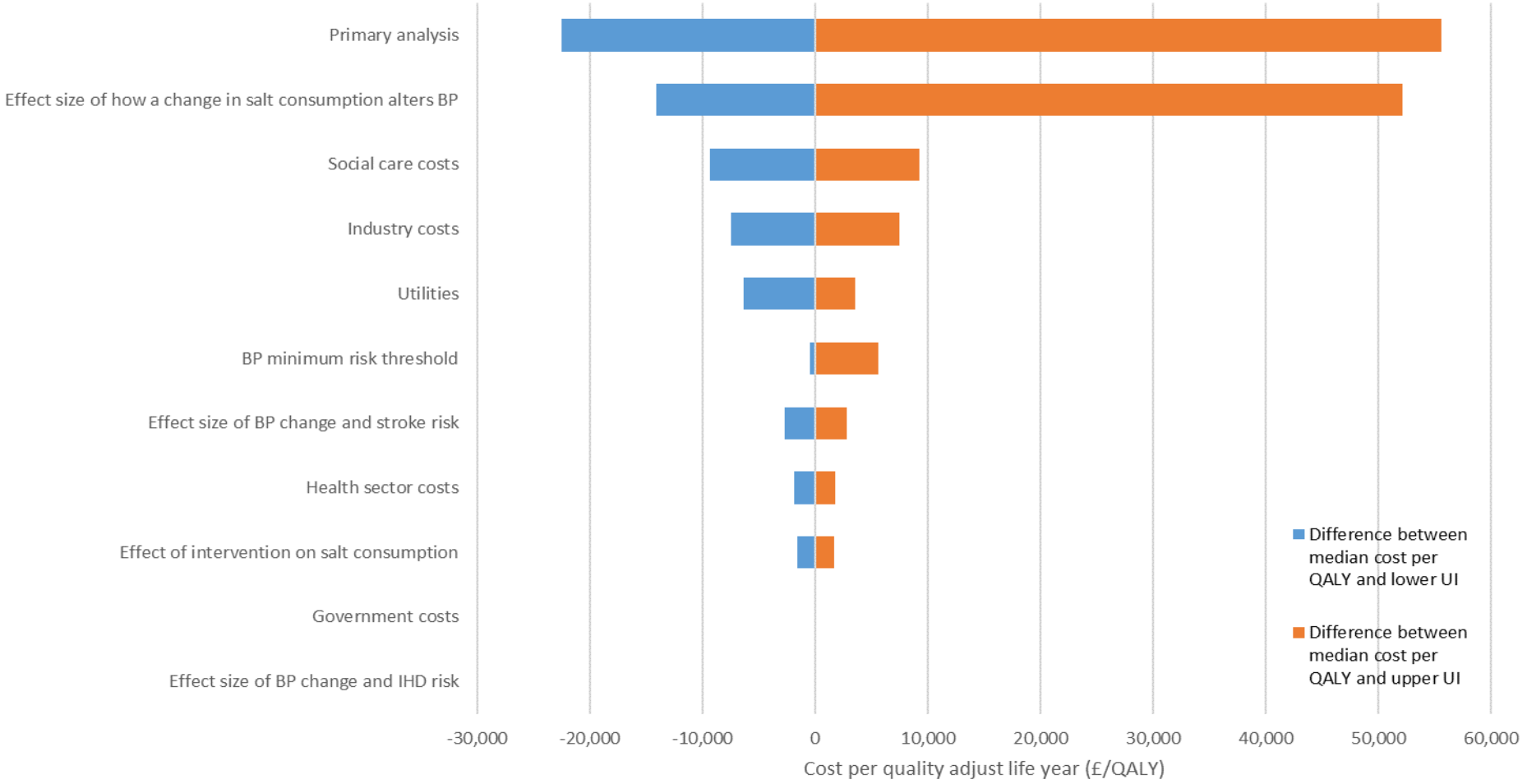
Sensitivity analysis	**Cost per QALY** (£)	**Return on investment** (£ for £1 spent)*
Primary analysis (for reference)	-17,046 (-39,553 to 38,547)	1.44 (0.50 to 2.94)
One year time horizon	2,101,398 (2,516,595 to 17,345,796)	-
Five year time horizon	124,628 (44,612 to 388,462)	-
20 year time horizon	-32,139 (-45,011 to -17,781)	4.63 (1.56 to 9.52)
100 year time horizon (lifetime of cohort)	-28,079 (-35,968 to -21,457)	16.52 (5.66 to 33.31)
No delay in the introduction of the intervention	-22,586 (-42,352 to 24,303)	1.73 (3.54 to 0.56)
No unrelated disease costs included	-20,110 (-41,807 to 34,848)	1.52 (0.52 to 3.05)
Results from an NHS perspective (no social care costs)	30,547 (13,898 to 90,823)	-
Results from a social care perspective (no NHS costs)	-7,824 (-29,778 to 48,329)	1.21 (0.42 to 2.46)
No industry costs included	-56,363 (-79,701 to -42,475)	110.32 (37.90 to 223.46)

Not including industry costs or government administrative costs	-56,646 (-79,160 to -42,895)	355.96 (125.38 to 743.49)
Including social care costs and productivity	-22,806 (-46,075 to 34,059)	1.57 (0.55 to 3.20)
Including all wider societal costs	-7,124 (-26,486 to 45,027)	1.18 (0.41 to 2.41)
Using a discount rate of 3.5%	-12,036 (-36,413 to 53,543)	1.27 (0.44 to 2.67)
No cancer included in model	-20,110 (-41,807 to 34,848)	1.52 (0.52 to 3.05)
Only including diseases directly related to blood pressure	-16,233 (-38,584 to 38,829)	1.41 (0.50 to 2.92)
Specialised services costs based on HES data	-17,550 (-37,822 to 40,375)	1.43 (0.51 to 2.95)
Primary care costs based on HES data	-16,921 (-38,869 to 33,141)	1.43 (0.51 to 2.87)
Maternity costs estimated using the general and acute cost curve	-16,177 (-39,087 to 55,978)	1.41 (0.42 to 2.89)

*Return on investment reported for analyses where the median cost per QALY is negative

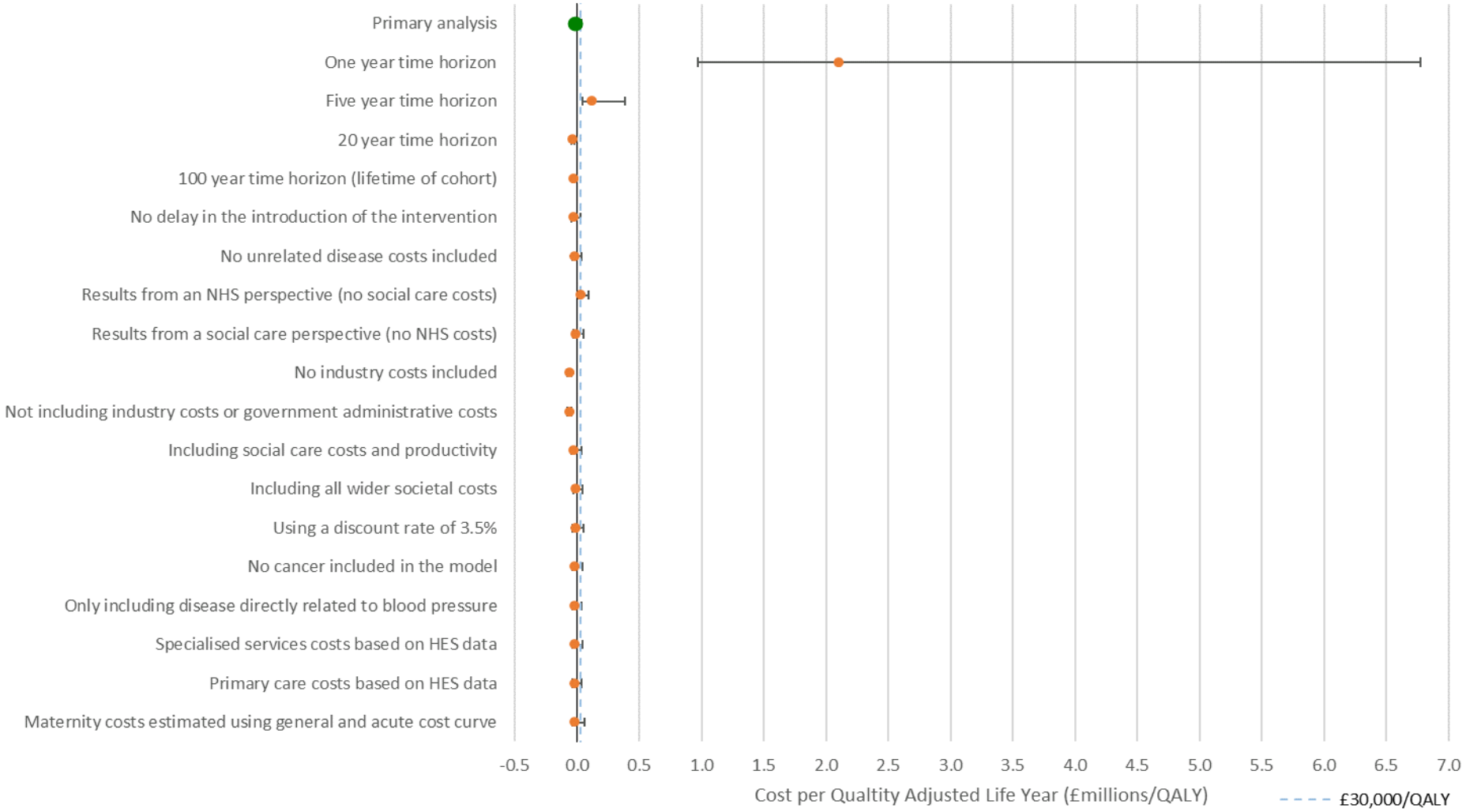
Median results reported from 2,000 iterations of a Monte Carlo simulation, with 95% uncertainty intervals in parentheses; QALY, quality adjusted life year; HES, hospital episode statistics

Figure 5.6 Tornado plot of uncertainty of model input parameters for diet intervention



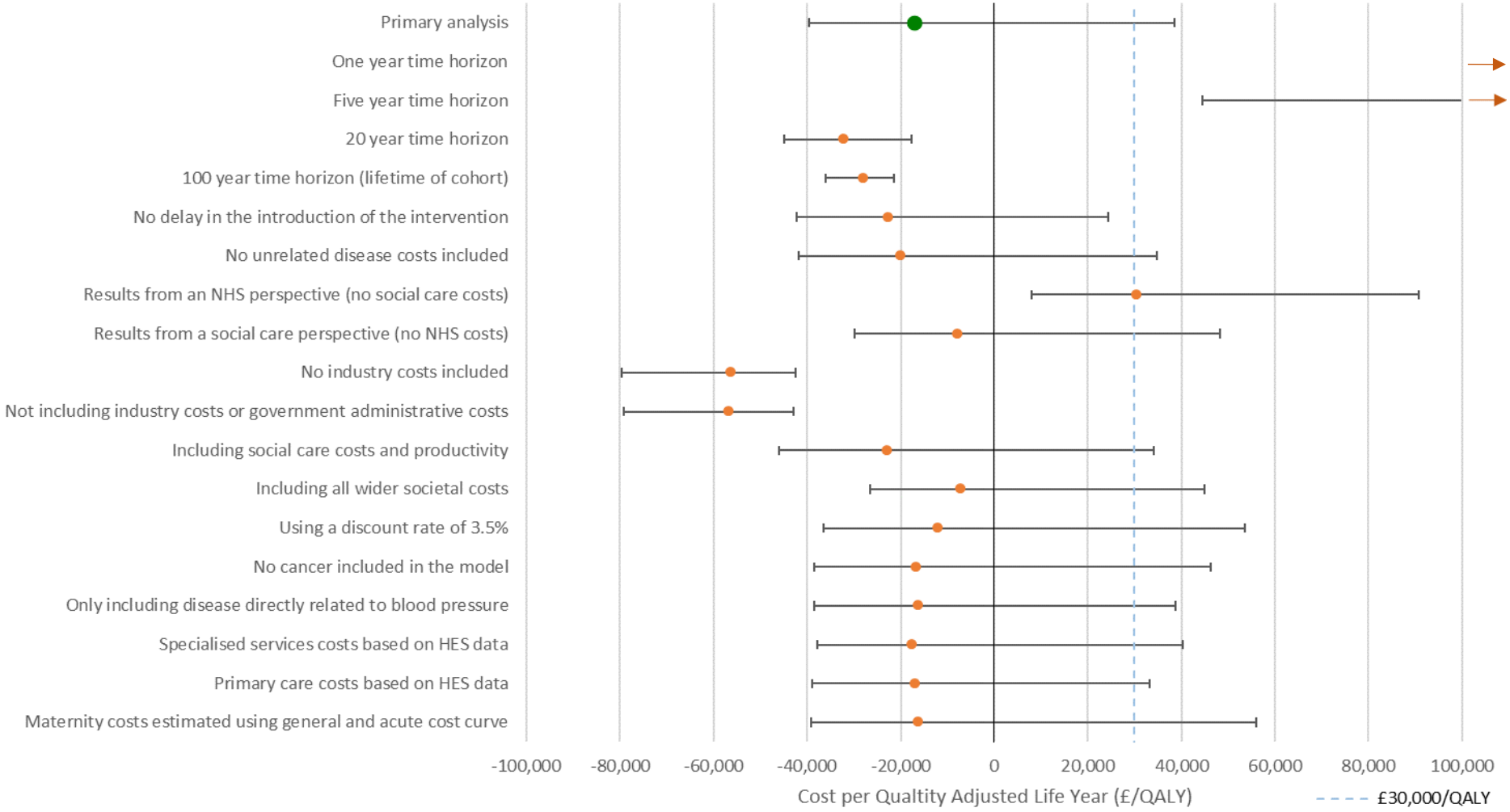
The bars illustrate the width of the uncertainty interval when only one parameter or collection of related parameters is allowed to vary according to their probability distribution. QALY, quality adjusted life year; UI, uncertainty interval; BP, blood pressure; IHD, ischaemic heart disease

Figure 5.7a Results of deterministic sensitivity analyses for the diet intervention



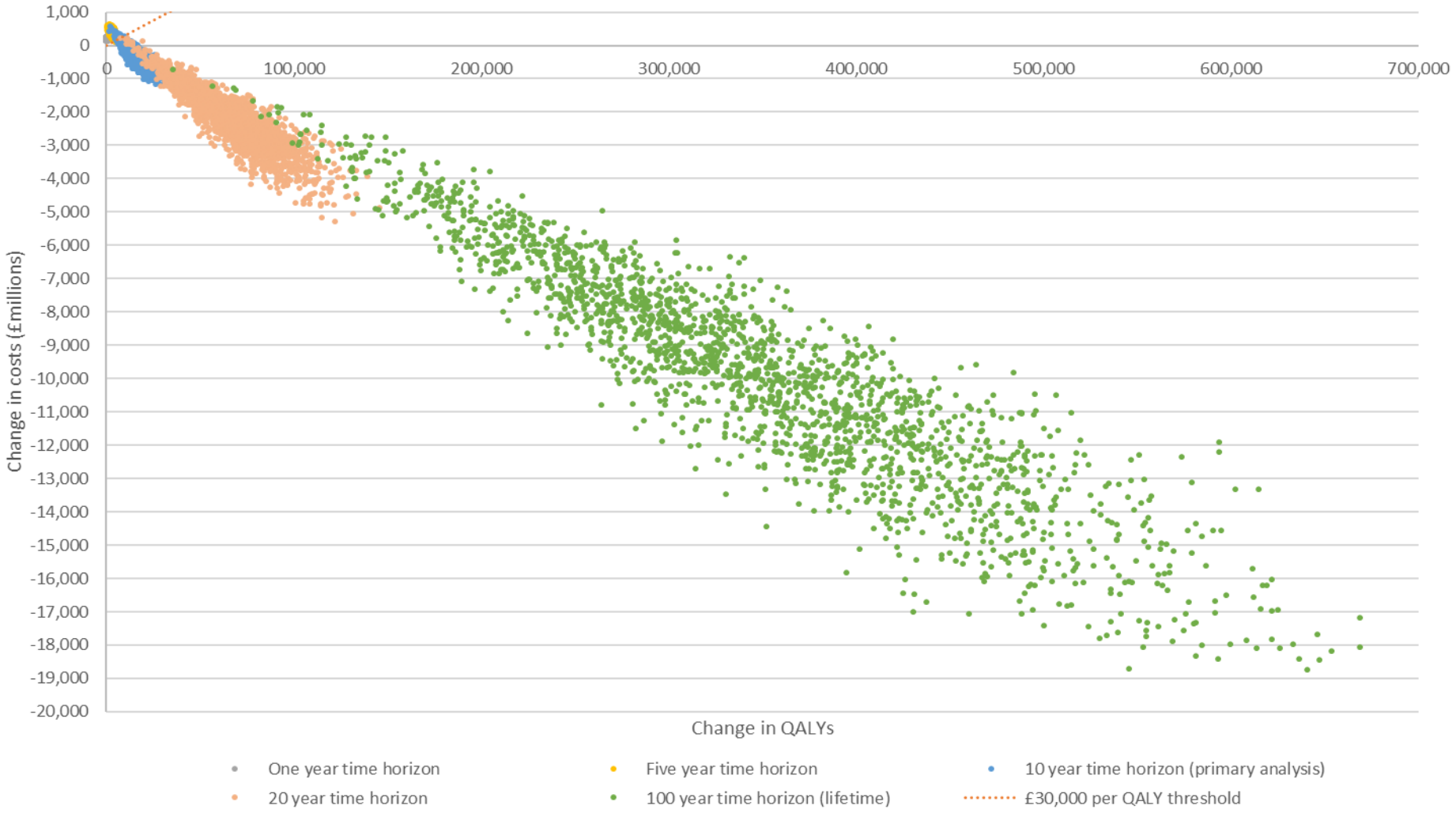
QALY, quality adjusted life year; HES, hospital episode statistics

Figure 5.7b Results of deterministic sensitivity analyses for the diet intervention with smaller scale on the x-axis



QALY, quality adjusted life year; HES, hospital episode statistics

Figure 5.8 Cost effectiveness plane showing results of the diet intervention with different time horizons



QALY, quality adjusted life year

5.2 Physical activity intervention

Expanding the Birmingham Be Active scheme to the whole of England was estimated to result in absolute percentage increases of 1.3% among men and 1.4% among women in the number of people achieving recommended levels of physical activity (see section 4.2.2).

5.2.1 Results of the primary analysis

The Be Active expansion resulted in a median increase of 1,600 (1,130 to 2,170) QALYs over 10 years with savings to the NHS of £10.40m (£7.19m to £14.58m) and social care of £14.18m (£7.84m to £22.41m). The intervention cost £1,177m (£882m to £1,602m) meaning the median cost per QALY was £727,000 (£514,000 to £1,064,000; mean, £743,000); there was no probability of the intervention being below £30,000 per QALY. To achieve a cost per QALY of £30,000, 10 year intervention costs would need to be £72.44m, approximately £2.20 per participant per year compared to the modelled costs of £53.80 per participant in year one and £29.80 in year two onwards. Results of the primary analysis are in table 5.3, and shown in figures 5.9 and 5.10. The disease with the largest reduction in cumulative incidence was diabetes with 4,200 (2,970 to 5,740) fewer cases over the 10 years of the intervention, followed by stroke (550 [300 to 870] fewer cases), IHD (210 [150 to 280]), colorectal cancer (130 [-60 to 400]), and breast cancer (130 [30 to 250]). Men experienced larger health benefits than women for all conditions except breast cancer (figure 5.11).

Results varied by age and sex. The largest average gain in QALYs was among men aged 40-44 years (5.6 additional QALYs per 100,000 people), and the greatest cost savings were among men aged 50-54 years for NHS costs (£0.41 per person) and men aged 75-79 years for social care costs (£2.72 per person). The overall cost per QALY ranged from £5.60m for 15-19 year olds to £106,000 for 85-89 year olds. Figures 5.12 and 5.13 show the change in QALYs and costs per person and in total by age and sex. Figure 5.14 shows how cost per QALY varies with age. The intervention does not become cost-saving over the life-time of the cohort meaning there is no time to return on investment.

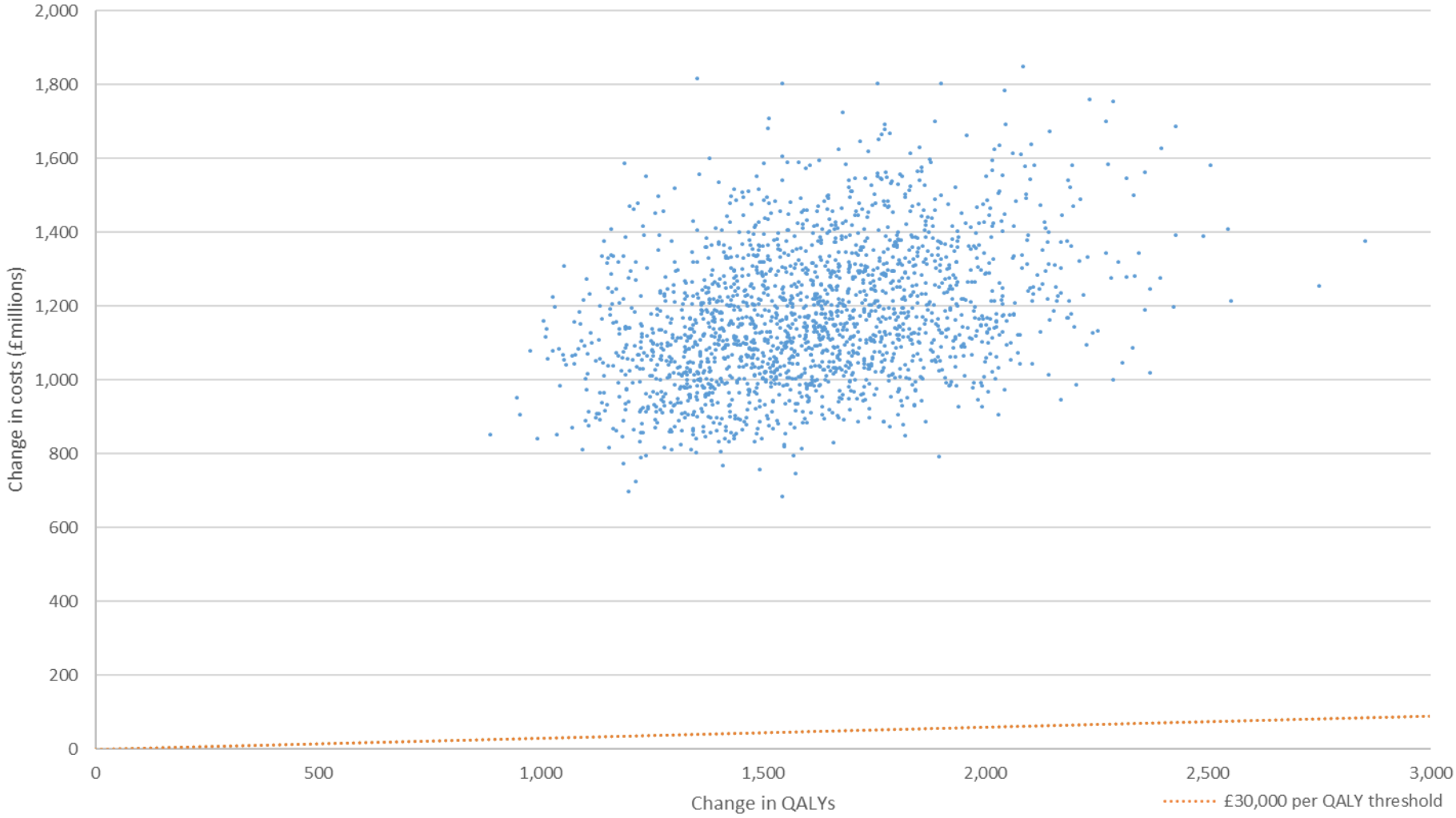
Table 5.3 Summary of results following the physical activity intervention primary analysis

	Males	**Females**	**Total**
Cost per QALY - overall	652,388	818,121	727,300
	(458,290 to 951,568)	(580,315 to 1,186,274)	(513,736 to 1,063,585)
Change in QALYs	872	723	1,597
	(614 to 1,197)	(518 to 984)	(1,134 to 2,169)
NHS savings (£)	5,772,494	4,627,015	10,402,110
	(3,975,940 to 8,107,746)	(3,206,297 to 6,477,924)	(7,190,945 to 14,583,876)
Social care savings (£)	6,997,737	7,186,580	14,181,040
	(3,810,312 to 11,073,841)	(3,995,764 to 11,333,750)	(7,838,893 to 22,409,996)
Intervention costs (£)	576,601,063	600,728,881	1,177,329,943
	(431,869,562 to 784,540,463)	(449,926,458 to 817,269,990)	(881,796,020 to 1,601,810,453)
Cost per QALY - NHS perspective (£)	660,208	829,675	735,851
	(466,633 to 959,372)	(589,731 to 1,196,543)	(521,112 to 1,071,650)
Cost per QALY - social care perspective	658,986	824,804	733,402
(£)	(464,292 to 958,713)	(586,489 to 1,193,704)	(519,764 to 1,069,979)
Change in IHD cumulative incidence	-144 [-0.6]	-61 [-0.3]	-205 [-0.5]
[incidence rate per 100,000]	(-194 to -102)	(-81 to -44)	(-275 to -147)
Change in stroke cumulative incidence	-305 [-1.4]	-241 [-1.1]	-545 [-1.2]
[incidence rate per 100,000]	(-488 to -163)	(-380 to -132)	(-867 to -295)
Change in type two diabetes cumulative	-2,264 [-10.5]	-1,932 [-8.6]	-4,199 [-9.5]
incidence [incidence rate per 100,000]	(-3,104 to -1,599)	(-2,631 to -1,369)	(-5,738 to -2,973)

Change in breast cancer cumulative incidence [incidence rate per 100,000]	0 [0]	-125 [-0.6] (-245 to -30)	-125 [-0.3] (-245 to -30)
Change in colorectal cancer cumulative incidence [incidence rate per 100,000]	-82 [-0.4] (-253 to 32)	-47 [-0.2] (-143 to 22)	-130 [-0.3] (-397 to 60)
Change in lung cancer cumulative incidence [incidence rate per 100,000]	0 [0] (0 to 1)	0 [0] (0 to 0)	1 [0] (0 to 1)
Change in stomach cancer cumulative incidence [incidence rate per 100,000]	0 [0] (0 to 0)	0 [0] (0 to 0)	0 [0] (0 to 0)
Change in pancreas cancer cumulative incidence [incidence rate per 100,000]	0 [0] (0 to 0)	0 [0] (0 to 0)	0 [0] (0 to 0))
Change in kidney cancer cumulative incidence [incidence rate per 100,000]	0 [0] (0 to 0)	0 [0] (0 to 0)	0 [0] (0 to 0))
Change in liver cancer cumulative incidence [incidence rate per 100,000]	0 [0] (0 to 0)	0 [0] (0 to 0)	0 [0] (0 to 0))
Change in liver cirrhosis cumulative incidence [incidence rate per 100,000]	0 [0] (0 to 0)	0 [0] (0 to 0)	0 [0] (0 to 0))

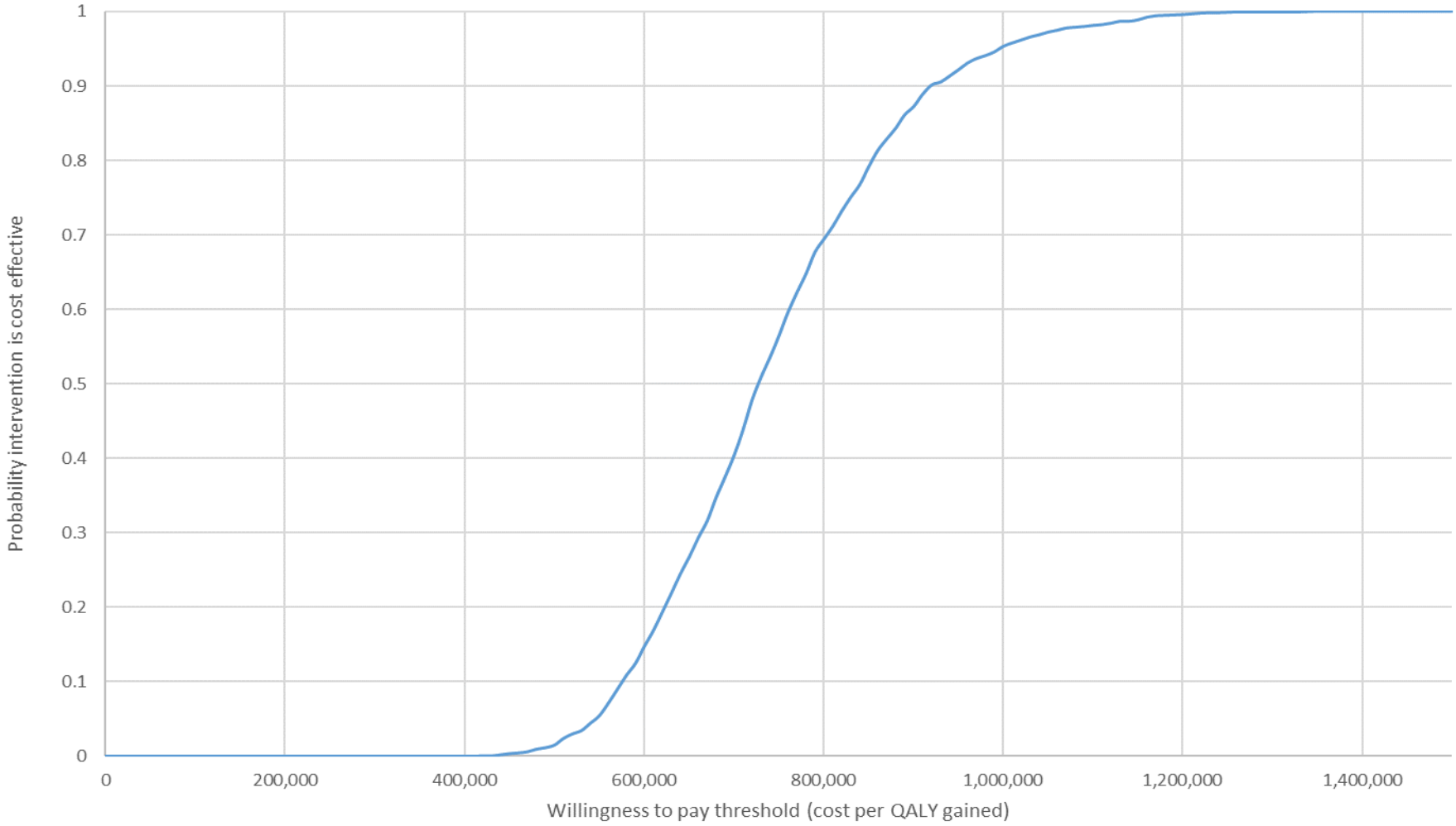
Results presented are the median values from the Monte Carlo simulation, as such, the numbers in the final column may not equal the sum of males and females; 95% uncertainty intervals in parentheses; QALY, quality adjusted life year; IHD, ischaemic heart disease

Figure 5.9 Cost effectiveness plane showing results of the physical activity intervention primary analysis



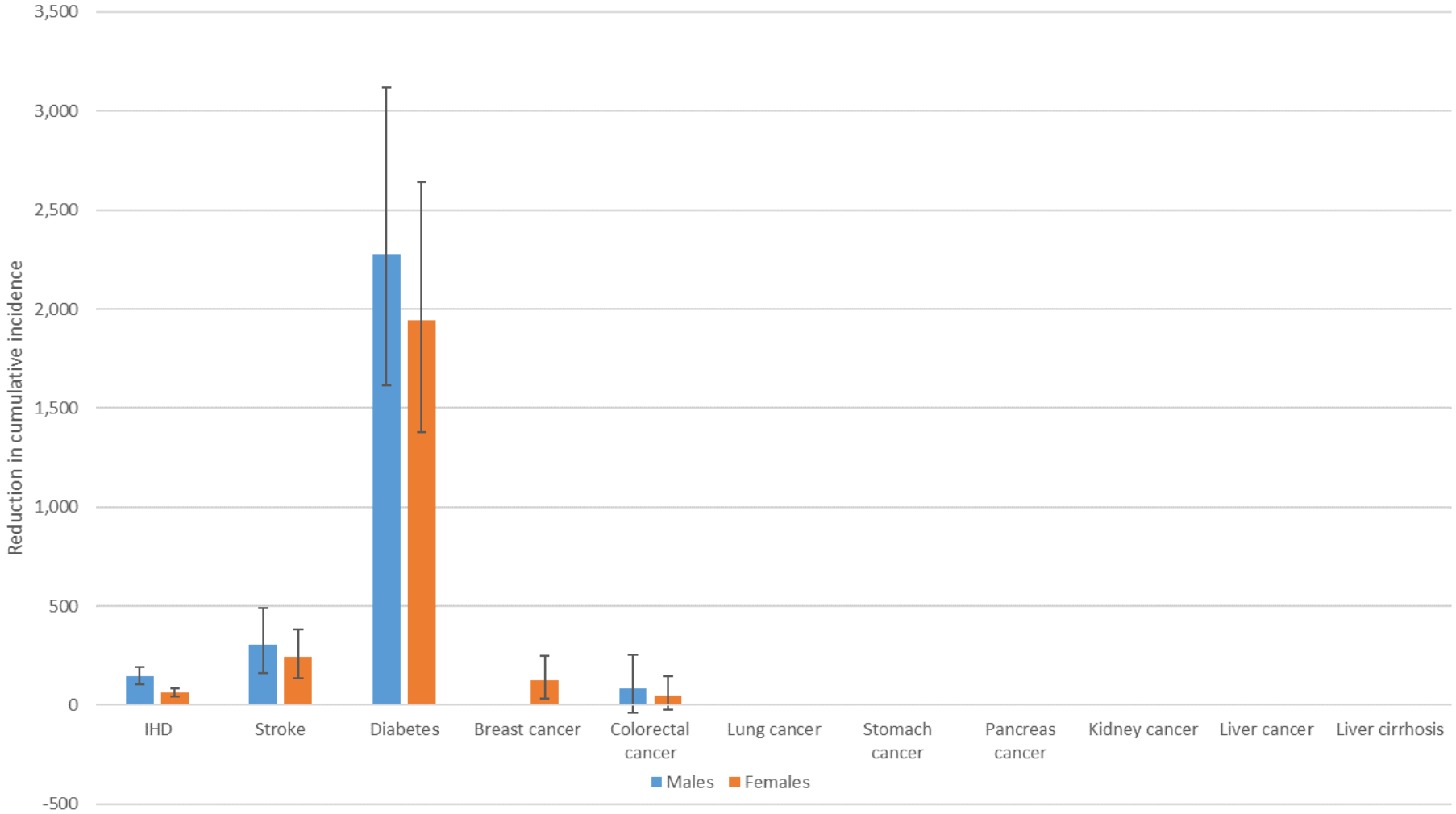
QALY, quality adjusted life year

Figure 5.10 Probability the physical activity primary analysis is cost-effective



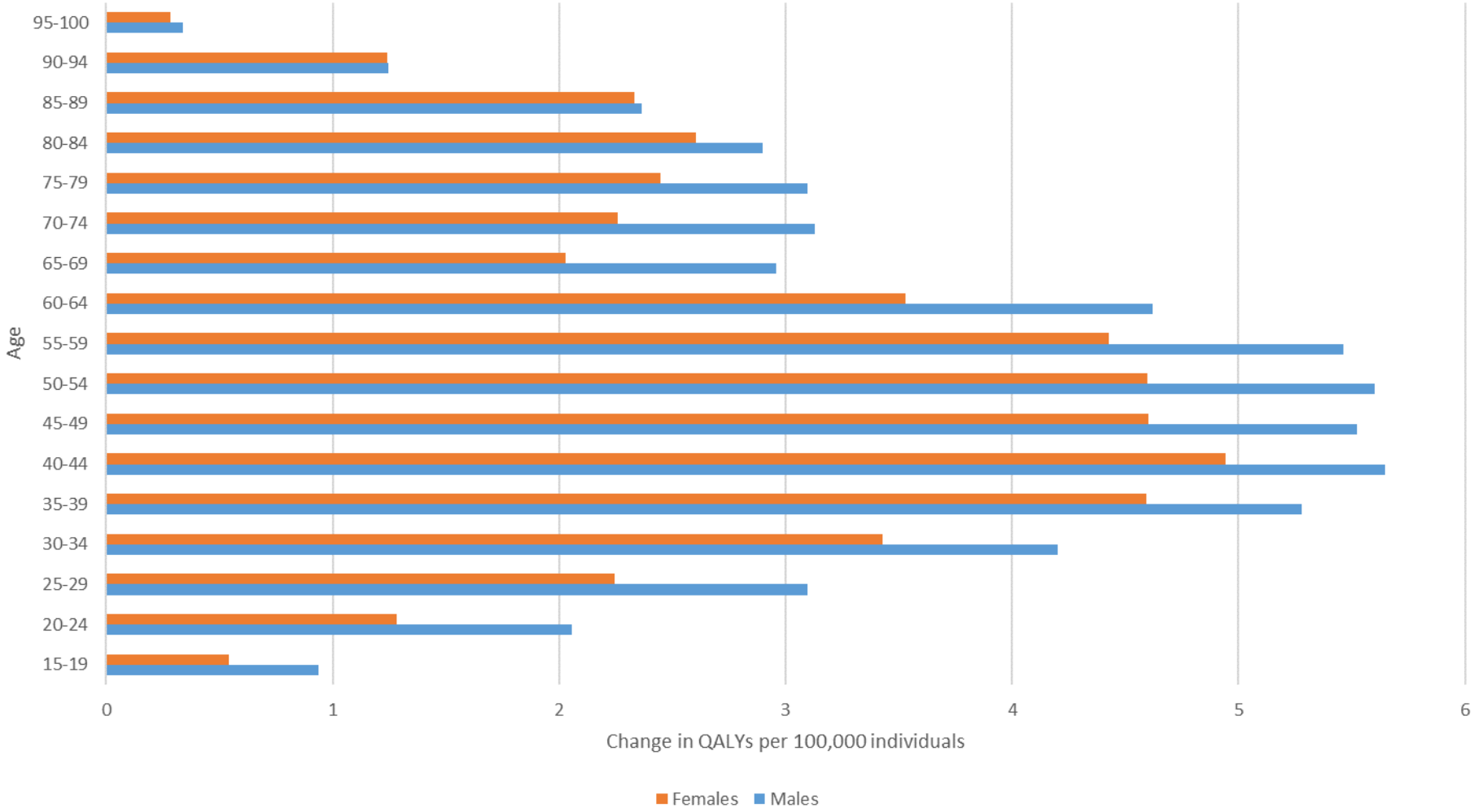
QALY, quality adjusted life year

Figure 5.11 Reduction in cumulative incidence by disease and by sex following the physical activity intervention primary analysis



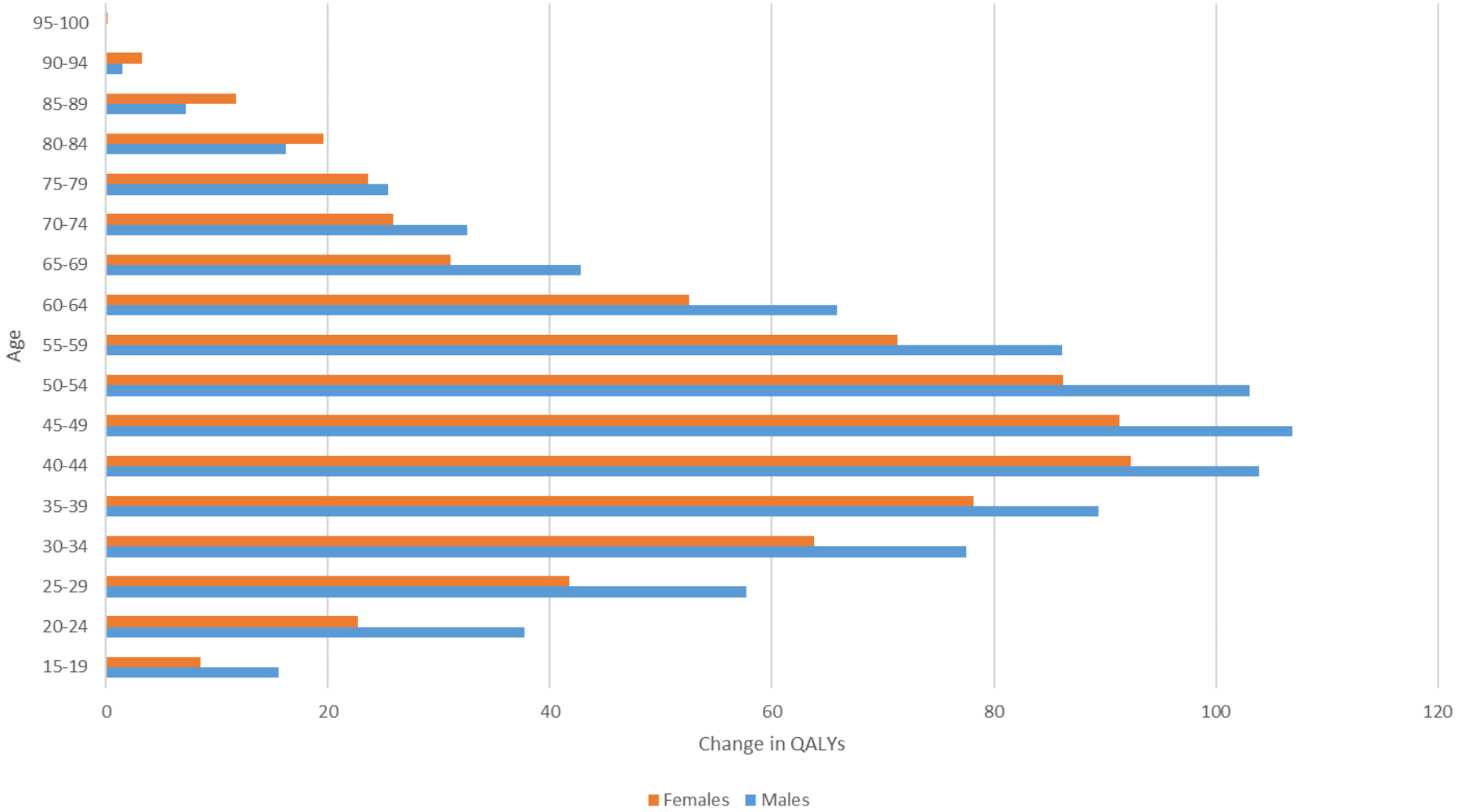
IHD, ischaemic heart disease

Figure 5.12a Change in QALYs per 100,000 individuals by age and sex for the physical activity primary analysis



QALYs, quality adjusted life years

Figure 5.12b Aggregate change in QALYs for each age and sex group for the physical activity primary analysis



QALY, quality adjusted life year

Figure 5.13a Change in health and social care costs per person by age and sex for the physical activity primary analysis

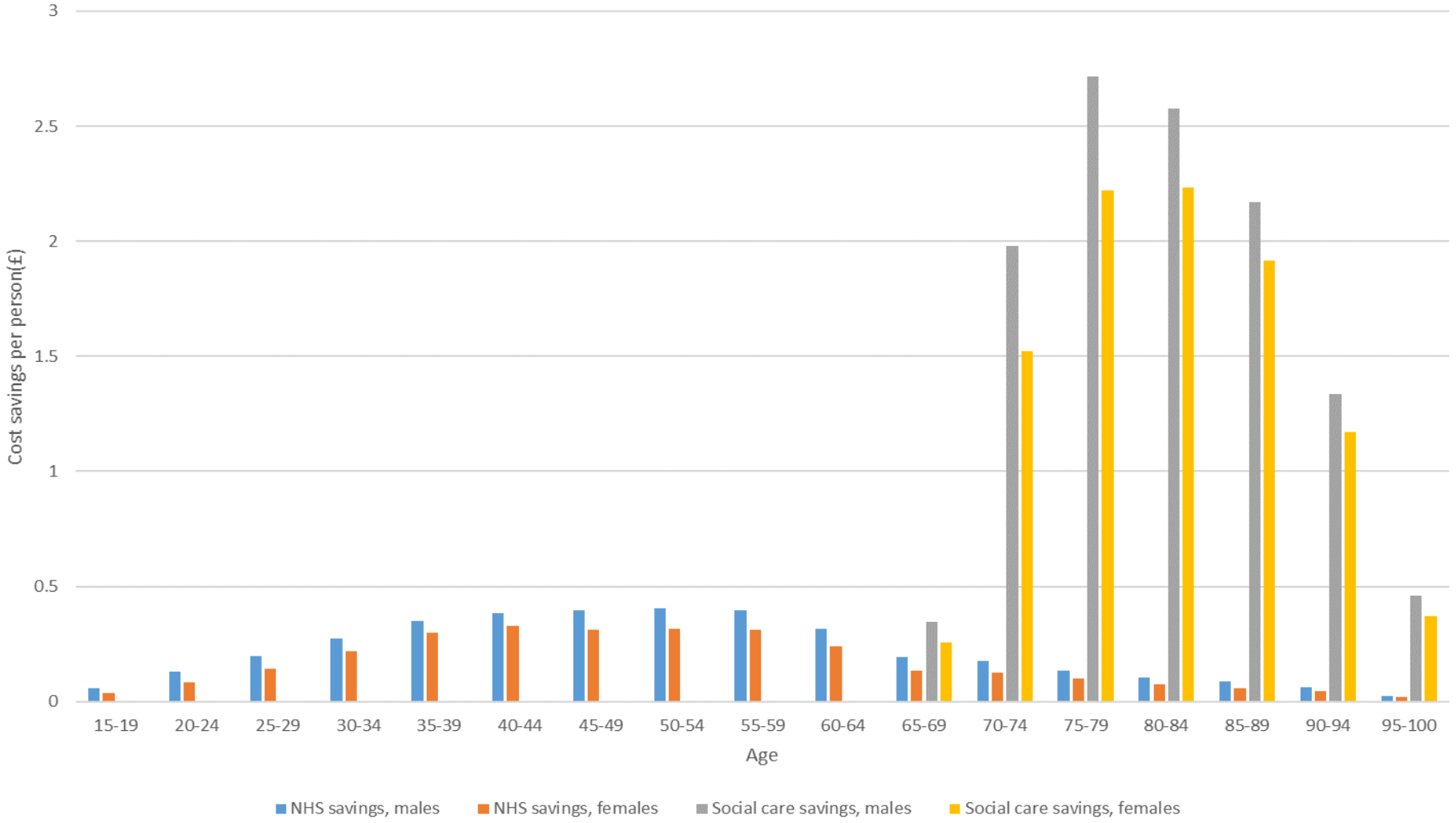


Figure 5.13b Change in aggregate health and social care costs by age and sex for the physical activity primary analysis

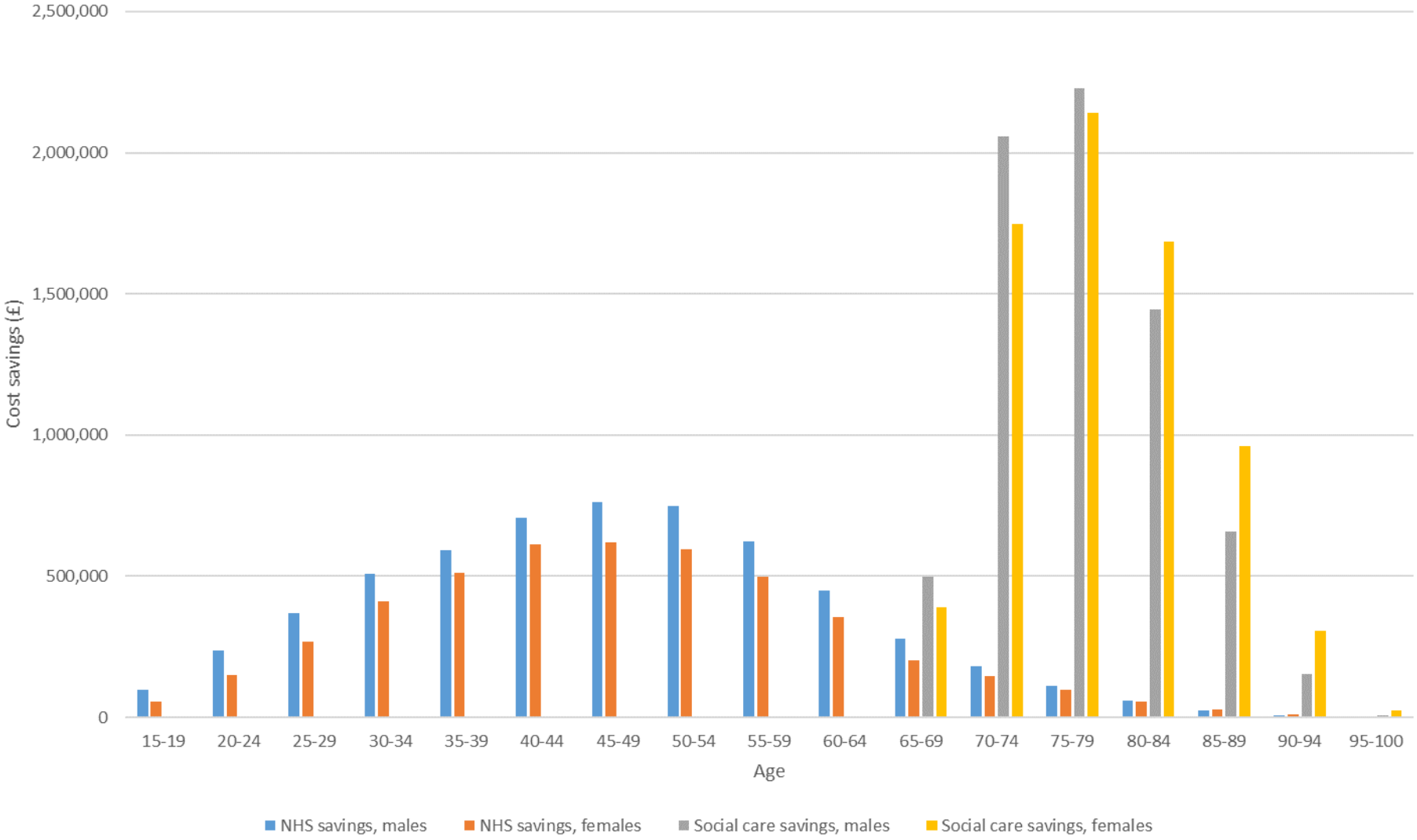
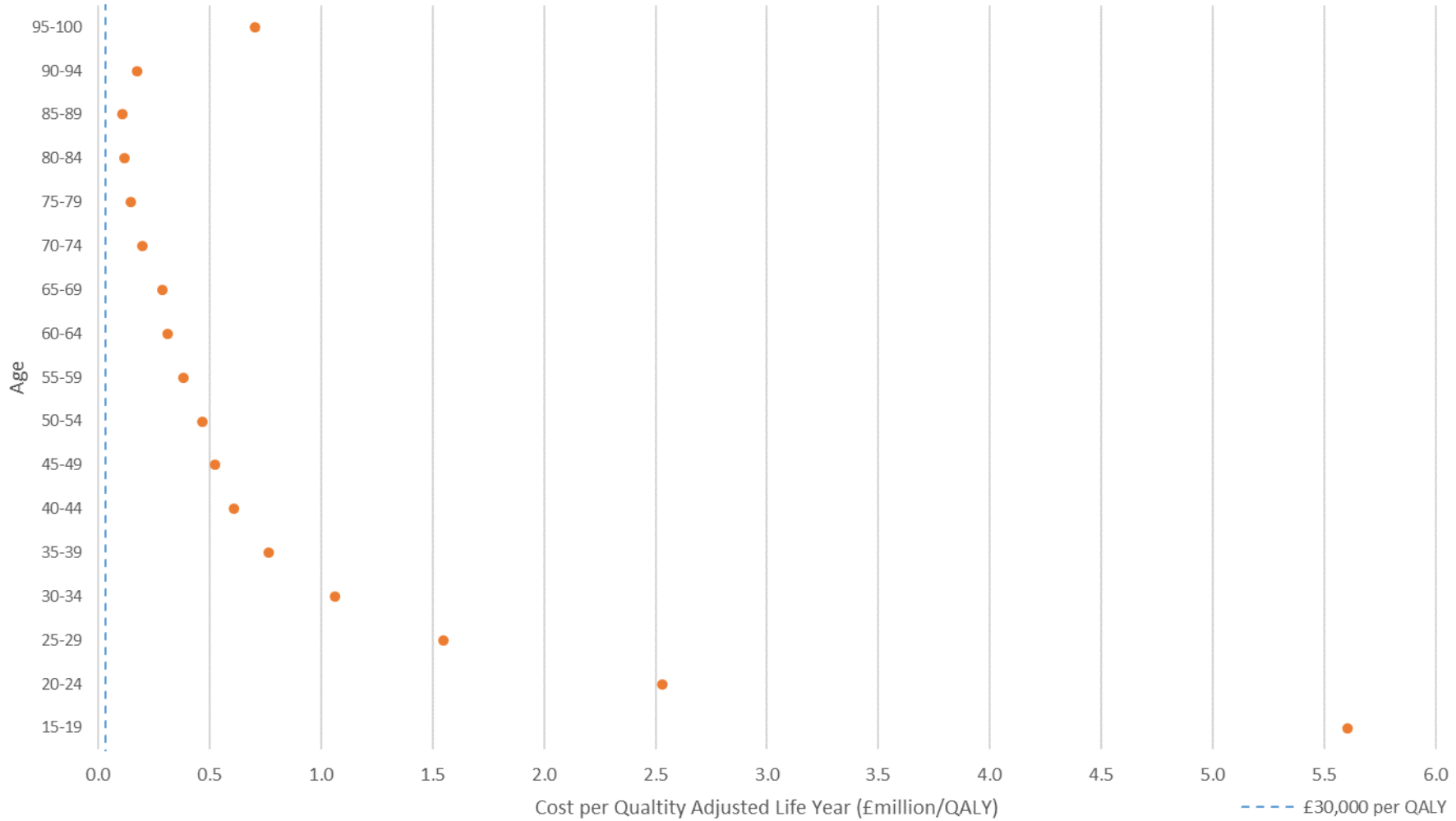


Figure 5.14 Cost per quality adjusted life year by age for the physical activity primary analysis



QALY, quality adjusted life year

5.2.2 Uncertainty and sensitivity analyses

The uncertainty interval for the cost per QALY in the primary analysis was £513,736 to £1,063,585. The tornado plot in figure 5.15 shows how different input parameters contributed to the overall uncertainty interval. The largest contributors were the uncertainty in the effect size of how changing physical activity levels alters the risk of diabetes, stroke, and colorectal cancer, the uncertainty of utility estimates, and uncertainty in the cost of delivering the intervention.

Sensitivity analyses and an explanation of their rationale are listed in appendix five, and results are shown in table 5.4 and figure 5.16. Changing the time-horizon significantly altered the cost-effectiveness of the intervention, with results ranging from £12.18m (£8.20m to £18.47m) per QALY over a one year time horizon, to £91,700 (£63,000 to £137,400) per QALY if estimating outcomes over the lifetime of the cohort (figures 5.16 and 5.17). Other sensitivity analyses did not significantly alter the cost-effectiveness of the intervention, except for when modelling a reduction in physical activity over time which resulted in a cost per QALY of £1.05m (£0.76m to £1.52m). The reduction in physical activity assumed that 50% of those who moved from active to recommended levels of physical activity following the intervention, moved back to active levels after the first year. None of the sensitivity analyses resulted in the median cost per QALY of less than £30,000.

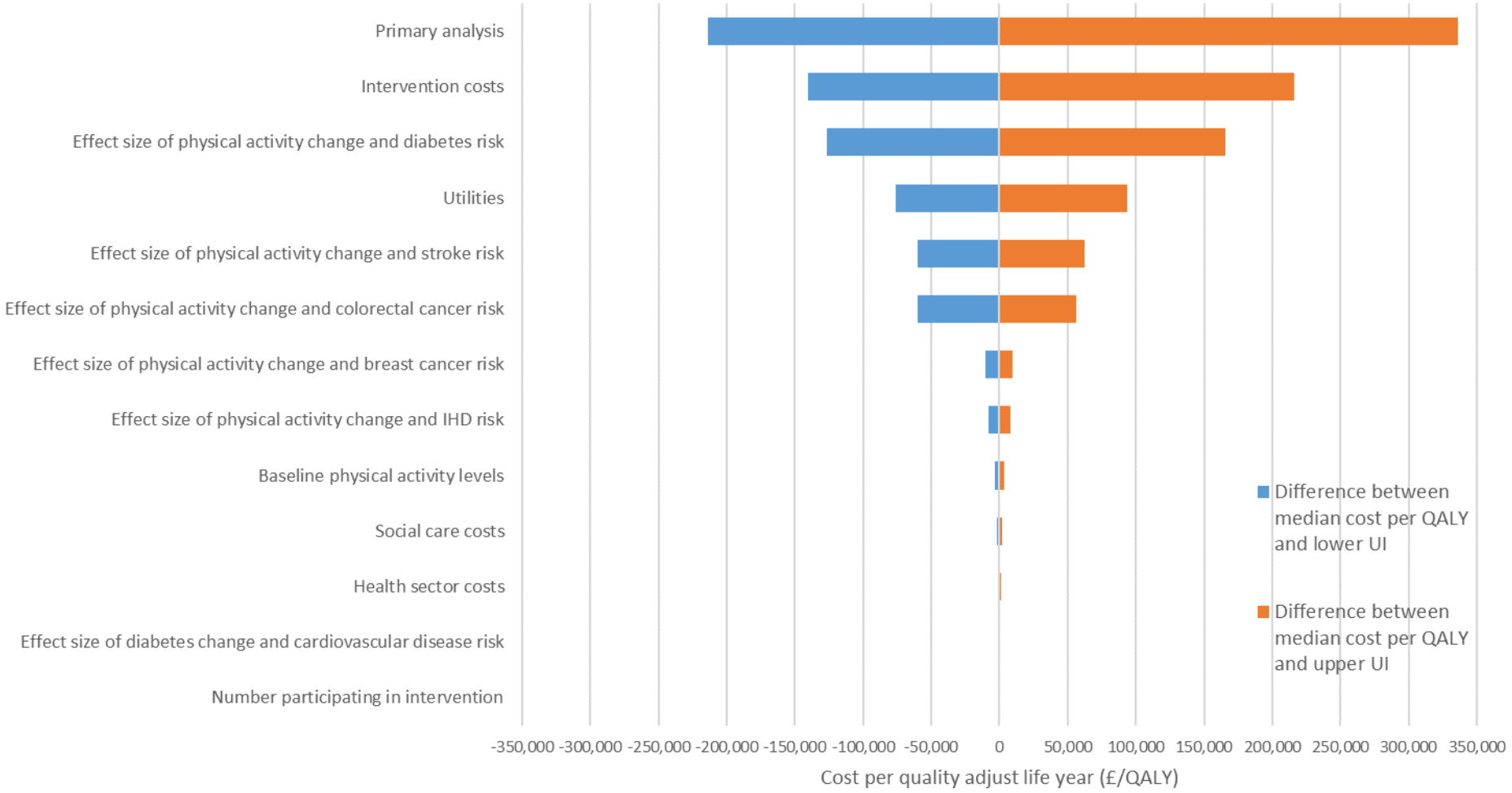
Table 5.4 Results of sensitivity analyses for the physical activity intervention

Sensitivity analysis	**Cost per QALY** (£)
Primary analysis (for reference)	727,300 (513,736 to 1,063,585)
One year time horizon	12,183,276 (8,198,390 to 18,469,397)
Five year time horizon	1,669,123 (1,218,480 to 2,363,574)
20 year time horizon	312,772 (216,893 to 461,198)
100 year time horizon (lifetime of cohort)	91,724 (62,950 to 137,421)

Fall in physical activity over time	1,052,103 (756,141 to 1,515,405)
No unrelated disease costs included	735,474 (521,921 to 1,040,670)
Results from an NHS perspective (no social care costs)	735,851 (521,112 to 1,071,650)
Results from a social care perspective (no NHS costs)	733,402 (519,764 to 1,069,979)
Including social care costs and productivity	720,677 (512,601 to 1,047,837)
Including all wider societal costs	705,043 (502,287 to 1,018,644)
Using a discount rate of 3.5%	765,447 (559,808 to 1,102,006)
No cancer included in model	784,240 (550,810 to 1,111,700)
Only including diseases directly related to physical activity	728,128 (518,729 to 1,065,997)
Specialised services costs based on HES data	723,223 (512,122 to 1,050,280)
Primary care costs based on HES data	726,140 (518,161 to 1,056,634)
Maternity costs estimated using general and acute cost curve	729,106 (516,460 to 1,065,997)

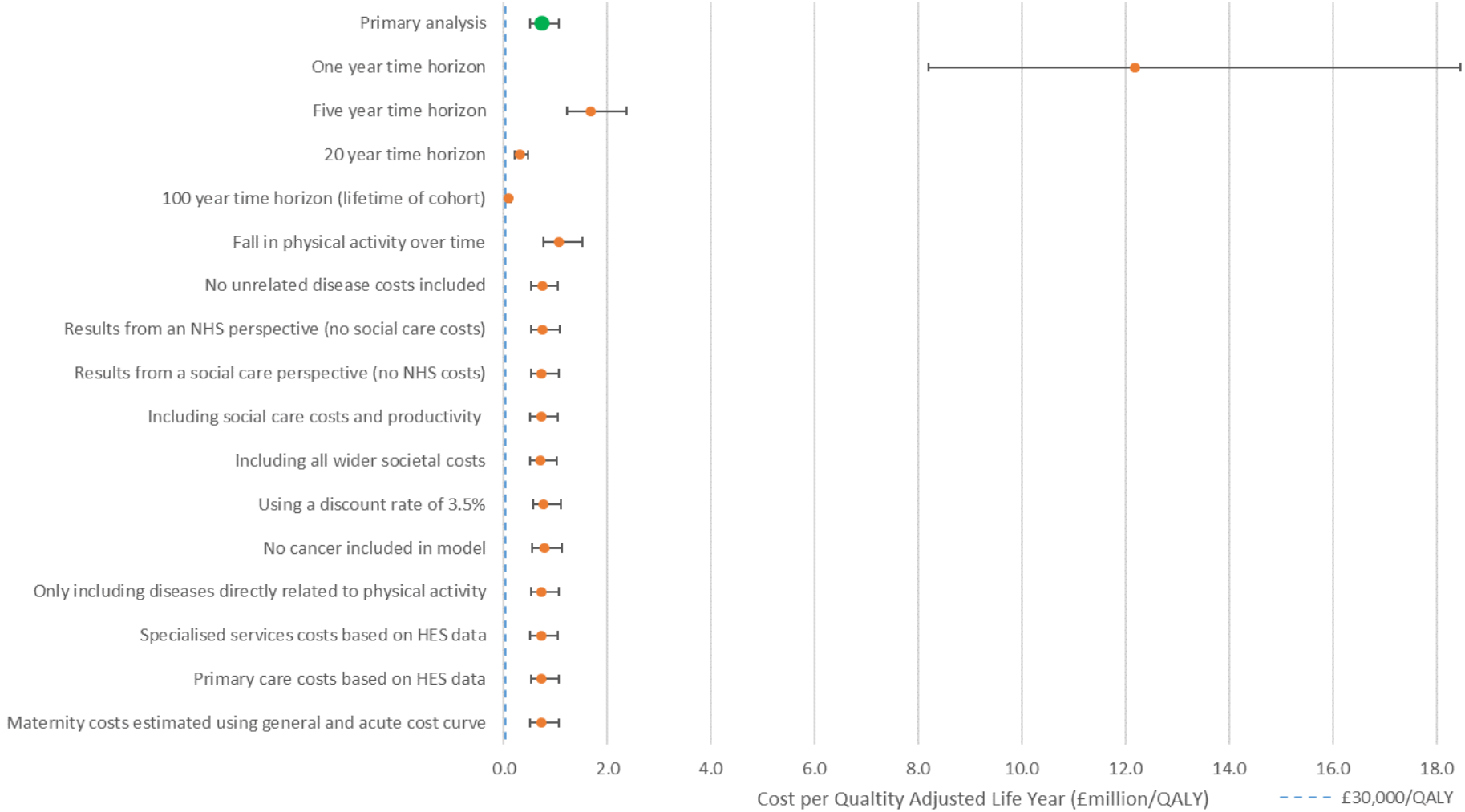
Median results reported from 2,000 iterations of a Monte Carlo simulation, with 95% uncertainty intervals in parentheses; QALY, quality adjusted life year; HES, hospital episode statistics

Figure 5.15 Tornado plot of uncertainty of model input parameters for the physical activity intervention



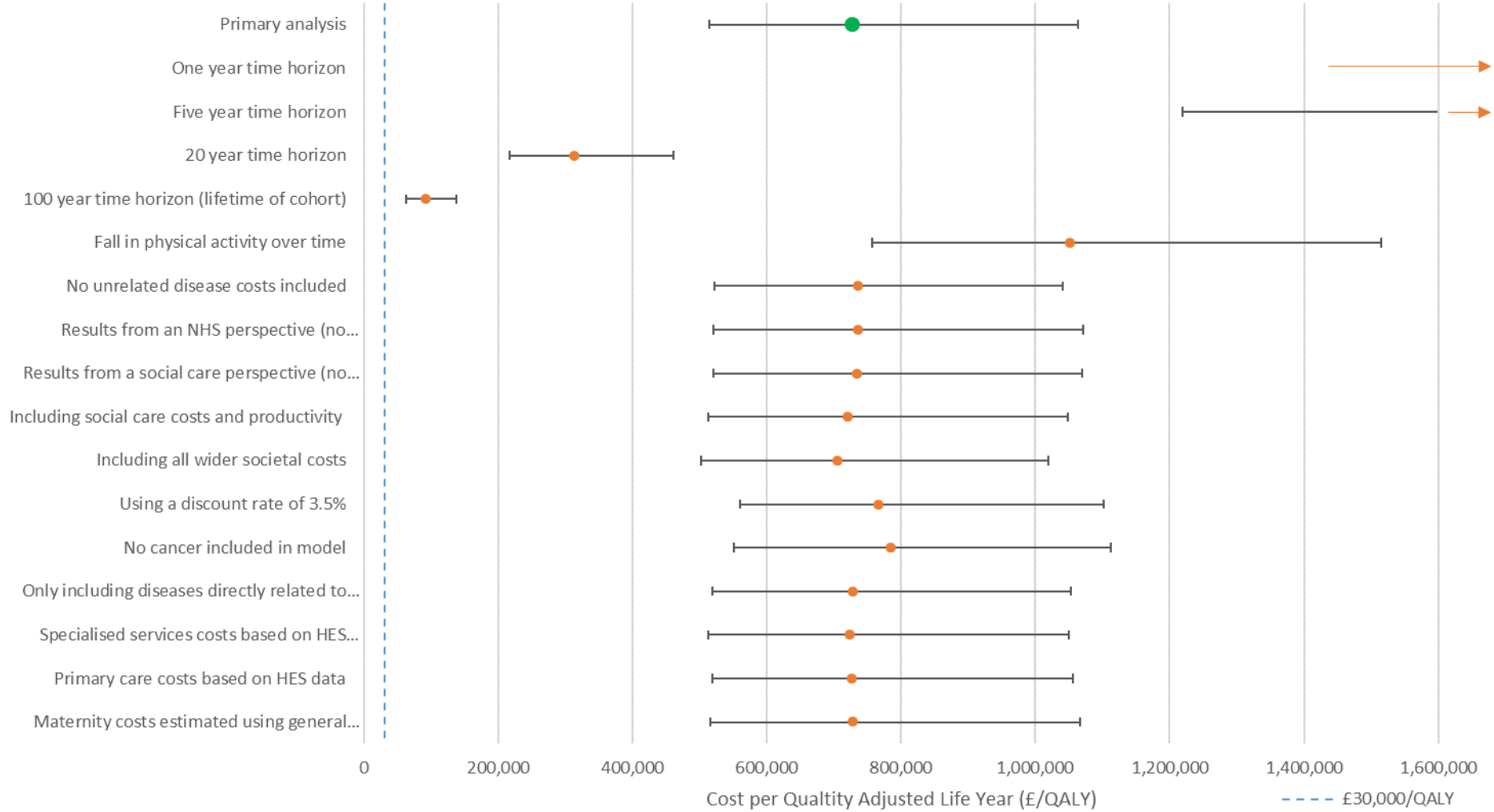
The bars illustrate the width of the uncertainty interval when only one parameter or collection of related parameters is allowed to vary according to their probability distribution. QALY, quality adjusted life year; UI, uncertainty interval; IHD, ischaemic heart disease

Figure 5.16a Results of sensitivity analyses for the physical activity intervention



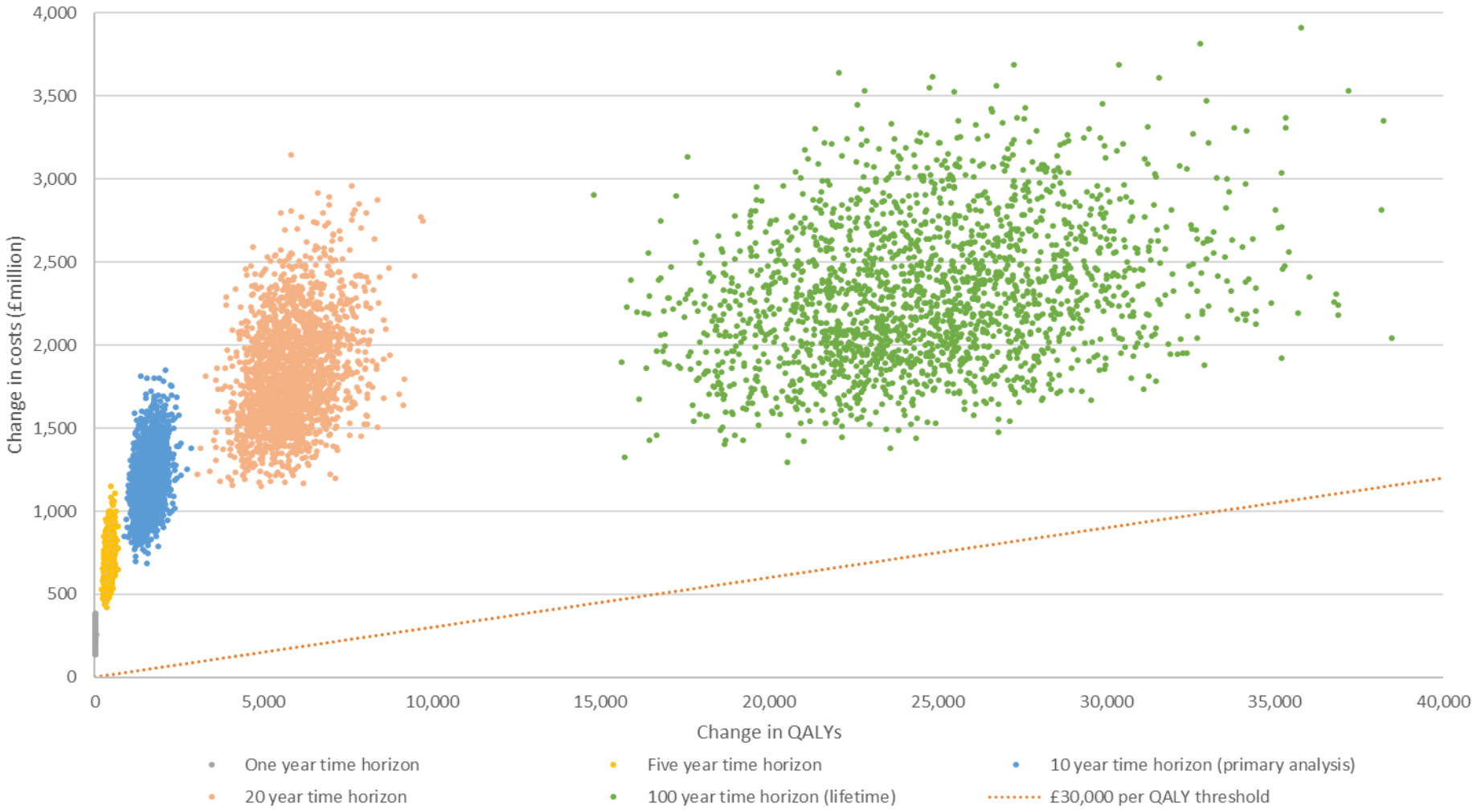
QALY, quality adjusted life year; HES, hospital episode statistics

Figure 5.16b Results of deterministic sensitivity analyses for the physical activity intervention with a smaller scale on the x-axis



QALY, quality adjusted life year; HES, hospital episode statistics

Figure 5.17 Cost effectiveness plane showing results of the physical activity intervention with different time horizons



QALY, quality adjusted life year

5.3 Discussion of results

Reformulating food to have less salt would be significantly more cost-effective than expanding the Be Active scheme to increase access to leisure centres across England. The physical activity intervention had greater implementation costs, lower health and social care savings, and led to fewer QALYs gained than salt reformulation.

Although for each intervention the uncertainty intervals generally overlapped between men and women, results of the primary analyses indicated that men would experience greater relative and absolute health benefits in terms of reductions in QALYs, and declines in cumulative incidence and incidence rates of disease (the only exception being the reduction in breast cancer rates following the physical activity intervention). This is due to men having a greater baseline burden of disease and because both interventions led to larger improvements in the targeted risk factor among men than women. Due to model input parameters generally being right-skewed, mean values were slightly larger in magnitude than medians for each intervention, and within the 95% UIs.

5.3.1 Diet intervention

Reformulating food to meet the 2017 salt targets led to both gains in QALYs and overall cost savings, and was therefore dominant. The cost per QALY was -£17,050 (-£39,550 to £38,550) with savings primarily arising from social care (£724.8m [£306.5m to £1,200m]) rather than the NHS (£141.7m [95% UI £58.8m to £235.9m]). This meant that the intervention was dominant from a social care perspective but not from an NHS perspective. The difference between NHS and social care savings was largely due to reductions in stroke incidence. Stroke is assumed to have additional social care costs over and above those estimated solely from a change in quality of life (see section 3.1.2.3). The largest gains in QALYs per person reflect the relationship between reduction in risk factor exposure (greatest among 18-34 year olds) and baseline disease burden (greatest among older age groups). Those aged 70 to 80 years had the greatest gain in QALYs per person and because social care costs are incurred at age 75 and over, the same age group had the largest social

care savings. By contrast, per person and aggregate healthcare savings peaked for 65 to 69 year olds, who also had the greatest aggregate gain in QALYs, all because of their larger population size (figures 5.4 and 5.5).

Uncertainty in the primary analysis was principally due to uncertainty in how salt consumption affects blood pressure (figure 5.6). This parameter was taken from He et al.[223] and is widely used in published studies estimating the effect of changing salt consumption on health.[66,115,219] The impact of changing this parameter is discussed in sections 4.1.3.2 and 6.1.3. Uncertainty in social care costs, industry costs, utilities, and the blood pressure minimum risk threshold (the systolic blood pressure below which there is no added health benefit from further blood pressure reductions) also contributed to the model's uncertainty but to a much lesser extent than the relationship between salt consumption and blood pressure. Importantly, despite the challenges of estimating government (section 4.1.1) and disease (section 3.1) costs, uncertainty in these parameters had little impact on the final cost per QALY uncertainty.

Some of the methodological assumptions used when estimating disease costs and societal costs (section 3.1) were tested with sensitivity analyses and had little impact on the final cost per QALY estimate (table 5.2 and figure 5.7), suggesting the primary analysis is fairly robust to these assumptions within the perspective and time horizon specified. These assumptions included using different methods for allocating healthcare costs to diseases, incorporating different societal costs, and varying the diseases included in the model beyond IHD and stroke.

Chapter six compares these results with other models simulating similar interventions and explores the reasons for differences in outcomes.

5.3.2 Physical activity intervention

The disease which saw the largest fall in cumulative incidence following the physical activity intervention was diabetes (figure 5.11). This is because there was a much greater baseline prevalence of diabetes in the population compared to other modelled diseases (46 cases per

100,000 compared to 8, 9, 8, and 4 per 100,000 for IHD, stroke, breast cancer, and colorectal cancer respectively). Furthermore, of the diseases modelled, the effect of changing physical activity on disease risk is greatest for diabetes.[242,243]

Cost savings were similar between the NHS (£10.4m [£7.2m to £14.6m]) and social care (£14.2m [£7.8m to £22.4m]) because, compared to the diet intervention, health benefits in terms of QALYs gained were greatest among 40 to 60 year olds, before social care costs are incurred (figure 5.12). When summed over the entire population, social care savings among over 75 year olds did not significantly outweigh the NHS savings accrued by younger age groups (figure 5.13). However as with the diet intervention, for those aged over 70 years, social care savings significantly outweighed NHS savings. Despite per person and aggregate QALY gains being greatest among 40-60 year olds, 85 to 89 year olds had the lowest cost per QALY of any age group. This is because 85 to 89 year olds had a larger baseline burden of disease with accrued cost savings from both the NHS and social care (figure 5.14).

The main sources of uncertainty in the physical activity intervention were the parameters describing the relationship between physical activity and disease, the intervention costs, and the utilities. The relationships between physical activity and disease were taken from reviews by Wahid et al.[242,243] These studies are used because they describe the dose-response relationship between physical activity and disease and are comparable across diseases included in the model (discussed in more detail in section 4.1.3.2). Intervention costs are the same as those reported by Frew et al. and their large contribution to model uncertainty is in part explained by the triangular distribution used which will sample more frequently from upper and lower estimates than normal, lognormal, or gamma distributions used elsewhere in the model (appendix four).[144]

As with the diet intervention, the time horizon used had a significant impact on results (table 5.4, figures 5.16 and 5.17). The only other sensitivity analysis that meaningfully altered the cost per QALY estimate was including a fall in physical activity over time (replicated from Frew et al.).[144] Other sensitivity analyses only slightly differed from the

primary analysis cost per QALY estimate meaning that as with the diet intervention, the result is likely to be robust to the assumptions tested by the sensitivity analyses within the context of the intervention tested, and time horizon and perspective used. Further evaluation of the stability of these results is explored in chapter six, concerning the validity of the PRIMEtime CE.

Chapter 6. Validation

Models simplify reality in order to estimate and predict real world outcomes, and model validation aims to understand how accurately a given model represents reality. There are different approaches to validating health economic and NCD models,[252–254] and this thesis uses the Assessment of the Validation Status of Health-Economic decision models (AdViSHE) tool.[254] The AdViSHE tool was developed by a panel of expert health economic modellers and decision makers to help model developers prioritise validation efforts. The tool divides validation into five parts:

A. validation of the conceptual model;
B. input data validation;
C. validation of the computerised model;
D. operational validation; and
E. other validation techniques.

6.1 Part A: Validation of the conceptual model

Validation of the conceptual model involves face validity testing by sharing it with experts, and cross validity testing by comparing it with other conceptual models in the literature.

In order to test the face validity of the conceptual model, the model was shared with stakeholders (expert decision makers) who were asked if it was an accurate representation of reality, and whether the logic of the model made sense. The results of the stakeholder engagement are reported in section 2.1.2 and appendix three. In summary, nine of 12 stakeholders responded that the model's logic was plausible and had face validity, with the remaining three suggesting that:

1. the range of interventions was too limited;
2. that it was unclear how productivity would be estimated; and
3. that wider social value outcomes should be included.

These points were addressed in chapter two when finalising the model's structure. The PRIMEtime CE conceptual model structure was not shared with expert modellers, however the model is based on PRIMEtime, a previously published model which has been peer-reviewed and therefore critiqued by model experts.[131]

The cross validity of PRIMEtime CE has been tested by comparing the model with similar multistate life table models in Australia and New Zealand (the Australian model is discussed in section 1.4.3.1) upon which PRIMEtime was originally based.[114,255,256] Both PRIMEtime CE and the Australian and New Zealand models describe the same relationships between diet and disease, and physical activity and disease. Furthermore, PRIMEtime CE has the same basic model structure with the same primary disease outcomes as other models simulating the relationships between salt consumption, physical activity, and disease: changing salt consumption affects blood pressure and then CVD outcomes,[61,66,96,219,257,258] and changing physical activity affects IHD, stroke, diabetes, colorectal cancer, and breast cancer.[60,73,85,144,244,245]

6.2 Part B: Input data validation

Input data validation assesses the appropriateness of data used to parameterise the model. The AdViSHE tool divides this into two parts, face validity testing of the input data and model fit testing.

The epidemiological data used to parameterise the model has been judged by experts as part of PRIME and PRIMEtime's peer-review and found to be of sufficient quality for publication.[131,132] Additional input data used by PRIMEtime CE are the costs, utilities, data relating to the interventions tested, and inputs parameterising the relationships between physical activity and disease. PRIMEtime CE has not yet been peer-reviewed for publication in a journal (this will form part of future work), however the costing methodology was peer-reviewed and presented as part of the Health Economists' Study Group meeting in 2017.[259]

This is a biannual meeting of health economists where accepted papers are presented and critiqued by designated discussants and attendees rather than by the author. Feedback from attendees was that this was a novel approach to estimating costs which did not have any obvious methodological flaws or major limitations. The utilities used in PRIMEtime CE are from a peer-reviewed journal and have also been peer-reviewed as part of other health economic models.[85,115,148,200] The majority of the data used to parameterise the interventions has not yet been judged by experts, an exception is the economic evaluation of Be Active which was used to inform the physical activity intervention.[144] Finally, the methodology used to parameterise the effect of changing physical activity on disease is from two meta-analyses using the same methods by the same research team, only one of which has been peer-reviewed.[242,243] Use of these data within PRIMEtime CE is also yet to be judged by experts.

Model fit testing asks whether the appropriate statistical tests have been performed where input parameters are based on regression models. Although many model input parameters used by PRIMEtime CE are based on regression models (such as utilities, disease relative risks, social care costs, and others), they are all from peer-reviewed secondary data sources, as shown in appendix four (except for Wahid et al.[242]). Where primary data have been used, such as estimating NHS costs or how the interventions affect behavioural risk factors, no regression models were required. For estimating the effect of the physical activity intervention on average MET hours/week, it was assumed that physical activity levels follow a log-normal distribution which was the best distribution available to fit the data of several tested (see section 4.1.2). However, these results were only for illustrative purposes and not used to parameterise PRIMEtime CE.

6.3 Part C: Validation of the computerised model

Validating the computerised model involves external review, extreme value testing, testing of traces, and unit testing.

As with other novel aspects of PRIMEtime CE, the computerised model is yet to be externally reviewed by experts, although this is planned for the future. Firstly, the Burden of Disease Epidemiology, Equity & Cost-Effectiveness (BODE^3) modelling team led by Prof Blakely at the University of Otago, New Zealand[260] will evaluate the model as part of a model comparison exercise, and secondly the model will be made available to peer-reviewers when its details are submitted for publication.

Extreme input values were modelled in PRIMEtime CE to see if outcomes were plausible and to detect any coding errors; results are shown in table 6.1. Results for inputting either a 0% or 100% reduction in salt consumption as reported by NDNS food diaries are in line with what would be expected given the gains in QALYs and cost savings of the diet intervention primary analysis. Similarly, results with 100% of the population achieving recommended levels of physical activity or being sedentary are not implausible. Both these policies assume that individuals enrol in the intervention, however under the scenario where 100% of the population are achieving recommended physical activity levels there are no additional health benefits from the intervention, and where 100% of the population are sedentary, the model assumes that the sedentary population does not change (see section 4.2.1) meaning the intervention has no health benefit yet costs remain largely the same.

Table 6.1 Testing of PRIMEtime CE with extreme input values

Scenario	Change in QALYs	NHS savings	Social care savings	Intervention costs
Diet intervention primary analysis	15,444	143,566,300	732,265,374	597,785,271
0% reduction in salt consumption	0	0	0	597,785,271
100% reduction in salt consumption	81,676	751,842,904	3,984,611,598	597,785,271
Physical activity intervention primary analysis	1,593	10,387,199	14,306,598	1,199,599,777

100% of the population achieve recommended physical activity levels	169,533	995,815,098	5,552,126,423	1,199,653,797
100% of the population are sedentary	-231,622	-1,467,160,814	-3,925,887,745	1,199,536,675

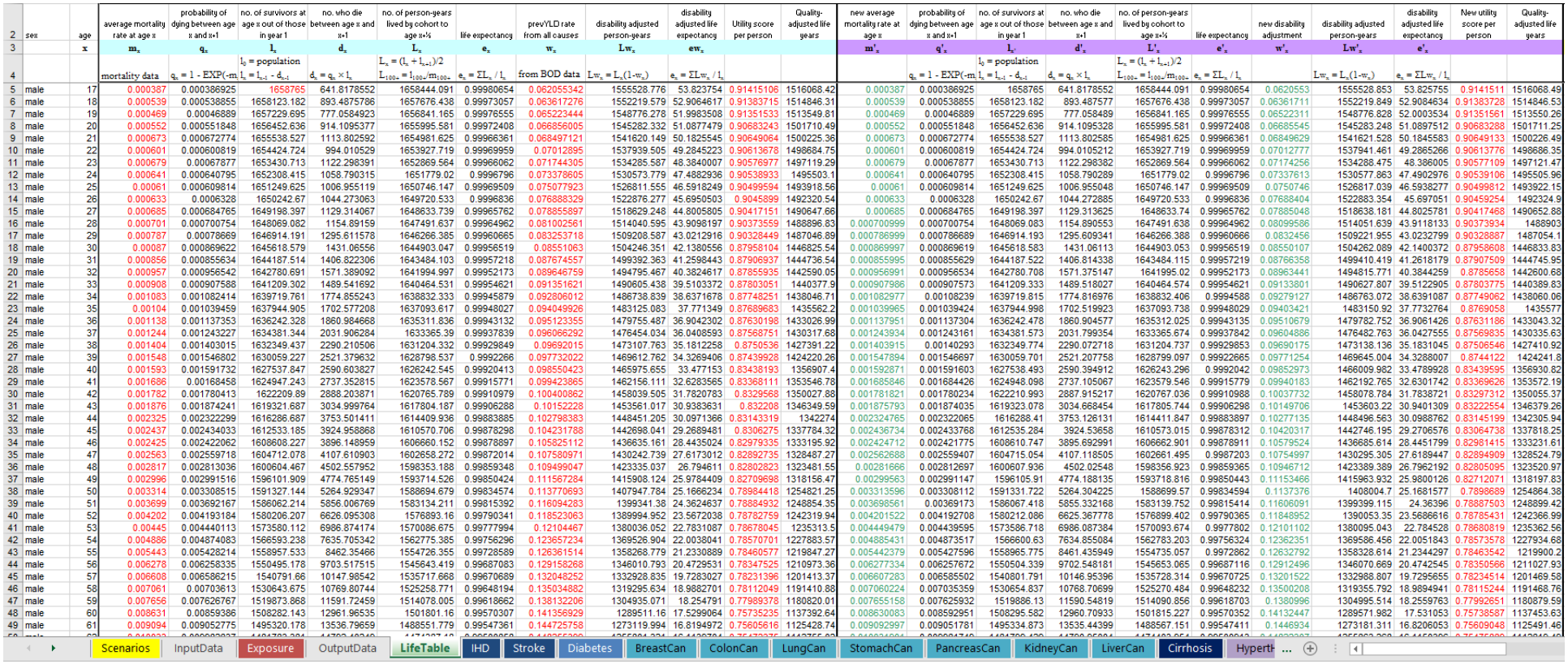
All results are point estimates using mean parameter estimates and therefore primary analysis results are not the same as the median results of the Monte Carlo simulation reported in chapter five. QALY, quality adjusted life year

Testing of traces involves tracking individuals through the model by listing the number of people simulated and their disease status at various time points to ensure that the entire modelled cohort is accounted for. This is explicitly done in multistate life table models as the simulated population is listed in the life table, enabling the cohort and their disutility to be accounted for at all years. The life table in PRIMEtime CE is shown in figure 6.1. The population illustrated is males aged 15-19 years, with data in the life table representing how this population's mortality and morbidity changes over time from an average age of 17. The left side of the life table coloured light blue represents the baseline population, and the right side coloured purple is the population following the physical activity intervention.

Figure 6.1 The PRIMEtime CE life table, with the modelled population and utility scores highlighted

Number of people in cohort

Average utility score per person



	sex	age	average mortality rate at age x	probability of dying between age x and x+1	no. of survivors at age x out of those in year 1	no. who die between age x and x+1	no. of person-years lived by cohort to age x+½	life expectancy	prevYLD rate from all causes	disability adjusted person-years	disability adjusted life expectancy	Utility score per person	Quality-adjusted life years	new average mortality rate at age x	probability of dying between age x and x+1	no. of survivors at age x out of those in year 1	no. who die between age x and x+1	no. of person-years lived by cohort to age x+½	life expectancy	new disability adjustment	disability adjusted person-years	disability adjusted life expectancy	New utility score per person	Quality-adjusted life years
2			m_x	q_x	l_x	d_x	L_x	e_x	w_x	Lw_x	ew_x			m'_x	q'_x	l'_x	d'_x	L'_x	e'_x	w'_x	Lw'_x	e'_x		
4			mortality data	$q_x = 1 - EXP(-m_x)$	l_0 = population $l_x = l_{x-1} - d_{x-1}$	$d_x = q_x \times l_x$	$L_x = (l_x + l_{x+1})/2$ $L_{100+} = l_{100+}/m_{100+}$	$e_x = \Sigma L_x / l_x$	from BOD data	$Lw_x = L_x(1-w_x)$	$e_x = \Sigma Lw_x / l_x$				$q'_x = 1 - EXP(-m'_x)$	l_0 = population $l'_x = l'_{x-1} - d'_{x-1}$	$d'_x = q'_x \times l'_x$	$L'_x = (l'_x + l'_{x+1})/2$ $L_{100+} = l_{100+}/m_{100+}$	$e'_x = \Sigma L'_x / l'_x$		$Lw'_x = L'_x(1-w'_x)$	$e'_x = \Sigma Lw'_x / l'_x$		
5	male	17	0.000387	0.000386925	1658765	641.8178552	1658444.091	0.99980654	0.062055342	1555528.776	53.823754	0.91415106	1516068.42	0.000387	0.000386925	1658765	641.8178552	1658444.091	0.99980654	0.0620553	1555528.853	53.825755	0.9141511	1516068.49
6	male	18	0.000539	0.000538855	1658123.182	893.4875786	1657676.438	0.99973057	0.063617276	1552219.579	52.9064617	0.91383715	1514846.31	0.000539	0.000538855	1658123.182	893.487577	1657676.438	0.99973057	0.06361711	1552219.849	52.9084634	0.91383728	1514846.53
7	male	19	0.000469	0.00046889	1657229.695	777.0584923	1656841.165	0.99976555	0.065223444	1548776.278	51.9983508	0.91351533	1513549.81	0.000469	0.00046889	1657229.695	777.058489	1656841.165	0.99976555	0.06522311	1548776.828	52.0003534	0.91351561	1513550.26
8	male	20	0.000552	0.000551848	1656452.636	914.1095377	1655995.581	0.99972408	0.066856005	1545282.332	51.0877479	0.90683243	1501710.49	0.000552	0.000551848	1656452.636	914.1095328	1655995.581	0.99972408	0.06685545	1545283.248	51.0897512	0.90683288	1501711.25
9	male	21	0.000673	0.000672774	1655538.527	1113.802592	1654981.625	0.99966361	0.068497121	1541620.149	50.1825545	0.90649064	1500225.36	0.000673	0.000672774	1655538.527	1113.802585	1654981.625	0.99966361	0.06849629	1541621.528	50.1845583	0.90649133	1500226.49
10	male	22	0.000601	0.000600819	1654424.724	994.010529	1653927.719	0.99969959	0.07012895	1537939.505	49.2845223	0.90613678	1498684.75	0.000601	0.000600819	1654424.724	994.0105212	1653927.719	0.99969959	0.07012777	1537941.461	49.2865266	0.90613776	1498686.35
11	male	23	0.000679	0.00067877	1653430.713	1122.298391	1652869.564	0.99966062	0.071744305	1534285.587	48.3840007	0.90576977	1497119.29	0.000679	0.00067877	1653430.713	1122.298382	1652869.564	0.99966062	0.07174256	1534288.475	48.386005	0.90577109	1497121.47
12	male	24	0.000641	0.000640795	1652308.415	1058.790315	1651779.02	0.9996796	0.073378605	1530573.779	47.4882936	0.90538933	1495503.1	0.000641	0.000640795	1652308.415	1058.790289	1651779.02	0.9996796	0.07337613	1530577.863	47.4902976	0.90539106	1495505.96
13	male	25	0.00061	0.000609814	1651249.625	1006.955119	1650746.147	0.99969509	0.075077923	1526811.555	46.5918249	0.90499594	1493918.56	0.00061	0.000609814	1651249.625	1006.955048	1650746.147	0.99969509	0.0750746	1526817.039	46.5938277	0.90499812	1493922.15
14	male	26	0.000633	0.0006328	1650242.67	1044.273063	1649720.533	0.9996836	0.076888329	1522876.277	45.6950503	0.9045899	1492320.54	0.000633	0.0006328	1650242.67	1044.272885	1649720.533	0.9996836	0.07688404	1522883.354	45.697051	0.90459254	1492324.9
15	male	27	0.000685	0.000684765	1649198.397	1129.314067	1648633.739	0.99965762	0.078855897	1518629.248	44.8005805	0.90417151	1490647.66	0.000685	0.000684765	1649198.397	1129.313625	1648633.74	0.99965762	0.07885048	1518638.181	44.8025781	0.90417468	1490652.88
16	male	28	0.000701	0.000700754	1648069.082	1154.89159	1647491.637	0.99964962	0.081002561	1514040.595	43.9098197	0.90373559	1488896.83	0.000700999	0.000700754	1648069.083	1154.890553	1647491.638	0.99964962	0.08099586	1514051.639	43.9118133	0.90373934	1488903
17	male	29	0.000787	0.00078669	1646914.191	1295.611578	1646266.385	0.99960665	0.083253718	1509208.587	43.0212916	0.90328449	1487046.89	0.000786999	0.000786689	1646914.193	1295.609341	1646266.388	0.99960666	0.0832456	1509221.955	43.0232799	0.90328887	1487054.1
18	male	30	0.00087	0.000869622	1645618.579	1431.06556	1644903.047	0.99956519	0.08551063	1504246.351	42.1380556	0.87958104	1446825.54	0.000869997	0.000869619	1645618.583	1431.06113	1644903.053	0.99956519	0.08550107	1504262.089	42.1400372	0.87958608	1446833.83
19	male	31	0.000856	0.000855634	1644187.514	1406.822306	1643484.103	0.99957218	0.087674557	1499392.363	41.2598443	0.87906937	1444736.54	0.000855995	0.000855629	1644187.522	1406.814338	1643484.115	0.99957219	0.08766358	1499410.419	41.2618179	0.87907509	1444745.95
20	male	32	0.000957	0.000956542	1642780.691	1571.389092	1641994.997	0.99952173	0.089646759	1494795.467	40.3824617	0.87855935	1442590.05	0.000956991	0.000956534	1642780.708	1571.375147	1641995.02	0.99952173	0.08963441	1494815.771	40.3844259	0.8785658	1442600.68
21	male	33	0.000908	0.000907588	1641209.302	1489.541692	1640464.531	0.99954621	0.091351621	1490605.438	39.5103372	0.87803051	1440377.9	0.000907986	0.000907573	1641209.333	1489.518027	1640464.574	0.99954621	0.09133801	1490627.807	39.5122905	0.87803775	1440389.83
22	male	34	0.001083	0.001082414	1639719.761	1774.855243	1638832.333	0.99945879	0.092806012	1486738.839	38.6371678	0.87748251	1438046.71	0.001082977	0.00108239	1639719.815	1774.816976	1638832.406	0.9994588	0.09279127	1486763.072	38.6391087	0.87749062	1438060.06
23	male	35	0.00104	0.001039459	1637944.905	1702.577208	1637093.617	0.99948027	0.094049926	1483125.083	37.771349	0.87689683	1435562.2	0.001039965	0.001039424	1637944.998	1702.519923	1637093.738	0.99948029	0.09403421	1483150.92	37.7732764	0.8769058	1435577
24	male	36	0.001138	0.001137353	1636242.328	1860.984668	1635311.836	0.99943132	0.095123355	1479755.487	36.9042302	0.87630198	1433026.99	0.001137951	0.001137304	1636242.478	1860.904577	1635312.025	0.99943135	0.09510679	1479782.752	36.9061426	0.87631186	1433043.32
25	male	37	0.001244	0.001243227	1634381.344	2031.906284	1633365.39	0.99937839	0.096066292	1476454.034	36.0408593	0.87568751	1430317.68	0.001243934	0.001243161	1634381.573	2031.799354	1633365.674	0.99937842	0.09604886	1476482.763	36.0427555	0.87569835	1430335.63
26	male	38	0.001404	0.001403015	1632349.437	2290.210506	1631204.332	0.99929849	0.09692015	1473107.763	35.1812258	0.8750536	1427391.22	0.001403915	0.00140293	1632349.774	2290.072718	1631204.737	0.99929853	0.09690175	1473138.136	35.1831045	0.87506546	1427410.92
27	male	39	0.001548	0.001546802	1630059.227	2521.379632	1628798.537	0.9992266	0.097732022	1469612.762	34.3269406	0.87439928	1424220.26	0.001547894	0.001546697	1630059.701	2521.207758	1628799.097	0.99922665	0.09771254	1469645.004	34.3288007	0.8744122	1424241.8
28	male	40	0.001593	0.001591732	1627537.847	2590.603827	1626242.545	0.99920413	0.098550423	1465975.655	33.477153	0.83438193	1356907.4	0.001592871	0.001591603	1627538.493	2590.394912	1626243.296	0.9992042	0.09852973	1466009.982	33.4789928	0.83439595	1356930.82
29	male	41	0.001686	0.00168458	1624947.243	2737.352815	1623578.567	0.99915771	0.099423865	1462156.111	32.6283565	0.83368111	1353546.78	0.001685846	0.001684426	1624948.098	2737.105067	1623579.546	0.99915779	0.09940183	1462192.765	32.6301742	0.83369626	1353572.19
30	male	42	0.001782	0.001780413	1622209.89	2888.203871	1620765.789	0.99910979	0.100400862	1458039.505	31.7820783	0.8329568	1350027.88	0.001781821	0.001780234	1622210.993	2887.915217	1620767.036	0.99910988	0.10037732	1458078.784	31.7838721	0.83297312	1350055.37
31	male	43	0.001876	0.001874241	1619321.687	3034.999764	1617804.187	0.99906288	0.10152228	1453561.017	30.9383631	0.832208	1346349.59	0.001875793	0.001874035	1619323.078	3034.668454	1617805.744	0.99906298	0.10149706	1453603.22	30.9401309	0.83222554	1346379.26
32	male	44	0.002325	0.002322299	1616286.687	3753.501411	1614409.936	0.99883885	0.102798383	1448451.205	30.0971366	0.83143319	1342274	0.002324765	0.002322065	1616288.41	3753.126131	1614411.847	0.99883897	0.10277135	1448496.563	30.0988762	0.83145199	1342305.94
33	male	45	0.002437	0.002434033	1612533.185	3924.958868	1610570.706	0.99878298	0.104231788	1442698.041	29.2689481	0.8306275	1337784.32	0.002436734	0.002433768	1612535.284	3924.53658	1610573.015	0.99878312	0.10420317	1442746.195	29.2706576	0.83064738	1337818.25
34	male	46	0.002425	0.002422062	1608608.227	3896.148959	1606660.152	0.99878897	0.105825112	1436635.161	28.4435024	0.82979335	1333195.92	0.002424712	0.002421775	1608610.747	3895.692991	1606662.901	0.99878911	0.10579524	1436685.614	28.4451799	0.82981415	1333231.61
35	male	47	0.002563	0.002559718	1604712.078	4107.610903	1602658.272	0.99872014	0.107580971	1430242.739	27.6173012	0.82892735	1328487.27	0.002562688	0.002559407	1604715.054	4107.118505	1602661.495	0.9987203	0.10754997	1430295.305	27.6189447	0.82894909	1328524.79
36	male	48	0.002817	0.002813036	1600604.467	4502.557952	1598353.188	0.99859348	0.109499047	1423335.037	26.794611	0.82802823	1323481.55	0.00281666	0.002812697	1600607.936	4502.02548	1598356.923	0.99859365	0.10946712	1423389.389	26.7962192	0.82805095	1323520.97
37	male	49	0.002996	0.002991516	1596101.909	4774.765149	1593714.526	0.99850424	0.111567284	1415908.124	25.9784409	0.82709698	1318156.47	0.00299563	0.002991147	1596105.91	4774.188135	1593718.816	0.99850443	0.11153466	1415963.932	25.9800126	0.82712071	1318197.83
38	male	50	0.003314	0.003308515	1591327.144	5264.929347	1588694.679	0.99834574	0.113770693	1407947.784	25.1666234	0.78984418	1254821.25	0.003313596	0.003308112	1591331.722	5264.304225	1588699.57	0.99834594	0.1137376	1408004.7	25.1681577	0.7898689	1254864.38
39	male	51	0.003699	0.003692167	1586062.214	5856.006769	1583134.211	0.99815392	0.116094283	1399341.38	24.3624637	0.78884932	1248854.35	0.003698561	0.00369173	1586067.418	5855.332168	1583139.752	0.99815414	0.11606091	1399399.115	24.36396	0.78887503	1248899.42
40	male	52	0.004202	0.004193184	1580206.207	6626.095308	1576893.16	0.99790341	0.118523063	1389994.952	23.5672038	0.78782759	1242319.94	0.004201522	0.004192708	1580212.086	6625.367778	1576899.402	0.99790365	0.11848952	1390053.35	23.5686616	0.78785431	1242366.99
41	male	53	0.00445	0.004440113	1573580.112	6986.874174	1570086.675	0.99777994	0.12104467	1380036.052	22.7831087	0.78678045	1235313.5	0.004449479	0.004439595	1573586.718	6986.087384	1570093.674	0.9977802	0.12101102	1380095.043	22.784528	0.78680819	1235362.56
42	male	54	0.004886	0.004874083	1566593.238	7635.705342	1562775.385	0.99756296	0.123657234	1369526.904	22.0038041	0.78570701	1227883.57	0.004885431	0.00487351	1566600.63	7634.855084	1562783.203	0.99756324	0.12362351	1369586.456	22.0051843	0.78573578	1227934.68
43	male	55	0.005443	0.005428214	1558957.533	8462.35466	1554726.355	0.99728589	0.126361514	1358268.779	21.2330889	0.78460577	1219847.27	0.005442379	0.005427596	1558965.775	8461.435949	1554735.057	0.9972862	0.12632792	1358328.614	21.2344297	0.78463542	1219900.2
44	male	56	0.006278	0.006258335	1550495.178	9703.517515	1545643.419	0.99687083	0.129158268	1346010.793	20.4729531	0.78347525	1210973.36	0.006277334	0.006257672	1550504.339	9702.548181	1545653.065	0.99687116	0.12912496	1346070.669	20.4742545	0.78350566	1211027.93
45	male	57	0.006608	0.006586215	1540791.66	10147.98542	1535717.668	0.99670689	0.132048252	1332928.835	19.7283027	0.78231396	1201413.37	0.006607283	0.006585502	1540801.791	10146.95396	1535728.314	0.99670725	0.13201522	1332988.807	19.7295655	0.78234514	1201469.58
46	male	58	0.007061	0.00703613	1530643.675	10769.80744	1525258.771	0.99648194	0.135034882	1319295.634	18.9882701	0.78112049	1191410.88	0.007060224	0.007035359	1530654.837	10768.70699	1525270.484	0.99648232	0.13500208	1319355.792	18.9894941	0.78115244	1191468.76
47	male	59	0.007656	0.007626767	1519873.868	11591.72459	1514078.005	0.99618662	0.138132206	1304935.071	18.254791	0.77989378	1180820.01	0.007655158	0.007625932	1519886.13	11590.54819	1514090.856	0.99618703	0.1380996	1304995.514	18.2559763	0.77992651	1180879.59
48	male	60	0.008631	0.00859386	1508282.143	12961.96535	1501801.16	0.99570307	0.141356929	1289511.16	17.5299064	0.75735235	1137392.64	0.008630083	0.00859295	1508295.582	12960.70933	1501815.227	0.99570352	0.14132447	1289571.982	17.531053	0.75738587	1137453.63
49	male	61	0.009094	0.009052775	1495320.178	13536.79659	1488551.779	0.99547361	0.144725758	1273119.994	16.8194972	0.75605616	1125428.74	0.009092997	0.009051781	1495334.873	13535.44399	1488567.151	0.99547411	0.1446934	1273181.311	16.8206053	0.75609048	1125491.46

Scenarios | InputData | Exposure | OutputData | LifeTable | IHD | Stroke | Diabetes | BreastCan | ColonCan | LungCan | StomachCan | PancreasCan | KidneyCan | LiverCan | Cirrhosis | Hyperth ...

Population shown is males aged 15-19 at baseline. The light blue coloured columns on the left represent the baseline scenario, and the purple columns on the right show the cohort post intervention

Unit testing is the separate analyses of model submodules. For PRIMEtime CE, the two main submodules are how the intervention affects a behavioural risk factor, and how the change in behavioural risk factor then affects disease outcomes. The effect of the intervention on the behavioural risk factor has been separately tested, with results detailed in chapter four demonstrating that the intervention submodule of PRIMEtime CE produces reasonable results. The effect of changing the risk factor on health was also separately investigated when comparing results of PRIMEtime CE with other models simulating similar interventions, this is discussed in more detail under section 6.4.2 below.

6.4 Part D: Operational validation

This part of AdViSHE aims to validate model outcomes. Four different forms of operational validation are presented: face validity testing of model outcomes, cross validation testing, validation against outcomes using alternative input data, and validation against empirical data.

6.4.1 Face validity testing of model outcomes

To assess the face validity of outcomes, the modelled results should be shared with appropriate experts to see if the results make sense (both to decision makers and modellers). Results have been presented to health economists at the Department of Health and PHE with verbal feedback suggesting that they are reasonable. Formal stakeholder feedback is currently being sought via an online questionnaire (the final part of stakeholder engagement, see appendix one). Finally, the results will also be shared with model experts when sent for publication in a peer-reviewed journal, which is planned for the near future.

6.4.2 Cross validation testing and validation of outcomes using alternative input data

Cross validation testing of model outcomes involves comparing results with other models addressing similar problems to see what differences arise and to identify possible

explanations. Validating outcomes using alternative input data identifies how results might be affected when using different data to parameterise the model.

Cross validation testing of PRIMEtime CE outcomes was explored in detail by comparing results of the diet intervention with similar analyses published by the UK Health Forum using their microsimulation model,[115] and by Collins et al. using the population level model, IMPACT CHD.[219] Results of the physical activity intervention were compared with results from Frew et al.'s bespoke population level Markov model.[144] In order to help explain difference in outcomes between Frew et al. and PRIMEtime CE, various input data from Frew et al. were used and results were again compared. These models were chosen for comparison because they either used a similar approach to estimating disease costs (the UK Health Forum model), or they provided input parameters for the interventions simulated in this thesis (IMPACT CHD and Frew et al.).

6.4.2.1 Cross validation testing using the UK Health Forum model

The UK Health Forum estimated the effect on CVD and NHS costs of four different salt consumption scenarios in England over a period of 20 years from 2014 to 2034:

1. following current trends and reaching 8.0g salt per day consumed by men and 6.1g per day consumed by women in 2034;
2. a best case where mean salt consumption for both men and women reaches 5.0g per day in 2034;
3. a worst case where salt consumption stays at 2014 levels of 8.9g per day for men, and 6.8g per day for women;
4. an expected scenario where mean consumption reaches 7.0g per day for men and women in 2034.[115]

In order to estimate the effect on health of the UK Health Forum scenarios using PRIMEtime CE, PRIMEtime CE was adapted to include many of the UK Health Forum microsimulation model baseline settings. For each of the scenarios, 2014 age and sex specific salt consumption in PRIMEtime CE was adjusted such that the average baseline

consumption was 8.9g per day for men and 6.8g per day for women to match that in the UK Health Forum scenario. Salt consumption was then assumed to change with a linear trend over 20 years towards the 2034 target, with the sex specific salt consumption target in each of the four scenarios being the same across all age groups. A discount rate of 0% was used and the time horizon was set to 20 years with only those over the age of 18 years included. As with PRIMEtime CE, the UK Health Forum model used a top down method to estimate the cost per prevalent case to the NHS with disaggregated NHS England programme budgeting data. The 2013/14 disease costs used in PRIMEtime CE were replaced by the aggregate 2012/13 disease cost estimates for MI and stroke used by the UK Health Forum; societal costs, intervention costs, and costs from other unrelated diseases were removed.

Other differences in input data between the two models remained. The UK Health Forum estimated the effect on MI rather than IHD, although the impact of this on model outcomes is likely to be minimal as costs were taken from the CHD NHS England programme budgeting category (the same as PRIMEtime CE), and the relative risk of CHD following a change in blood pressure was also the same in both models.[224] Baseline disease incidence and prevalence data differed between the two models. The average baseline prevalence rates for MI and stroke used by the UK Health Forum were 1,474 per 100,000 and 1,188 per 100,000 compared to 1,819 per 100,000 and 2,141 per 100,000 for IHD and stroke in PRIMEtime CE. Baseline incidence rates were also higher in PRIMEtime CE at 143 per 100,000 and 178 per 100,000 compared to 131 per 100,000 and 108 per 100,000 used by the UK Health Forum. The UK Health Forum used routinely reported British Heart Foundation data from between 2006 and 2010[261,262] for incidence, prevalence, and case fatality rates (based on various sources including the HSE and GPRD). By contrast, PRIMEtime CE used the sources listed in appendix four, including primary research studies and DISMOD II. Larger baseline prevalence and incidence rates in PRIMEtime CE would mean that the model is likely to estimate larger absolute changes in disease cumulative incidence than the UK Health Forum.

Estimates of baseline salt consumption by age and sex differed between the models. Both models used NDNS data with the UK Health Forum model estimating trends for each sex and for three age categories for men (20-39 years, 40-59 years, over 60 years) and two for women (20-39 years and over 40 years) based on urinary sodium excretion studies from between 2001 and 2011.[263] The proportion of the population in each age and sex group consuming less than 6g, between 6g and 12g and more than 12g salt per day was estimated for 2014 and each subsequent year by fitting a lognormal distribution to the data, with trends in the proportion of the population in each consumption category for each age and sex group estimated through to 2034. In PRIMEtime CE, age and sex specific mean salt consumption was estimated using NDNS dietary data from 2008 to 2012[218] across six age groups (0-3 years, 4-12 years, 13-17 years, 18-34 years, 35-54 years, and over 55 years; only those aged 18 years and older were included in the comparison with the UK Health Forum) and the mean was then scaled to match urinary sodium excretion with uncertainty estimated using standard errors from NDNS data. The UK Health Forum did not report their baseline salt consumption by age and sex, only the average for each gender, so it is difficult to know the direction in which these differences will affect results. For the purposes of the comparison, the change in salt consumption over time was modelled in the same way as the UK Health Forum.

The other important remaining difference between the two models was how the change in blood pressure was estimated following a change in salt consumption. Both models refer to He et al.,[223] with PRIMEtime CE using He et al.'s estimate that systolic blood pressures falls by an average 5.8mmHg following a 100mmol reduction in 24 hour urinary sodium (adjusted for age and baseline blood pressure, equivalent to 5.9g salt consumption). The UK Health Forum, by contrast, used results from a subgroup analysis suggesting that blood pressure among hypertensive adults falls by 5.4mmHg following a 75mmol reduction in 24 hour urinary sodium (7.2mmHg for a 100mmol reduction) and by 2.4mmHg (3.2mmHg) among normotensive adults. Furthermore, based on data from both He et al.[223] and Cappuccio et al.[264] suggesting that older adults have a greater fall in blood pressure for the

same reduction in 24 hour urinary sodium excretion, the UK Health Forum model assumed that the effect size for hypertensive adults from He et al. applied to anybody over the age of 65 years. As those over 65 years have a greater burden of CVD, these differences mean that the UK Health Forum model is likely to estimate larger overall effect sizes than PRIMEtime CE.

The results of the UK Health Forum model alongside the modelling of equivalent scenarios in PRIMEtime CE are shown in table 6.2 and figure 6.2. The UK Health Forum estimated larger increases in stroke and IHD cumulative incidence than PRIMEtime CE when comparing the following trends scenario (salt consumption falling from 8.9g to 8.0g among men and 6.8g to 6.1g among women) with the worst case scenario (no change in salt consumption). This is likely due to the differences outlined above in how the relationship between salt consumption and blood pressure is parameterised. By contrast, changes in disease incidence measured by the UK Health Forum model and PRIMEtime CE were similar for the comparison of the best case scenario (salt consumption falls to 5.0g among both men and women) with the following trends scenario. In PRIMEtime CE, the 3.0g and 1.1g reductions in salt consumption among men and women respectively produced changes to cumulative incidence over the 20 years of the intervention that were much greater in magnitude than those estimated by the 0.9g and 0.7g increases in salt consumption simulated in the worst case scenario versus follows trends scenario. However, the UK Health Forum estimated that the reduction in disease cumulative incidence arising following the best case versus follows trends scenario was similar in magnitude to the increase estimated in the worst case versus follows trends scenario. This may also be due to differences in how the relationship between salt consumption and blood pressure is parameterised. As salt consumption declines to increasingly low levels, reductions in blood pressure would mean that fewer individuals in the UK Health Forum model are hypertensive and therefore those under 65 years old would have a smaller reduction in their risk of developing CVD compared to the population simulated in PRIMEtime CE where the risk reduction is the same irrespective of hypertension status. This means that at lower

levels of salt consumption, PRIMEtime CE may be less likely than at higher levels of salt consumption to underestimate the health benefit when compared with the UK Health Forum model.

The UK Health Forum model also predicted larger changes to IHD incidence than stroke incidence for both the worst case versus follows trends and best case versus follows trends scenarios, with the pattern being the reverse with PRIMEtime CE. This reflects the differences in baseline disease incidence and prevalence rates between the two models.

Compared with the UK Health Forum model, PRIMEtime CE estimated the change to NHS costs to be proportionately larger for the number of cases averted. Healthcare costs were estimated in both models based on the annual prevalence of disease and therefore cases of MI and stroke that are averted soon after the intervention accrue more years' worth of cost savings than those that are averted later. In PRIMEtime CE, the fall in salt consumption occurred evenly over the 20 years of the intervention with the same relationship between salt consumption and blood pressure irrespective of age and hypertension status. In the UK Health Forum model, the effect of reducing salt consumption on blood pressure was smaller than in PRIMEtime CE for normotensive individuals aged under 65 years, but higher for those aged over 65 years or with hypertension. As more of the 20 years were simulated, an increasing proportion of the modelled population would have either turned 65 years old or developed hypertension. Therefore this would accelerate the decline in incidence rate towards the end of the 20 years with fewer years in which to accrue savings.

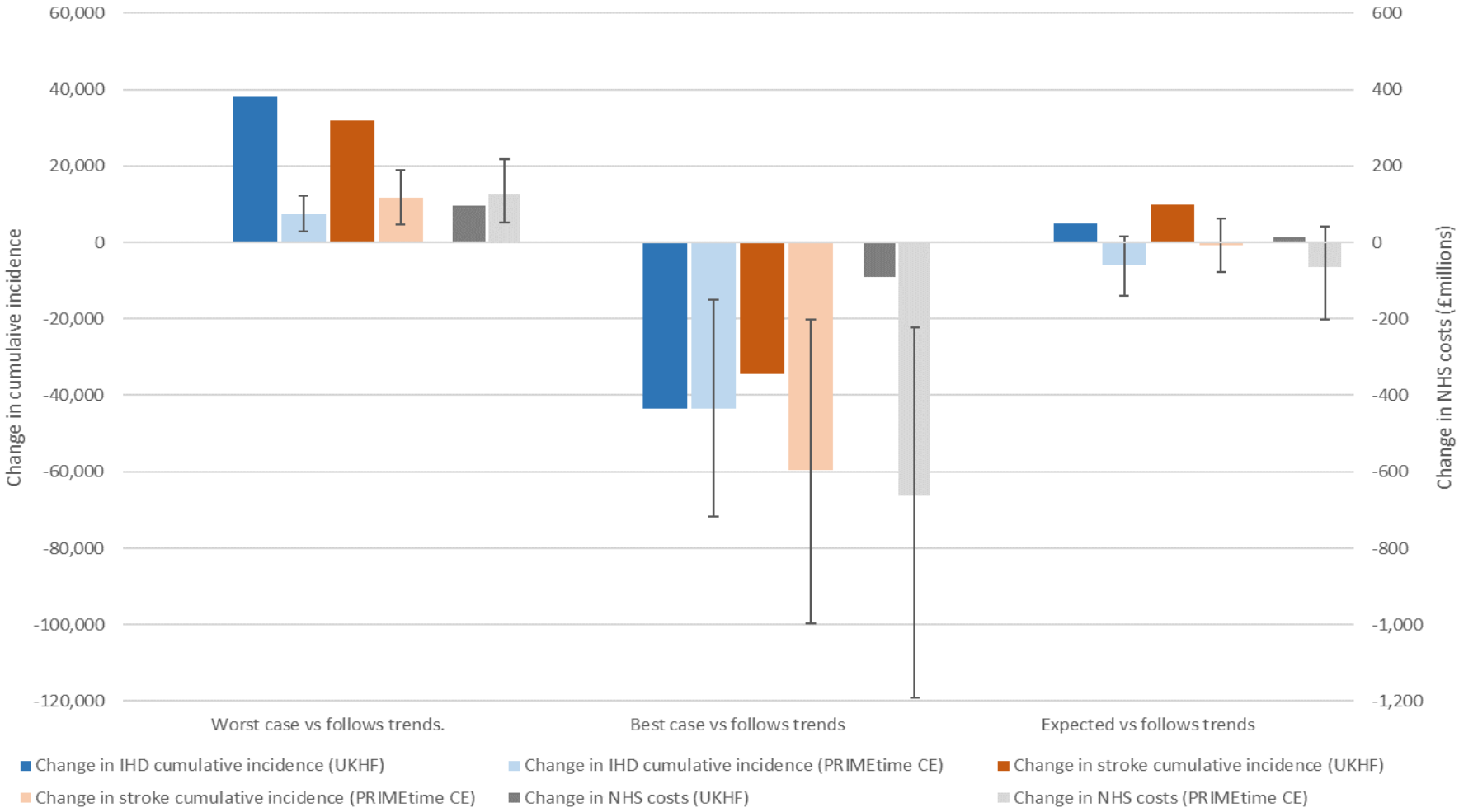
Finally, PRIMEtime CE results suggest that the 'expected' scenario would lead to small health gains when compared to following trends (although uncertainty intervals do overlap with zero), this is in contrast to the UK Health Forum model, which predicted small health losses. This is because the health gained by men moving from 8g salt consumption a day to 7g outweighs the health loss from women moving from 6.1g to 7 g in PRIMEtime CE, whereas the balance of outcomes by sex is the other way around in the UK Health Forum model.

Table 6.2 Results from the UK Health Forum microsimulation model and PRIMEtime CE modelling the UK Health Forum salt reduction scenarios

Scenario	Model	Change in NHS costs (£million)	Change in IHD cumulative incidence	Change in stroke cumulative incidence	Change in IHD incidence rate/100,000	Change in stroke incidence rate/100,000
Worst case vs follows trends	UK Health Forum	95	37,996	31,868	62	52
	PRIMEtime CE	128 (52 to 217)	7,553 (2,955 to 12,219)	11,778 (4,687 to 18,910)	14 (5 to 22)	22 (9 to 35)
Best case vs follows trends	UK Health Forum	-90	-43,512	-34,319	-71	-56
	PRIMEtime CE	-663 (-1,191 to -223)	-43,529 (-71,825 to -15,075)	-59,510 (-99,747 to -20,115)	-80 (-132 to -27)	-109 (-184 to -37)
Expected vs follows trends	UK Health Forum	14	4,903	9,805	8	16
	PRIMEtime CE	-65 (-201 to 42)	-5,873 (-14,095 to 1,581)	-689 (-7,883 to 6,188)	-11 (-26 to 3)	-1 (-14 to 11)

Median values presented for PRIMEtime CE results with uncertainty intervals in parentheses; IHD, ischaemic heart disease

Figure 6.2 Results of the UK Health Forum microsimulation model and PRIMEtime CE modelling the UK Health Forum salt reduction scenarios



Median values presented for PRIMEtime CE results; IHD, ischaemic heart disease; UKHF, UK Health Forum

6.4.2.2 Cross validation testing using the IMPACT CHD model

Collins et al. estimated the effect on health and NHS costs of reformulating food to contain less salt using the IMPACT CHD model.[219] The IMPACT CHD model is a population level model that estimates CHD mortality rates over time based on historical trends and the attributable disease burden of risk factors such as salt consumption. The authors estimated the effect on NHS costs and life years gained in England between 2010 and 2020 of 20%, 15%, and 2% reductions in salt consumption. Consumption was assumed to fall at the beginning of the 10 year period with health gains quantified for seven separate patient groups – acute MI admissions, unstable angina admissions, secondary prevention after acute MI, secondary prevention after revascularisation, angina in the community, heart failure admissions, and heart failure in the community. To enable comparisons of results, PRIMEtime CE was adapted to simulate the same reductions in salt consumption estimated by Collins et al., stroke was removed from the model so that only the effect on IHD was calculated, no unrelated disease costs were included, and the same 3.5% discount rate was used.

There remained some important differences between the two models, such as each model including different diseases. PRIMEtime CE modelled IHD, whereas IMPACT CHD simulated the effects of the intervention on a wider range of coronary heart diseases. Furthermore, the data sources for baseline disease incidence and prevalence differed between the two groups, with IMPACT CHD using pre-2011 data from HES,[163] the Myocardial Ischaemia National Audit Project,[265] and the GPRD,[147] compared to the data sources used by PRIMEtime CE (appendix four). Collins et al. did not publish their baseline disease incidence and prevalence rates making it unclear how this would affect results. Both models used He et al. to estimate the effect of a change to salt consumption on blood pressure.[223] However, the IMPACT CHD model used the separate estimates for adults with and without hypertension to quantify an average effect size for the population in each of the seven patient groups. Again, the effect of this difference on results is not clear because the final effect sizes used by Collins et al. are not reported.

A further difference was that although both models used Lewington et al. to estimate how a change in blood pressure affects vascular disease,[224] this was only used for mortality estimates by IMPACT CHD to quantify life years gained, compared to incidence in PRIMEtime CE. In order to estimate the number of life years gained in IMPACT CHD, the number of deaths delayed or postponed following the 10 years of the intervention was quantified based on the change in blood pressure and multiplied by the age specific median-survival for the those with diagnosed CHD, undiagnosed CHD, and without CHD. This therefore assumes that the same survival benefit occurs, irrespective of from when the death is postponed following the intervention. In PRIMEtime CE, the life years gained was estimated based on the difference in 10 year life expectancy for each age group calculated as the difference in the sum of person-years lived pre- and post-intervention. This method does not use an estimate of median-life expectancy and is likely to underestimate the number of life years gained compared to IMPACT CHD because rather than assuming all deaths are delayed from the beginning of the intervention, deaths are delayed across the ten years.

Finally, there are significant differences in cost estimates. In contrast to PRIMEtime CE, IMPACT CHD used bottom-up estimates from reference costs based on expected treatments, medications, and professional health visits for each patient group simulated. For example, a patient with heart failure in the community had annual costs based on using four different medications, and having six GP and six practice nurse appointments a year. Costs are saved as a result of fewer individuals developing disease, based on the attributable fraction of CHD cases for the change in blood pressure using odds ratios from the INTERHEART study.[266] Neither the number of cases averted nor the unit costs for each patient group are reported by Collins et al. making it difficult to understand how these differences impact on the final NHS cost savings.

Results of the interventions modelled by Collins et al. using IMPACT CHD and PRIMEtime CE are shown in table 6.3.

Table 6.3 Results from IMPACT CHD and PRIMEtime CE modelling the Collins et al. salt reduction scenarios

Scenario	Model	Life years gained (95% UI)	NHS cost savings (£)
Salt reduction of 20%	IMPACT CHD	19,365 (11,967 to 27,887)	686,572,778
	PRIMEtime CE	10,437*	187,694,772
Salt reduction of 15%	IMPACT CHD	14,593 (9,000 to 21,049)	601,843,338
	PRIMEtime CE	7,880*	141,798,746
Salt reduction of 2%	IMPACT CHD	1,970 (1,209 to 2,854)	434,248,212
	PRIMEtime CE	1,069*	19,277,616

*Uncertainty intervals not estimated due to computational time required to simulate each age group separately to calculate life years gained; UI, uncertainty interval

The estimated number of life years gained are smaller using PRIMEtime CE than IMPACT CHD, however they are in the same order of magnitude. It is not possible to know how much of the variation is due to differences in the diseases simulated and the calculation of life years gained, and how much is due to other aspects of each model's structure. The changes to cost estimates are markedly different with NHS savings using PRIMEtime CE being roughly proportional to the number of lives saved. By contrast, significantly higher cost savings are estimated using IMPACT CHD, with only a 37% difference in savings for a 2% reduction in salt consumption compared to a 20% reduction, despite a 90% difference in life years gained. This implies that the IMPACT CHD cost savings are more aligned with changes in morbidity than mortality, about which there is very little detail in the paper and its supplementary materials.

6.4.2.3 Cross validation testing and validation against outcomes using alternative input data using Frew et al.

PRIMEtime CE was compared with Frew et al.'s Markov model, which estimated the cost-effectiveness of the Be Active intervention.[144] The input parameters quantifying the effect of Be Active on physical activity used by PRIMEtime CE in section 4.2 were taken from Frew

et al. and when the cost-effectiveness of scaling up the intervention to England was estimated (the physical activity intervention primary analysis), the cost per QALY was £727,000 (95% UI £514,000 to £1,064,000), compared to £400 per QALY estimated by Frew et al. However, the PRIMEtime CE results are calculated using a different time horizon, cost perspective, and discount rate than Frew et al. When using 3.5% discounting, a five year time horizon, and NHS costs only with no unrelated disease costs, as used by Frew et al., the cost per QALY was even higher at £1,711,000 (£1,252,000 to £2,400,000).

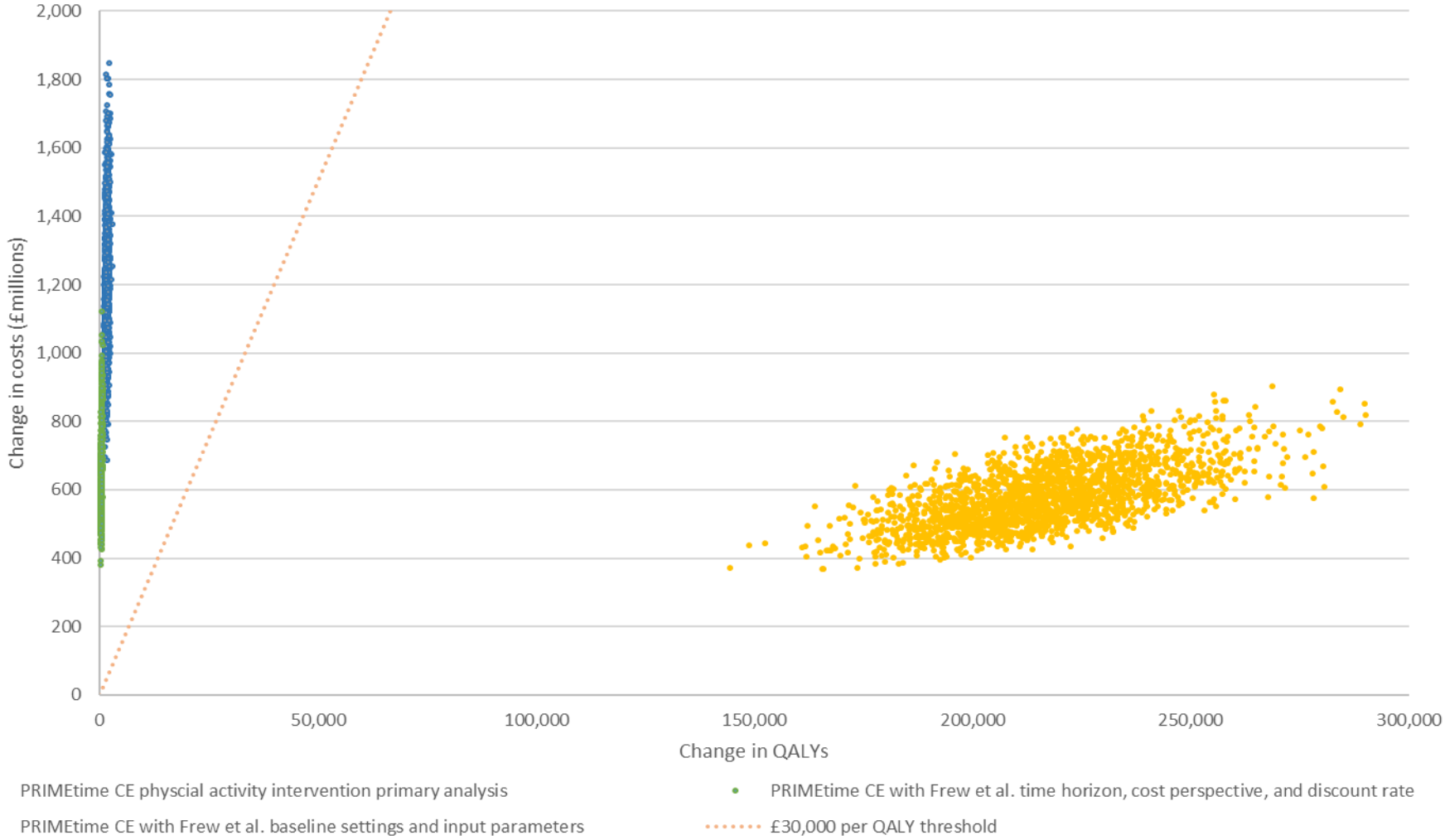
However, the two models had other important differences between their input parameters, in particular the disease specific utility and cost estimates. As discussed in chapter three, Frew et al. used a variety of sources for their utility decrements and costs, with both being generally larger than those used in PRIMEtime CE. For example, the utility for stroke used by Frew et al. was 0.612 (from a baseline of 0.805, 0.875, or 0.912 depending on the baseline state of physical activity) compared with a decrement of -0.112 (incident case) and -0.073 (prevalent case) in PRIMEtime CE. Furthermore, the cost per incident and prevalent stroke case was £9,630 and £2,396 respectively in Frew et al.'s model, compared with £843 for both incident and prevalent cases in PRIMEtime CE.

The other important difference between the two models was how they quantified the risk of disease following a change in physical activity levels. Frew et al. conducted their own systematic reviews and meta-analyses focusing on LTPA (see section 4.2.3.2). By contrast, PRIMEtime CE used results of two meta-analyses providing a dose response relationship of total physical activity and disease outcome,[242,243] and used these to estimate the age and sex specific change in relative risk of disease based on physical activity levels pre- and post-intervention. Table 4.5 compares the relative risks used by both studies, with those used by Frew et al. for people moving between under active to recommended levels of physical activity being significantly smaller than those used by PRIMEtime CE. Furthermore, the annual disease risks used by Frew at al. are significantly greater than the baseline disease incidence rates in PRIMEtime CE. For example, the annual risk of breast cancer used by

Frew et al. was 0.021, 0.018, and 0.017 depending on whether an individual was under active, active, or achieving recommend physical activity levels. By contrast, the baseline annual incidence rate of breast cancer among those aged 16 years and over used in PRIMEtime CE was 0.0008. Therefore Frew et al.'s model is likely to lead to both greater relative and absolute health gains than PRIMEtime CE.

Using costs and utilities used by Frew et al. in PRIMEtime CE, as well as 2010 intervention costs, a five year time horizon, NHS perspective, and 3.5% discount rate, the cost per QALY was £2,662 (£2,176 to £3,327), much closer to Frew et al.'s estimate of £400 per QALY. Figure 6.3 shows the results of the PRIMEtime CE comparisons with Frew et al. on the cost effectiveness plane. The residual difference in cost per QALY between the two models was due to costs: the incremental QALY estimate was 0.06 per person in Frew et al. and 0.06 (0.05 to 0.07) in PRIMEtime CE, but the incremental cost was £24 in Frew et al. compared to £153 (£116 to £206) in PRIMEtime CE. This cost difference may in part be due to how disease relative risk estimates vary between the two models, as differences in individual disease outcomes may have led to a similar overall effect on QALYs but not on costs, and in part be due to different underlying model structures.

Figure 6.3 Results of PRIMEtime CE comparisons with Frew et al.



QALY, quality adjusted life year

6.4.2.4 Validation against outcomes using alternative input data using disability weights

Estimates of the change in QALYs following the primary analyses were compared with disease disability weights taken from the Global Burden of Disease study to calculate health adjusted life years (HALYs).[267] Like QALYs, in PRIMEtime CE, HALYs are calculated by weighting the total number of years of life lived by the population by the level of ill-health after the intervention compared with the baseline. For QALYs, disease weighting uses utilities, where perfect health has a score of one and death has a score of zero, with PRIMEtime CE utilities measured with the EQ-5D questionnaire and valued using TTO (see section 3.2.2). Being in ill-health results in a utility decrement and a reduction in the average population utility. By contrast, disability weights value perfect health as zero, with ill health leading to additional disability on scale of zero to one, and are measured based on paired comparison questions where study participants are asked to choose which of two hypothetical individuals with different disease states they think is healthier.[268] When estimating HALYs in PRIMEtime CE, the Global Burden of Disease study disease disability weights were subtracted from one and multiplied by the total number of years lived by the population with and without the intervention. This provides a comparison with the utility decrements used for estimating QALYs. Although measured very differently, comparisons between QALYs and HALYs are shown in table 6.4 and are broadly similar, providing additional evidence that the QALY estimates are valid.

Table 6.4 Comparison of change in QALYs with a change in HALYs for the primary analyses of the diet and physical activity interventions

Scenario	Change in QALYs	Change in HALYs
Diet intervention primary analysis	15,000 (5,928 to 25,520)	19,531 (8,520 to 30,764)
Physical activity primary analysis	1,597 (1,134 to 2,169)	1,785 (1,308 to 2,375)

95% uncertainty intervals in parentheses. QALY, quality adjusted life year; HALY, health adjusted life year

6.4.3 Validation against empirical data

Validation against empirical data can take two forms, dependent (or internal) validation against data used to parameterise model, and independent (or external) validation against data not used by the model.

6.4.3.1 Dependent validation using empirical data

This form of validation checks that model inputs have been appropriately combined to give expected results. One of the challenges of dependent validation for models simulating public health behavioural interventions is the difficulty of isolating the disease burden attributable to a single behavioural risk factor from the population level dataset used to parameterise the model. Studies that attempt to quantify this attributable disease burden, for example how trends in salt consumption have affected CVD outcomes, generally use modelling themselves.[269,270] Another approach is to compare model results with future disease trends using the same datasets that informed the model parameters. However, not only does this method encounter the same problem of quantifying the independent effect of a change in a behavioural risk factor on disease outcomes, but disease outcomes can take many years to manifest with the additional difficulties of separating out other secular epidemiological trends. Therefore, dependent validation of PRIMEtime CE has not been attempted.

6.4.3.2 Independent validation using empirical data

Independent validation can be done of the entire model or of different parts of the model depending on data availability: how an intervention affects a behavioural risk factor could be validated separately to how the change in risk factor subsequently affects health. Independent validation of the estimate of how an intervention affects a change in the risk factor relies on data from a similar intervention in a different setting, either following a trial or natural experiment. Such data are not available for the two interventions modelled in this thesis, with trials of population level public health interventions such as salt reformulation being difficult to design. In the absence of trial data, independently

validating how the modelled change in risk factor affects health using natural experiments has the same difficulties as dependent validation, discussed in section 6.4.3.1.

Newer large cohort studies such as the UK Biobank are providing datasets with enough information on both predictor variables and outcomes to potentially enable independent validation of public health models affecting behavioural risk factors.[271] For example, a crude check of independent validity could be made if within the cohort of 500,000 adults it is possible to identify two subpopulations that match the modelled salt consumption pre- and post-intervention. Health outcomes for each subpopulation could then be adjusted for measured known confounders and the independent effect of salt on stroke and IHD estimated. However, two such populations may have been consuming their reported levels of salt all their life, and as such they represent a very different population to one whose dietary habits change following an intervention. A more appropriate independent validation would require identifying one subpopulation that consumes a steady level of sodium throughout the cohort's follow up, and another subpopulation whose salt consumption falls at some point post recruitment. The effect of this change in salt consumption could then be modelled with resulting change in health outcomes compared with the cohort's subpopulations. Finding a large enough subpopulation of individuals each with the same fall in salt consumption would likely be difficult. Furthermore, such large and comprehensive data sources remain rare and are still relatively young, with the UK Biobank finishing recruiting in 2010.

Independent validation of PRIMEtime CE has not yet been attempted. The coding and linking of hospital admissions for those in the UK Biobank has recently been completed and therefore could be a potential option for independently validating PRIMEtime CE in the future.

6.5 Part E: Other validation techniques

Part E suggests other types of validation that may also be used, such as guiding people through the conceptual model or computer model step by step (structured walk-throughs), or using naïve benchmarking against what might be expected from rough calculations. Health economists from the Department of Health and PHE have been guided through the conceptual model, and both DPhil supervisors and Dr Cobiac from the BODE^3 research group have been guided through the computer model. Other validation techniques have not been performed.

6.6 Discussion of model validation

This chapter illustrates the challenges of validating PRIMEtime CE and public health economic models more generally. The face validity of PRIMEtime CE's conceptual model and its model inputs appears to be good, and the model also produces expected results when analyses are run using extreme input variables. However, cross-validation of outcomes showed that once input settings have been changed to match those used by comparison studies, PRIMEtime CE estimated similar disease outcomes to Frew et al.[144] and slightly smaller effect sizes than Collins et al.,[219] but costs were significantly lower than those estimated by Collins et al. and higher than those estimated by Frew et al. When compared with the UK Health Forum model,[115] differences in disease outcomes varied depending on the scenario simulated, but costs were much larger relative to the change in cumulative disease incidence. Some of the differences between PRIMEtime CE and the model to which it was compared can be explained by differences in input parameters and the model's underlying assumptions, however unexplained differences remained and are likely to relate to other elements of the model's structure (structural uncertainty). The cross validation testing provides some evidence that PRIMEtime CE does not systematically over or under estimate cost and disease outcomes compared with other models, but instead the magnitude and direction of the differences varies with each comparison.

A further important result from the cross validation testing of model outcomes is the impact on model outcomes of the choice of input parameters. The PRIMEtime CE cost per QALY estimate of the physical activity intervention was £1.7m when using PRIMEtime CE disease cost and utility estimates and £2,700 when using those from Frew et al. (both estimated with a five year time horizon and an NHS perspective). This stark variation between results highlights the importance of being transparent about the choice of model parameters and presenting sensitivity analyses to illustrate how outcomes may differ under alternative assumptions.

The challenges of cross validation and structural uncertainty are well known with the Mount Hood challenge in diabetes modelling, the COPD modelling meeting, CISNET cancer model comparisons, and the Agricultural Model Intercomparison and Improvement Project all set up to try to understand how model structural differences affect results.[25–27,272,273] These academic collaborations provide important settings in which to appraise model validity, address methodological difficulties, and subsequently improve model quality. No such collaboration exists for public health economic models or non-communicable disease models more broadly, although future work is planned in this area with the BODE^3 research group in New Zealand[260] and is expected to provide useful insights and to facilitate model improvements.

Finally, independent validation of PRIMEtime CE is a priority for future work. This would provide key information about which aspects of the model's structure do and do not adequately predict outcomes in real world settings, and enable appropriate improvements.

Chapter 7. Discussion

The aim of this thesis was to develop a model to compare the cost-effectiveness of primary prevention policies affecting diet and physical activity in England. The intention is that the model, PRIMEtime CE, will be used to enable better local and national public health decision making and to improve population health. Chapters one and two outline the background to the problem and explain the development and structure of PRIMEtime CE; chapters three and four describe how model inputs were identified and used; and chapters five and six present the results and model validation. Each chapter has a discussion section, where the major strengths and weaknesses of the methods described in that chapter are outlined. This final chapter reiterates the key results and conclusions of the thesis, highlights PRIMEtime CE's principal strengths and limitations, and discusses the work in the context of informing public health policy.

7.1 Main findings

When directly comparing the diet intervention with the physical activity intervention, PRIMEtime CE estimates that reformulating food to have less salt would be considerably more cost-effective than replicating the Be Active scheme across England (salt reformulation is dominant compared to £727,000 [£514,000 to £1,064,000] per QALY for expanding Be Active, see chapter five).

A key aspect of PRIMEtime CE is that it adopts a standardised and replicable approach to identifying costs and utilities across multiple diseases related to diet and physical activity. Furthermore, as public health interventions can prevent disease and prolong life, PRIMEtime CE quantifies the health and social care costs, as well as any reduction in quality of life, arising as a consequence of developing diseases unrelated to those modelled. This makes it possible for decision makers to be reasonably confident (within the

uncertainty intervals presented) that one intervention is likely to be more cost effective than another given the time horizon and economic perspective used.

However, the absolute cost per QALY is heavily dependent on decisions made in three aspects of the model: the choice of model scope and baseline settings, the choice of model parameters, and the choice of model structure.

7.1.1 Choice of model scope and baseline settings

The model's scope and baseline settings are decided upon during the model's development. They include the model's time horizon, discount rates, and economic perspective. Decisions about the scope and baseline settings can have significant implications for the final cost per QALY estimated, and as such they are usually scrutinised with univariate sensitivity analyses. For example, sensitivity analyses with different time horizons resulted in the cost per QALY for the dietary intervention being as high as £2.1m (£0.9m to £6.8m) after one year, compared to being dominant after just nine years, and having a return on investment of £17 (£6 to £33) for every £1 spent over the lifetime of the cohort. Similarly, implementing Be Active across England had a cost per QALY of £12.2m (£8.2m to £18.5m) after one year compared to £92,000 (£62,000 to £137,000) over the cohort's lifetime. The importance of the economic perspective is highlighted by the diet intervention sensitivity analysis where only government costs, and not industry costs, were included. Under this scenario the intervention's cost per QALY was -£56,000 (-£80,000 to -£42,000) compared to -£17,000 (-£40,000 to £39,000) in the primary analysis.

Importantly for informing policy, these results show how the cost-effectiveness of an intervention can vary depending on the time horizon and choice of economic perspective. At a national level, the five year fixed term parliament act means that it may be politically important for cost-effectiveness estimates of different policies to use a five year time horizon, and at local authorities, the political timeframe is even shorter with councillors being re-elected every four years. Although a 10 year time horizon was the preferred time horizon for the majority of stakeholders based on the ranking of their top three choices, the

same number of respondents ranked five and 10 years as their first choice (see appendix three). Using a five year time horizon for the primary analyses would have led to significantly higher cost per QALY estimates for each intervention.

7.1.2 Choice of model parameters

The absolute cost per QALY also depends on the choice of model input parameters. Parameters used in PRIMEtime CE were chosen based on the aim of developing a cost-effectiveness model the can compare different interventions affecting multiple diseases. Although the uncertainty associated with each parameter used in PRIMEtime CE was quantified and represented by uncertainty intervals, the median cost per QALY and its uncertainty intervals would change were different parameters chosen during the model's development. For example, comparing the cost per QALY of the physical activity intervention estimated using PRIMEtime CE with Frew et al.'s model, the two model's results differed markedly despite having the same scope and baseline settings.[144] PRIMEtime CE estimated the cost per QALY to be £1,711,000 (£1,252,000 to £2,400,000) whereas Frew et al. estimated it to be just £400. However, changing the intervention costs, disease costs, and utility input parameters in PRIMEtime CE to match those included in Frew et al. reduced the cost per QALY estimate to £2,700 (£2,200 to £3,300, see section 6.2.3). This result is over 600 times smaller than the previous estimate and emphasises how the choice of model parameters can dramatically affect results, and therefore subsequent policy decisions, despite parametric uncertainty being quantified and reported.

7.1.3 Choice of model structure

Finally, the model structure has an impact on the model's final results. The model structure determines what input data are required and the relationships between the interventions, risk factors, and outcomes modelled. For example, as a multistate life table model, PRIMEtime CE includes age and sex specific data on baseline disease incidence, prevalence, case fatality, and trends, none of which are required by the Markov model used by Frew et al.[144] Both models also relate their input data to outcomes differently, for example, the

prevalence of type two diabetes affects the risk of developing IHD and stroke in PRIMEtime CE, but not in Frew et al.'s model. These structural differences go some way to explain residual variations in model outcomes when model scope, baseline settings, and parameters have been standardised.

7.1.4 Implications of model choices for decision makers

The problem underlying the decisions regarding model scope, parameters, and structure is that the true cost per QALY is not known, there is no gold standard against which results should be compared. Independent validation has a key role in improving models and understanding which ones best represent reality. However as discussed in chapter six, this can be very challenging for public health interventions – if the true costs and consequences of an intervention were known then there would be no need for the model.

The problem of the true cost per QALY being unknown can place decision makers in a difficult situation – how much should models be relied upon, and do decision makers have the time and expertise to understand the methodological challenges described above? Following his tenure as Chief Scientific Adviser at the UK Department for International Development (DFID), Whitty wrote an editorial titled 'What makes an academic paper useful for health policy?'[125] In it, Whitty describes the common features of different pieces of research that successfully influence policy. These commonalities include attempting to minimise bias, being explicit about a model's methodology and limitations, and being timely. When discussing modelling, Whitty highlights how commonly models are used by government but *"the key is usually not that they are 'right', but that they provide an unpredicted insight."* Allowing decision makers to change the starting assumptions and providing sensitivity analyses of key variables are particularly important, as is keeping the model simple. Although the point is made that economic models are generally better understood by decision makers than other model types, the underlying principles remain. Another source of advice is guidance produced the ISPOR-SMDM Modeling Good Research Practices Task Force.[12] This was set up in response to a need to increase the credibility of

models used to inform decisions about health technologies and like Whitty, this series of papers highlights the importance of model uncertainty, validation, and transparency. Therefore, comparing PRIMEtime CE with advice from Whitty and the ISPOR-SMDM guidance may help to highlight how PRIMEtime CE can be improved to maximise its policy impact.

7.1.5 Policy strengths of PRIMEtime CE

When assessed against the points made by Whitty and ISPOR-SMDM guidelines about how models can best inform policy, PRIMEtime CE has several strengths. Bias is minimised by adopting a consistent and systematic approach to identifying model input parameters, in particular disease costs and utilities; the model's methodology and its limitations are explained in detail in this thesis; model uncertainty is quantified and reported; and multiple sensitivity analyses are reported. Furthermore, PRIMEtime CE has the option to include changes to costs and quality of life from diseases unrelated to those modelled, thereby more accurately representing the cost per QALY of interventions that prolong life.

The development of PRIMEtime CE followed Squires et al.'s conceptual modelling framework which included engaging relevant stakeholders early in the model's development to ensure the model suited stakeholders' needs.[46] In terms of timeliness, one of the benefits of multistate life table models is that additional disease outcomes can be added without significantly increasing complexity. This means that PRIMEtime CE can be updated to respond to decision makers' needs and newly available evidence. As well as disseminating the model, more work needs to be done both with validation (see section 6.6), and to make the model more user friendly to enable decision makers to change many of the starting assumptions. This is discussed under section 7.3.

Finally, public health policies in England are often assessed against their ability to save money and provide a return on investment, by comparison healthcare interventions are usually assessed on their cost effectiveness.[17] Prof Brian Ferguson, Chief Economist at Public Health England, has written about this in a series of blog posts stressing the need to

"measure prevention and treatment by equal-standards" and value the long-term benefits of preventive interventions.[274] Being able to directly compare interventions using PRIMEtime CE may be a powerful way to emphasise the benefits of public health interventions alongside common healthcare interventions. This is also discussed further in section 7.3.

7.1.6 Policy limitations of PRIMEtime CE

There are various methodological limitations of both the data sources used by PRIMEtime CE and the model's structure. These are discussed in more detail in the relevant chapters of the thesis, however it is important to reiterate some of the major ones here, and their possible impact on the model's results and its subsequent use for informing public health policy.

Firstly, as discussed in chapters one and two, multistate life table models assume diseases are independent of each other. This means that the proportion of the cohort existing in one disease state does not affect the probability of the cohort developing any other disease states. To counter this limitation, PRIMEtime CE has been developed such that there is a dynamic relationship between type two diabetes and CVD (an increase in diabetes prevalence increases the risk of the modelled cohort developing IHD and stroke). However, this does not occur for other diseases which, depending on the intervention simulated, may over- or under-estimate the resulting cost per QALY. Secondly, no interactions between individuals within the population or between the population and its environment are simulated. Adding these interactions requires data on their effects and would generate more uncertainty in model results, but it may also mean the model better represents reality and is more accurate. An example is that if a physical activity intervention affects attitudes towards being more active, there may be a positive feedback loop between the numbers active and uptake of the intervention, which might subsequently decrease the cost per QALY. However, as discussed in chapter one, if multiple interactions are important to include, a different model structure may be more appropriate. Thirdly, PRIMEtime CE only

models the effect of changing a risk factor on disease incidence and not case fatality rates. If simulated interventions led to reductions in case fatality rates, the model may underestimate health and cost outcomes arising from increased longevity (both for modelled and unrelated diseases). Finally, the impact of policies on health inequalities are important to public health decision makers (section 1.1) and although it is possible to estimate the impact of different policies on different population subgroups using cohort models, microsimulation models are better suited to this.

Further limitations are introduced through using NHS England programme budgeting data to estimate disease costs. Although using these data mean that results are comparable with one another and that 78% of the NHS England budget is accounted for, £21bn remains unaccounted. This is for two main reasons, firstly it was not possible to allocate expenditure in category *23, other* (20% of 2013/14 programme budgeting costs) to specific diseases because often the condition related to the care provided was not known. Secondly, over £7bn of expenditure in 2014 by NHS England was not included as it was not directly allocated through programme budgeting, primary care, or specialised services. Examples of how this money was spent include CCG running costs, military and offender care, and NHS England central health programmes such as clinical excellence awards. If some modelled diseases are miscoded to category 23, then PRIMEtime CE would underestimate savings from these diseases. Conversely if some NHS England administrative costs (which in reality would be unlikely to change following small changes to disease prevalence rates) are included in programme budgeting categories related to specific diseases, then these will have contributed to disease costs in PRIMEtime CE and savings following an intervention would be overestimated.

7.2 Comparisons with other research

Previous reviews have found that public health interventions are often cost effective or cost saving.[2,3,10,11] Owen et al. reviewed 200 cost effectiveness estimates from NICE public health

guidance and found that 89% had a cost per QALY of less than £30,000.[3] Most analyses reviewed by Owen et al. used a lifetime horizon and the cost per QALY estimates using a lifetime horizon of the two interventions investigated in this thesis would be consistent with these analyses. PRIMEtime CE estimated that salt reformulation would be cost saving over the life time of the cohort and that expanding Be Active would have a cost per QALY of £92,000 (£62,000 to £137,000), not dissimilar to free swimming for children and young people intervention included in the Owen et al. review which had a cost per QALY of £40,462. A review of economic evaluations of preventive health interventions published in 2008 by van Gils et al. found 232 relevant studies, of which 80% were cost-effective at a threshold of €50,000 (£39,500 in 2008) per QALY, and 60% at €20,000 (£15,800) per QALY.[10] The authors concluded that the majority of studies show favourable cost-effectiveness ratios, however the cost per QALY of specific interventions is not reported meaning that results cannot be directly compared with the interventions evaluated in this thesis. Masters et al. reviewed studies estimating the return on investment or cost benefit ratio of different public health interventions and found that the median return on investment among the 52 studies included was 14.3 to 1 (PRIMEtime CE estimated that the lifetime return on investment of the salt intervention was £17 for every £1 spent). They concluded that local and national public health policies are likely to be cost-saving, however the extent to which included papers are representative of broader research on the cost effectiveness of public health interventions is unclear.[11] Finally, the WHO Regional Office for Europe also reviewed the evidence on the cost-effectiveness of public health interventions finding that many are cost-effective with several (such as vaccination programmes and violence prevention) showing a return on investment within 1-2 years.[2] This review cited the WHO report on the "best buys" for the most cost-effective approaches to reducing the burden of NCDs in low and middle-income countries.[275] The report suggests that in low- and middle-income countries, reducing salt consumption through reformulating unhealthy food and mass media campaigns would be very cost effective (the cost of an additional year of healthy life would be less than the national per person annual

gross domestic product), as would promoting physical activity through mass media campaigns.

The intended use of PRIMEtime CE is not to add to the evidence of such reviews, but to allow comparable cost-effectiveness estimates to be made of different public health policies including ranking interventions to help decision makers prioritise resources. Similar rankings have been done in the UK and elsewhere; for example, the NICE physical activity return on investment tool,[87,276] the ACE Prevention programme of research in Australia,[277] and the BODE³ Programme in New Zealand.[260] The NICE physical activity return on investment tool compares how different physical activity interventions would affect health and costs for the user's geographic region and population demographic. The tool is based on a Markov model that estimates how changes to physical activity affect the prevalence of CHD, stroke, and type two diabetes. The tool has a user-friendly interface whereby various model inputs can be manipulated and the outputs analysed. Unlike PRIMEtime CE, however, the disease costs and utilities used by the model are from a variety of different sources and population groups.

The research programmes in Australia and New Zealand both used systematic and comparable approaches to estimating the cost-effectiveness of different public health interventions, with the explicit aim of informing policy. As with PRIMEtime CE, these were developed using country specific routine data where possible with standardised approaches to estimating utilities and costs. The BODE³ programme of research derived age, sex, and disease specific health costs from Health Tracker, national data that individually links costs with health events;[96] New Zealand specific disability weights taken from the Global Burden of Disease study.[278] The ACE programme used the DISMOD tool to estimate disability weights based on Australian burden of disease data (see supplementary file one from Cobiac et al.[279]), and costs were taken from a national dataset with disease specific healthcare cost estimates.[280] The use of country specific data sources mean that although results from the ACE and BODE³ programmes of research may not be directly applicable to England, both

can rank interventions. Neither programme has investigated a scheme similar to Be Active, however both ACE and BODE³ simulated interventions aimed at reducing salt consumption, finding that as with PRIMEtime CE, reformulation would dominant over a lifetime horizon.[94,256]

7.3 Future work and dissemination

7.3.1 Developing and expanding PRIMEtime CE

The results of this thesis have been shared with health economists from the Department of Health and PHE. Following presenting the results at a meeting in June 2017, both groups gave feedback on how they would ideally like PRIMEtime CE to be developed to maximise its policy relevance (further stakeholder feedback is expected following the online questionnaire circulated as the final part of stakeholder engagement, see appendix one). Health economists from PHE recommended designing a user friendly interface whereby key model assumptions and simulated interventions could be changed by decision makers. Adding in additional risk factors to broaden the scope of interventions that can be compared was suggested by Department of Health economists as a potential priority for future work. Two key risk factors that are not currently included are alcohol misuse and tobacco use, and alongside adding these to PRIMEtime CE there may be benefits to including routinely prescribed medications used in primary and secondary prevention such as anti-hypertension drugs, statin-lowering drugs, and aspirin. Not only would the inclusion of medications broaden the policy appeal of PRIMEtime CE, they would help to address some of the challenges posed by Prof Ferguson.[274] Using PRIMEtime CE to rank the cost-effectiveness of interventions affecting the four behavioural risk factors alongside the cost-effectiveness of prescribing drugs for secondary prevention may illustrate how some public health policies affecting behavioural risk factors are more cost-effective than routinely used healthcare interventions.

Health economists at PHE were also interested in developing the tool for use at the local authority level using local demographic data. This is an important future development of PRIMEtime CE that would allow it to better inform local public health decision makers working in local authorities. Further possible developments of PRIMEtime CE would be to include estimates by SES (also suggested by PHE) and by ethnicity. This would require SES and ethnicity specific demographic and burden of disease data, plus updated parameters where appropriate to reflect any differences in how these characteristics affect the relationships between interventions, risk factors, and diseases. This update would allow the model to estimate how interventions impact inequalities by socioeconomic group and by ethnicity.[16,46] Finally, in the long term the model could be developed for use in different international settings, where data are available.

7.3.2 How to value outcomes

There is a wider debate surrounding how best to value outcomes from public health interventions, underpinned by extra-welfarism discussed in section 1.3.6. Using CUAs and reporting cost per QALY is useful to health policy makers in England because it is consistent with the NICE approach to assessing medical interventions and is therefore relatively easy to interpret and compare with other health economic assessments.[13] However, public health interventions may have societal and environmental costs and benefits beyond health, not quantified using a CUA.[14–17,19,46,281,282] Building on work by Drummond et al. and Weatherly et al.,[15,16] Marsh et al. discussed the merits of various alternatives to CUA, ranging from CBAs to adopting alternative measures of wellbeing.[14]

Marsh et al. initially propose using a CCA. This involves presenting a list of different possible outcomes, or consequences, of an intervention which may have an impact on different diseases, populations, or institutions, see Trueman and Anokye for an example.[31] Although this has in the past been recommended by NICE for public health interventions,[283] the broad range of outcomes presented makes it difficult to compare CCAs of different interventions. A CBA would be an alternative approach, again previously recommended by

NICE,[283] where all outcomes are assigned a monetary value; however for reasons discussed in section 1.3.5, CBAs are rare in evaluations of public health interventions. Marsh et al. discuss using either a capability approach as developed by Amartya Sen, or a subjective wellbeing approach to quantifying an intervention's outcomes. The capability approach is suggested by NICE as a way to measure and value the effects of non-health interventions.[13] The approach aims to capture how an intervention affects an individual's ability (capability) to function in a given way.[43,284] Subjective wellbeing differs in that it quantifies an individual's self-assessed wellbeing before, during, and after an intervention. Both of these approaches aim to represent a wider perspective than health alone. Finally, Marsh et al. mention multi-criteria decision analysis (MCDA) approaches which involve developing outcome criteria against which an intervention is assessed, before weighting each criteria and summing the results so that different interventions can be ranked. There are examples of this from both the UK and the US however these have been developed for bespoke purposes and have not been more widely adopted.[285,286]

Both Edwards and Banke-Thomas discuss social return on investment (SROI) as a method of quantifying non-health outcomes.[17,287] This is used in public sector settings by the UK Cabinet Office to measure policy outcomes and, like CBAs, SROI requires assigning a monetary value to all outcomes so that they can be summed.[288] If SROI is used to assess public health interventions but only health outcomes are quantified, then the outcome is no different from simply estimating the return on investment using a health perspective and suffers the same unfair comparison with healthcare interventions discussed by Prof Ferguson.[274] However, if non-health outcomes are included, such as those quantified by the Department of Health wider societal benefits tool used by PRIMEtime CE (discussed in section 3.1) then public health interventions could be more readily compared with a range of other non-health government policies.[150] However, this still does not resolve the issues of how to assign monetary values to wider societal and environmental outcomes, and whether public health interventions should be compared with non-health government policies when funding decisions are prioritised.

It is evident that CUAs do not fully capture the costs and benefits associated with public health interventions and using alternative methods such as SROI, CCA, CBA, MCDA, subjective wellbeing, or capability approaches may better represent the breadth of outcomes. However, the use of CUAs has a significant advantage as a result of their dominance in how health and public health outcomes have historically been evaluated in England. By quantifying social care costs as part of the primary analysis and wider societal costs as sensitivity analyses, this thesis does include some outcomes beyond health, however, a wider debate and subsequent consensus is required if public health economic interventions are to move away from CUAs. This would substantially improve the quantification of costs and benefits from public health interventions, although care must be taken to ensure that any changes enhance rather than diminish the mechanisms by which public health policies are prioritised.

7.4 Conclusion

This thesis aimed to develop a model to compare the cost-effectiveness of interventions affecting diet and physical activity. As part of the thesis, an approach to choosing a modelling structure for public health economic models has been published, comparable methods for identifying disease costs and utilities have been developed, the PRIMEtime CE model has been created, and the cost-effectiveness of diet and physical activity interventions have been evaluated.

The intention is for PRIMEtime CE to be made freely available and for it to be used widely to evaluate public health interventions and inform public health policy. Future developments of the model will increase its policy relevance and help deliver on the underlying aim of this thesis, to improve population health.

References

1 Institute for Health Metrics and Evaluation (IHME). GBD Compare. 2017. http://vizhub.healthdata.org/gbd-compare/ (accessed June 22, 2017).

2 World Health Organization Regional Office for Europe. The case for investing in public health. Copenhagen, 2014 http://www.euro.who.int/__data/assets/pdf_file/0009/278073/Case-Investing-Public-Health.pdf.

3 Owen L, Morgan A, Fischer A, Ellis S, Hoy A, Kelly MP. The cost-effectiveness of public health interventions. *J Public Health* 2012; **34**: 37–45.

4 World Health Organization. Interventions on diet and physical activity: what works: summary report. Geneva, 2009 http://www.who.int/dietphysicalactivity/summary-report-09.pdf.

5 National Institute for Health and Care Excellence. Physical activity. Local government briefing [LGB3]. London, 2012 https://www.nice.org.uk/advice/lgb3/chapter/introduction.

6 NHS England. Five Year Forward View. London, 2014 https://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf.

7 UK Government. Health and Social Care Act 2012 c.7. 2012 http://www.legislation.gov.uk/ukpga/2012/7/contents.

8 Local Government Association. Future funding outlook for councils 2019/20. Interim 2015 update. London, 2015 https://www.local.gov.uk/sites/default/files/documents/future-funding-outlook-co-18b.pdf.

9 House of Commons Health Committee. Public health post-2013. London, 2016 https://www.publications.parliament.uk/pa/cm201617/cmselect/cmhealth/140/14002.htm?utm_source=140&utm_medium=fullbullet&utm_campaign=modulereports.

10 van Gils PF, Tariq L, Verschuuren M, van den Berg M. Cost-effectiveness research on preventive interventions: a survey of the publications in 2008. *Eur J Public Health* 2011; **21**: 260–4.

11 Masters R, Anwar E, Collins B, Cookson R, Capewell S. Return on investment of public health interventions: a systematic review. *J Epidemiol Community Health* 2017; published online March 29. DOI:10.1136/jech-2016-208141.

12 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*; **15**: 796–803.

13 National Institute for Health and Care Excellence. Developing NICE guidelines: the manual. London, 2014 https://www.nice.org.uk/media/default/about/what-we-do/our-programmes/developing-nice-guidelines-the-manual.pdf.

14 Marsh K, Phillips CJ, Fordham R, Bertranou E, Hale J. Estimating cost-effectiveness in public health: a summary of modelling and valuation methods. *Health Econ Rev* 2012; **2**: 17.

15 Weatherly H, Drummond M, Claxton K, *et al.* Methods for assessing the cost-effectiveness of public health interventions: key challenges and recommendations. *Health Policy* 2009; **93**: 85–92.

16 Drummond M, Weatherly H, Claxton K, *et al.* Assessing the Challenges of Applying Standard Methods of Economic Evaluation to Public Health Interventions. York, 2007 http://phrc.lshtm.ac.uk/papers/PHRC_D1-05_Final_Report.pdf.

17 Edwards RT, Charles JM, Lloyd-Williams H. Public health economics: a systematic review of guidance for the economic evaluation of public health interventions and discussion of key methodological issues. *BMC Public Health* 2013; **13**: 1001.

18 Payne K, McAllister M, Davies LM. Valuing the economic benefits of complex interventions: when maximising health is not sufficient. *Health Econ* 2013; **22**: 258–71.

19 Kelly MP, McDaid D, Ludbrook A, Powell J. Economic appraisal of public health interventions. London, 2005 http://www.cawt.com/Site/11/Documents/Publications/Population Health/Economics of Health Improvement/Economic_appraisal_of_public_health_interventions.pdf.

20 Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. *BMJ* 2008; **336**: 1281–3.

21 Cooper K, Brailsford SC, Davies R. Choice of modelling technique for evaluating health care interventions. *J Oper Res Soc* 2007; **58**: 168–76.

22 Squires H, Chilcott J, Akehurst R, Burr J, Kelly MP. A systematic literature review of the key challenges for developing the structure of public health economic models. *Int J Public Health* 2016; **61**: 289–98.

23 Kelly M, Morgan A, Ellis S, Younger T, Huntley J, Swann C. Evidence based public health: A review of the experience of the National Institute of Health and Clinical Excellence (NICE) of developing public health guidance in England. *Soc Sci Med* 2010; **71**: 1056–62.

24 Kaltenthaler E, Tappenden P, Paisley S, Squires H. NICE DSU technical support document 13: Idenfitfying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models. Sheffield, 2011 https://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0092948/.

25 Palmer AJ, Clarke P, Gray A, *et al.* Computer modeling of diabetes and its complications: a report on the Fifth Mount Hood challenge meeting. *Value Health* 2013; **16**: 670–85.

26 Hoogendoorn M, Feenstra TL, Asukai Y, *et al.* External Validation of Health Economic Decision Models for Chronic Obstructive Pulmonary Disease (COPD): Report of the Third COPD Modeling Meeting. *Value Health* 2017; **20**: 397–403.

27 Mandelblatt JS, Stout NK, Schechter CB, *et al.* Collaborative Modeling of the Benefits and Harms Associated With Different U.S. Breast Cancer Screening Strategies. *Ann Intern Med* 2016; **164**: 215–25.

28 Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes, 4th edn. Oxford: Oxford University Press, 2015.

29 Weinstein MC, Manning WG. Theoretical issues in cost-effectiveness analysis. *J Health Econ* 1997; **16**: 121–8.

30 Gray AM, Clarke PM, Wolstenholme JL, Wordsworth S. Applied Methods of Cost-effectiveness Analysis in Health Care. Oxford: Oxford University Press, 2011.

31 Trueman P, Anokye NK. Applying economic evaluation to public health

interventions: the case of interventions to promote physical activity. *J Public Health* 2013; **35**: 32–9.

32 Dallongeville J, Dauchet L, de Mouzon O, Réquillart V, Soler L-G. Increasing fruit and vegetable consumption: a cost-effectiveness analysis of public policies. *Eur J Public Health* 2011; **21**: 69–73.

33 McIntosh E, Louviere J, Frew EJ, Clarke PM. Applied Methods of Cost-Benefit Analysis in Health Care. Oxford: Oxford University Press, 2010.

34 Cookson R, McCabe C, Tsuchiya A. Public healthcare resource allocation and the Rule of Rescue. *J Med Ethics* 2008; **34**: 540–4.

35 Boadway RW, Bruce N. Welfare economics. Wiley-Blackwell, 1984.

36 Brouwer WBF, Culyer AJ, van Exel NJA, Rutten FFH. Welfarism vs. extra-welfarism. *J Health Econ* 2008; **27**: 325–38.

37 Briggs A, Claxton K, Sculpher M. Decision Modelling for Health Economic Evaluation. Oxford: Oxford University Press, 2006.

38 Weisbrod BA. Costs and Benefits of Medical Research: A Case Study of Poliomyelitis. *J Polit Econ* 1971; **79**: 527–44.

39 Coast J. Maximisation in extra-welfarism: A critique of the current position in health economics. *Soc Sci Med* 2009; **69**: 786–92.

40 Elixhauser A, Luce BR, Taylor WR, Reblando J. Health care CBA/CEA: an update on the growth and composition of the literature. *Med Care* 1993; **31**: JS1-11, JS18-149.

41 Elixhauser A, Halpern M, Schmier J, Luce BR. Health care CBA and CEA from 1991 to 1996: an updated bibliography. *Med Care* 1998; **36**: MS1-9, MS18-147.

42 Garber AM, Phelps CE. Economic foundations of cost-effectiveness analysis. *J Health Econ* 1997; **16**: 1–31.

43 Coast J, Smith R, Lorgelly P. Should the capability approach be applied in Health Economics? *Health Econ* 2008; **17**: 667–70.

44 Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. *J Health Serv Res Policy* 2004; **9**: 110–8.

45 Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ* 2006; **15**: 1295–310.

46 Squires H, Chilcott J, Akehurst R, Burr J, Kelly MP. A Framework for Developing the Structure of Public Health Economic Models. *Value Health* 2016; **19**: 588–601.

47 Squires H. A methodological framework for developing the structure of Public Health economic models. 2014; published online Jan 1. http://etheses.whiterose.ac.uk/5316/1/Final doctoral thesis with appendices.pdf (accessed June 22, 2017).

48 National Institute for Health and Care Excellence. Section 3, The manufacturer's submission. In, Ticagrelor for the treatment of acute coronary syndromes. NICE technology appraisal guidance TA 236. London, 2011 https://www.nice.org.uk/guidance/ta236/chapter/3-The-manufacturers-submission#cost-effectiveness.

49 Groot Koerkamp B, Stijnen T, Weinstein MC, Hunink MGM. The combined analysis of uncertainty and patient heterogeneity in medical decision models. *Med*

Decis Making; **31**: 650–61.

50 Trogdon J, Finkelstein EA, Reyes M, Dietz WH. A return-on-investment simulation model of workplace obesity interventions. *J Occup Environ Med* 2009; **51**: 751–8.

51 Howard P, Knight C, Boler A, Baker C. Cost-utility analysis of varenicline versus existing smoking cessation strategies using the BENESCO Simulation model: application to a population of US adult smokers. *Pharmacoeconomics* 2008; **26**: 497–511.

52 Bolin K, Wilson K, Benhaddi H, *et al.* Cost-effectiveness of varenicline compared with nicotine patches for smoking cessation-results from four European countries. *Eur J Public Health* 2009; **19**: 650–4.

53 Hoogendoorn M, Welsing P, Rutten-van Mölken MPMH. Cost-effectiveness of varenicline compared with bupropion, NRT, and nortriptyline for smoking cessation in the Netherlands. *Curr Med Res Opin* 2008; **24**: 51–61.

54 Kautiainen K, Ekroos H, Puhakka M, *et al.* Re-treatment with varenicline is a cost-effective aid for smoking cessation. *J Med Econ* 2017; **20**: 246–52.

55 Macmillan A, Connor J, Witten K, Kearns R, Rees D, Woodward A. The societal costs and benefits of commuter bicycling: simulating the effects of specific policies using system dynamics modeling. *Environ Health Perspect* 2014; **122**: 335–44.

56 The CDC Diabetes Cost-effectiveness Group. Cost-effectiveness of Intensive Glycemic Control, Intensified Hypertension Control, and Serum Cholesterol Level Reduction for Type 2 Diabetes. *JAMA* 2002; **287**: 2542.

57 Herman WH, Hoerger TJ, Brandle M, *et al.* The cost-effectiveness of lifestyle modification or metformin in preventing type 2 diabetes in adults with impaired glucose tolerance. *Ann Intern Med* 2005; **142**: 323–32.

58 Fone D, Hollinghurst S, Temple M, *et al.* Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *J Public Health Med* 2003; **25**: 325–35.

59 Nandi A, Megiddo I, Prabhakaran D, Laxminarayan R. An agent-based simulation modelling approach to extended cost-effectiveness analysis of health interventions. *Lancet* 2013; **381**: S96.

60 Cobiac LJ, Vos T, Barendregt JJ. Cost-effectiveness of interventions to promote physical activity: a modelling study. *PLoS Med* 2009; **6**: e1000110.

61 Cobiac LJ, Vos T, Veerman JL. Cost-effectiveness of interventions to reduce dietary salt intake. *Heart* 2010; **96**: 1920–5.

62 Trueman P, Haynes SM, Felicity Lyons G, *et al.* Long-term cost-effectiveness of weight management in primary care. *Int J Clin Pract* 2010; **64**: 775–83.

63 Petrou S, Gray A. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. *BMJ* 2011; **342**: d1766.

64 The EConDA project. Cost-effectiveness of interventions to prevent, screen and treat chronic diseases: A review. 2014 http://econdaproject.eu/publications_18_361135336.pdf.

65 Asaria P, Chisholm D, Mathers C, Ezzati M, Beaglehole R. Chronic disease prevention: health effects and financial costs of strategies to reduce salt intake and control tobacco use. *Lancet* 2007; **370**: 2044–53.

66 Webb M, Fahimi S, Singh GM, *et al.* Cost effectiveness of a government supported policy strategy to decrease sodium intake: global analysis across 183 nations. *BMJ* 2017; **356**: i6699.

67 Jacobs-van der Bruggen MAM, Bos G, Bemelmans WJ, Hoogenveen RT, Vijgen SM, Baan CA. Lifestyle interventions are cost-effective in people with different levels of diabetes risk: results from a modeling study. *Diabetes Care* 2007; **30**: 128–34.

68 Bemelmans W, van Baal P, Wendel-Vos W, *et al.* The costs, effects and cost-effectiveness of counteracting overweight on a population level. A scientific base for policy targets for the Dutch national plan for action. *Prev Med* 2008; **46**: 127–32.

69 van Gils PF, Over EAB, Hamberg-van Reenen HH, *et al.* The polypill in the primary prevention of cardiovascular disease: cost-effectiveness in the Dutch population. *BMJ Open* 2011; **1**: e000363.

70 Hoogenveen RT, van Baal PHM, Boshuizen HC. Chronic disease projections in heterogeneous ageing populations: approximating multi-state models of joint distributions by modelling marginal distributions. *Math Med Biol* 2010; **27**: 1–19.

71 Hurley SF, Matthews JP. The Quit Benefits Model: a Markov model for assessing the health benefits and health care cost savings of quitting smoking. *Cost Eff Resour Alloc* 2007; **5**: 2.

72 Hurley SF, Matthews JP. Cost-effectiveness of the Australian National Tobacco Campaign. *Tob Control* 2008; **17**: 379–84.

73 Roux L, Pratt M, Tengs TO, *et al.* Cost effectiveness of community-based physical activity interventions. *Am J Prev Med* 2008; **35**: 578–88.

74 Liew D, Lim SS, Bertram M, McNeil JJ, Vos T. A model for undertaking effectiveness and cost-effectiveness analyses of primary preventive strategies in cardiovascular disease. *Eur J Cardiovasc Prev Rehabil* 2006; **13**: 515–22.

75 Lazar LD, Pletcher MJ, Coxson PG, Bibbins-Domingo K, Goldman L. Cost-effectiveness of statin therapy for primary prevention in a low-cost statin era. *Circulation* 2011; **124**: 146–53.

76 Pletcher MJ, Lazar L, Bibbins-Domingo K, *et al.* Comparing impact and cost-effectiveness of primary prevention strategies for lipid-lowering. *Ann Intern Med* 2009; **150**: 243–54.

77 Palmer AJ, Roze S, Valentine WJ, Spinas GA, Shaw JE, Zimmet PZ. Intensive lifestyle changes or metformin in patients with impaired glucose tolerance: modeling the long-term health economic implications of the diabetes prevention program in Australia, France, Germany, Switzerland, and the United Kingdom. *Clin Ther* 2004; **26**: 304–21.

78 Adarkwah CC, Gandjour A, Akkerman M, Evers SM. Cost-effectiveness of angiotensin-converting enzyme inhibitors for the prevention of diabetic nephropathy in The Netherlands-a Markov model. *PLoS One* 2011; **6**: e26139.

79 Galani C, Schneider H, Rutten FFH. Modelling the lifetime costs and health effects of lifestyle intervention in the prevention and treatment of obesity in Switzerland. *Int J Public Health* 2007; **52**: 372–82.

80 Lindgren P, Lindström J, Tuomilehto J, *et al.* Lifestyle intervention to prevent diabetes in men and women with impaired glucose tolerance is cost-effective. *Int J Technol Assess Health Care* 2007; **23**: 177–83.

81 Peura P, Martikainen J, Soini E, Hallinen T, Niskanen L. Cost-effectiveness of

statins in the prevention of coronary heart disease events in middle-aged Finnish men. *Curr Med Res Opin* 2008; **24**: 1823–32.

82 Menn P, Leidl R, Holle R. A lifetime Markov model for the economic evaluation of chronic obstructive pulmonary disease. *Pharmacoeconomics* 2012; **30**: 825–40.

83 Pignone M, Earnshaw S, Tice JA, Pletcher MJ. Aspirin, statins, or both drugs for the primary prevention of coronary heart disease events in men: a cost-utility analysis. *Ann Intern Med* 2006; **144**: 326–36.

84 Smith KJ, Hsu HE, Roberts MS, *et al.* Cost-effectiveness analysis of efforts to reduce risk of type 2 diabetes and cardiovascular disease in southwestern Pennsylvania, 2005-2007. *Prev Chronic Dis* 2010; **7**: A109.

85 Gulliford MC, Charlton J, Bhattarai N, Charlton C, Rudisill C. Impact and cost-effectiveness of a universal strategy to promote physical activity in primary care: population-based cohort study and Markov model. *Eur J Heal Econ* 2014; **15**: 341–51.

86 Ward S, Lloyd Jones M, Pandor A, *et al.* A systematic review and economic evaluation of statins for the prevention of coronary events. *Health Technol Assess* 2007; **11**: 1–160, iii–iv.

87 Matrix. Estimating Return on Investment for interventions and strategies to increase physical activity - Technical Report. 2014 https://www.nice.org.uk/Media/Default/About/what-we-do/Into-practice/Return-on-Investment/NICE-return-on-investment-physical-activity-technical-report.pdf.

88 Dalziel K, Segal L. Time to give nutrition interventions a higher profile: cost-effectiveness of 10 nutrition interventions. *Health Promot Int* 2007; **22**: 271–83.

89 Palmer AJ, Tucker DMD. Cost and clinical implications of diabetes prevention in an Australian setting: a long-term modeling analysis. *Prim Care Diabetes* 2012; **6**: 109–21.

90 Homer JB, Hirsch GB. System dynamics modeling for public health: background and opportunities. *Am J Public Health* 2006; **96**: 452–8.

91 Kaplan EH, Craft DL, Wein LM. Emergency response to a smallpox attack: the case for mass vaccination. *Proc Natl Acad Sci U S A* 2002; **99**: 10935–40.

92 Dangerfield BC. System dynamics applications to European health care issues. *J Oper Res Soc* 1999; **50**: 345–53.

93 Hoogenveen RT, van Baal PH, Boshuizen HC, Feenstra TL. Dynamic effects of smoking cessation on disease incidence, mortality and quality of life: The role of time since cessation. *Cost Eff Resour Alloc* 2008; **6**: 1.

94 Nghiem N, Blakely T, Cobiac LJ, Pearson AL, Wilson N, Pearson A. Health and Economic Impacts of Eight Different Dietary Salt Reduction Interventions. *PLoS One* 2015; **10**: e0123915.

95 Cobiac LJ, Magnus A, Lim S, Barendregt JJ, Carter R, Vos T. Which Interventions Offer Best Value for Money in Primary Prevention of Cardiovascular Disease? *PLoS One* 2012; **7**: e41842.

96 Wilson N, Nghiem N, Eyles H, *et al.* Modeling health gains and cost savings for ten dietary salt reduction targets. *Nutr J* 2015; **15**: 44.

97 Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J. Modeling using discrete

event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. *Med Decis Making* 2012; **32**: 701–11.

98 Schlessinger L, Eddy DM. Archimedes: a new model for simulating health care systems--the mathematical formulation. *J Biomed Inform* 2002; **35**: 37–50.

99 Eddy DM, Schlessinger L, Kahn R. Clinical outcomes and cost-effectiveness of strategies for managing people at high risk for diabetes. *Ann Intern Med* 2005; **143**: 251–64.

100 Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. *J Natl Cancer Inst* 2006; **98**: 774–82.

101 Cancer Intervention and Surveillance Modeling Network. Breast Cancer Model Profiles. National Cancer Institute. 2015. http://cisnet.cancer.gov/breast/profiles.html (accessed June 22, 2017).

102 Siebers PO, Macal CM, Garnett J, Buxton D, Pidd M. Discrete-event simulation is dead, long live agent-based simulation! *J Simul* 2010; **4**: 204–10.

103 Chhatwal J, He T. Economic Evaluations with Agent-Based Modelling: An Introduction. *Pharmacoeconomics* 2015; **33**: 423–33.

104 Christakis NA, Fowler JH. The Spread of Obesity in a Large Social Network over 32 Years. *N Engl J Med* 2007; **357**: 370–9.

105 Megiddo I, Colson AR, Nandi A, *et al.* Analysis of the Universal Immunization Programme and introduction of a rotavirus vaccine in India with IndiaSim. *Vaccine* 2014; **32**: A151–61.

106 Nandi A, Megiddo I, Ashok A, Verma A, Laxminarayan R. Reduced burden of childhood diarrheal diseases through increased access to water and sanitation in India: A modeling analysis. *Soc Sci Med* 2017; **180**: 181–92.

107 Haby MM, Vos T, Carter R, *et al.* A new approach to assessing the health benefit from obesity interventions in children and adolescents: the assessing cost-effectiveness in obesity project. *Int J Obes* 2006; **30**: 1463–75.

108 Forster M, Veerman JL, Barendregt JJ, Vos T. Cost-effectiveness of diet and exercise interventions to reduce overweight and obesity. *Int J Obes* 2011; **35**: 1071–8.

109 Cobiac LJ, Vos T, Veerman JL. Cost-effectiveness of interventions to promote fruit and vegetable consumption. *PLoS One* 2010; **5**: e14148.

110 Cobiac L, Vos T, Veerman L. Cost-effectiveness of Weight Watchers and the Lighten Up to a Healthy Lifestyle program. *Aust N Z J Public Health* 2010; **34**: 240–7.

111 Murray CJL, Lauer JA, Hutubessy RCW, *et al.* Effectiveness and costs of interventions to lower systolic blood pressure and cholesterol: a global and regional analysis on reduction of cardiovascular-disease risk. *Lancet* 2003; **361**: 717–25.

112 Rubinstein A, García Martí S, Souto A, Ferrante D, Augustovski F. Generalized cost-effectiveness analysis of a package of interventions to reduce cardiovascular disease in Buenos Aires, Argentina. *Cost Eff Resour Alloc* 2009; **7**: 10.

113 Gortmaker SL, Long MW, Resch SC, *et al.* Cost Effectiveness of Childhood Obesity Interventions. *Am J Prev Med* 2015; **49**: 102–11.

114 Blakely T, Cobiac LJ, Cleghorn CL, *et al.* Health, Health Inequality, and Cost

Impacts of Annual Increases in Tobacco Tax: Multistate Life Table Modeling in New Zealand. *PLoS Med* 2015; **12**: e1001856.

115 Divajeva D, Retat L, Shaw A, Brown M, Webber L. Risk factor based modelling for Public Health England. London, 2014 http://nhfshare.heartforum.org.uk/RMAssets/UKHFreports/PHE Report_October 2014 final.pdf.

116 Sassi F, Cecchini M, Lauer JA, Chisholm D. Improving Lifestyles, Tackling Obesity: The Health and Economic Impact of Prevention Strategies. OECD Health Working Papers, No. 48. Paris, 2009 http://dx.doi.org/10.1787/220087432153.

117 Cecchini M, Sassi F, Lauer JA, Lee YY, Guajardo-Barron V, Chisholm D. Tackling of unhealthy diets, physical inactivity, and obesity: health effects and cost-effectiveness. *Lancet* 2010; **376**: 1775–84.

118 National Institute for Health and Care Excellence. Obesity prevention. Section 6: Health economics. Clinical guideline 43. London, 2006 https://www.nice.org.uk/guidance/cg43/evidence/full-guideline-section-6-health-economics-evidence-statements-and-reviews-195027235.

119 Bertram MY, Lim SS, Barendregt JJ, Vos T. Assessing the cost-effectiveness of drug and lifestyle intervention following opportunistic screening for pre-diabetes in primary care. *Diabetologia* 2010; **53**: 875–81.

120 Kang H-Y, Ko S-K, Liew D. Results of a Markov model analysis to assess the cost-effectiveness of statin therapy for the primary prevention of cardiovascular disease in Korea: the Korean Individual-Microsimulation Model for Cardiovascular Health Interventions. *Clin Ther* 2009; **31**: 2919–30.

121 Johansson P, Ostenson C-G, Hilding AM, Andersson C, Rehnberg C, Tillgren P. A cost-effectiveness analysis of a community-based diabetes prevention program in Sweden. *Int J Technol Assess Health Care* 2009; **25**: 350–8.

122 Feldman I, Hellström L, Johansson P. Heterogeneity in cost-effectiveness of lifestyle counseling for metabolic syndrome risk groups -primary care patients in Sweden. *Cost Eff Resour Alloc* 2013; **11**: 19.

123 Basu S, Andrews J. Complexity in mathematical models of public health policies: a guide for consumers of models. *PLoS Med* 2013; **10**: e1001540.

124 Sculpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models. A suggested framework and example of application. *Pharmacoeconomics* 2000; **17**: 461–77.

125 Whitty CJM. What makes an academic paper useful for health policy? *BMC Med* 2015; **13**: 301.

126 New Economy. Cost Benefit Analysis Guidance and Model. 2015. http://www.neweconomymanchester.com/our-work/research-evaluation-cost-benefit-analysis/cost-benefit-analysis/cost-benefit-analysis-guidance-and-model (accessed Aug 9, 2017).

127 Thompson R. Stakeholder Analysis. Mind Tools. 2015. http://www.mindtools.com/pages/article/newPPM_07.htm (accessed June 22, 2017).

128 WHO Framework Convention on Tobacco Control. The WHO Framework Convention on Tobacco Control: an overview. Geneva, 2015 http://www.who.int/fctc/WHO_FCTC_summary_January2015_EN.pdf?ua=1.

129 Brennan A, Meier P, Purshouse R, *et al.* The Sheffield Alcohol Policy Model - A Mathematical Description. *Health Econ* 2015; **24**: 1368–88.

130 Marsh K, Phillips CJ, Fordham R, Bertranou E, Hale J. Estimating cost-effectiveness in public health: a summary of modelling and valuation methods. *Health Econ Rev* 2012; **2**: 17.

131 Cobiac LJ, Scarborough P, Kaur A, Rayner M. The Eatwell Guide: Modelling the Health Implications of Incorporating New Sugar and Fibre Guidelines. *PLoS One* 2016; **11**: e0167859.

132 Scarborough P, Harrington RA, Mizdrak A, *et al.* The Preventable Risk Integrated ModEl and Its Use to Estimate the Health Impact of Public Health Policy Scenarios. *Scientifica* 2014; **2014**: 748750.

133 Barendregt JJ. EpiGearXL. 2016. http://www.epigear.com/Products/EpigearXL/epigearxl.html.

134 Barendregt JJ. Ersatz. 2016. http://www.epigear.com/index_files/ersatz.html.

135 Microsoft. Microsoft Excel. 2013.

136 HM Treasury. The Green Book: appraisal and evaluation in central government. London, 2016 https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-governent.

137 Cobiac L, Scarborough P. Translating the World Health Organization 25x25 goals into a United Kingdom context: The PROMISE modelling study. *BMJ Open* 2017; **7**: e012805.

138 Barbaric M, Brooks E, Moore L, Cheifetz O. Effects of Physical Activity on Cancer Survival: A Systematic Review. *Physiother Canada* 2010; **62**: 25–34.

139 Hay JW, Smeeding J, Carroll N V, *et al.* Good research practices for measuring drug costs in cost effectiveness analyses: issues and recommendations: the ISPOR Drug Cost Task Force report-Part I. *Value Health* 2010; **13**: 3–7.

140 Chapko MK, Liu C-F, Perkins M, Li Y-F, Fortney JC, Maciejewski ML. Equivalence of two healthcare costing methods: bottom-up and top-down. *Health Econ* 2009; **18**: 1188–201.

141 Olsson TM. Comparing top-down and bottom-up costing approaches for economic evaluation within social welfare. *Eur J Health Econ* 2011; **12**: 445–53.

142 Mercier G, Naro G. Costing hospital surgery services: the method matters. *PLoS One* 2014; **9**: e97290.

143 Yabroff KR, Borowski L, Lipscomb J. Economic studies in colorectal cancer: challenges in measuring and comparing costs. *J Natl Cancer Inst Monogr* 2013; **2013**: 62–78.

144 Frew EJ, Bhatti M, Win K, *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* 2014; **48**: 207–12.

145 Adam T, Evans DB, Murray CJ. Econometric estimation of country-specific hospital costs. *Cost Eff Resour Alloc* 2003; **1**: 3.

146 World Health Organization. Cost effectiveness and strategic planning (WHO-CHOICE). 2014. http://www.who.int/choice/cost-effectiveness/en/ (accessed June 22, 2017).

147 Medicines & Healthcare products Regulatory Authority. The Clinical Practice Research Datalink. 2017. DOI:22nd June 2017.

148 Bhimjiyani A, Knuchel-Takano A, Jaccard A, Brown M, Webber L. Tipping the Scales: why preventing obesity makes economic sense. Technical Report. London, 2016 http://nhfshare.heartforum.org.uk/RMAssets/Modelling/CRUK UKHF Obesity Technical summary Published.pdf.

149 NHS England. 2012-13 Programme Budgeting Benchmarking Tool. 2014. https://www.networks.nhs.uk/nhs-networks/health-investment-network/news/2012-13-programme-budgeting-data-is-now-available (accessed June 22, 2017).

150 Claxton K, Sculpher M, Palmer S, Culyer AJ. Causes for concern: is NICE failing to uphold its responsibilities to all NHS patients? *Health Econ* 2015; **24**: 1–7.

151 van Baal PHM, Feenstra TL, Polder JJ, Hoogenveen RT, Brouwer WBF. Economic evaluation and the postponement of health care costs. *Health Econ* 2011; **20**: 432–45.

152 Morton A, Adler AI, Bell D, *et al*. Unrelated Future Costs and Unrelated Future Benefits: Reflections on NICE Guide to the Methods of Technology Appraisal. *Health Econ* 2016; **25**: 933–8.

153 National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. London, 2013 http://www.nice.org.uk/article/pmg9/resources/non-guidance-guide-to-the-methods-of-technology-appraisal-2013-pdf.

154 van Baal P, Meltzer D, Brouwer W. Future Costs, Fixed Healthcare Budgets, and the Decision Rules of Cost-Effectiveness Analysis. *Health Econ* 2016; **25**: 237–48.

155 NHS England. 2013-14 CCG Programme Budgeting Benchmarking Tool. London, 2015 https://www.england.nhs.uk/wp-content/uploads/2015/06/13-14-ccg-prog-bug-benchmarking-tool.xlsm.

156 NHS England. NHS allocations for 2013/14. 2012. http://webarchive.nationalarchives.gov.uk/20161103155843/https://www.england.nhs.uk/allocations-2013-14/ (accessed June 22, 2017).

157 NHS England. Specialised services expenditure by disease category. Personal Communication. 2016.

158 NHS England. Manual for prescribed specialised services for 2013/14. London, 2014.

159 NHS England. 2014-15 Programme Budgeting Mappings and Definitions v1.0 (10.08.15). London, 2015.

160 Luengo-Fernandez R. Resource costs, health outcomes and cost-effectivenes in stroke care: evidence from the Oxford Vascular Study. 2009. https://ora.ox.ac.uk/objects/uuid:a5012d51-6794-48a7-bb78-4e5166e8cfdf.

161 Husereau D, Drummond M, Petrou S, *et al*. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMJ* 2013; **346**: f1049.

162 Critical Appraisal Skills Programme. CASP Economic Evaluation Checklist. Oxford, 2013 http://media.wix.com/ugd/dded87_3b2bd5743feb4b1aaac6ebdd68771d3f.pdf.

163 Health and Social Care Information Centre. Hospital Episode Statistics. 2015.

http://www.hscic.gov.uk/hes (accessed June 22, 2017).

164 NHS England. 2012-13 Programme Budgeting Mappings & Definitions. London, 2013 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216917/Programme-Budgeting-Mappings-Definitions.xls.

165 NHS England. Technical Guide to Allocation Formulae and Pace of Change. Leeds, 2016 https://www.england.nhs.uk/wp-content/uploads/2016/04/1-allctins-16-17-tech-guid-formulae.pdf.

166 Gardiner L, Everard K. Primary medical care - new workload formula for allocations to CCG areas. Leeds, 2016 https://www.england.nhs.uk/wp-content/uploads/2016/04/5-primary-care-allctins-16-17.pdf.

167 Office for National Statistics. Births by Parents' Characteristics, England and Wales: 2014. London, 2015 https://www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/datasets/birthsbyparentscharacteristics/2014/3birthsbyparentscharacteristics2014final.xls.

168 PSSRU. Personal Social Services Research Unit. 2015. http://www.pssru.ac.uk/index-kent-lse.php (accessed June 22, 2017).

169 The Money Advice Service. Means tests for help with care costs – how they work - Money Advice Service. 2017. https://www.moneyadviceservice.org.uk/en/articles/means-tests-for-help-with-care-costs-how-they-work (accessed June 22, 2017).

170 The NHS Information Centre. Personal Social Services Adult Social Care Survey, England - 2011-12. London, 2012 https://data.gov.uk/dataset/personal_social_services_adult_social_care_survey_england/resource/2a34d0e9-a758-4f1c-89af-0ed63e1328b2.

171 Department of Health. GP Patient Survey 2011. London, 2012 https://gp-patient.co.uk/.

172 Institute for Social and Economic Research. Understanding Society. 2017. https://www.understandingsociety.ac.uk/ (accessed June 22, 2017).

173 Office for National Statistics. Annual Survey of Hours and Earnings: 2014. London, 2014 http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/rel/ashe/annual-survey-of-hours-and-earnings/2014-revised-results/index.html.

174 Curtis L. Unit Costs of Health and Social Care 2014. Canterbury, 2014 http://www.pssru.ac.uk/project-pages/unit-costs/2014/.

175 Hex N, Bartlett C, Wright D, Taylor M, Varley D. Estimating the current and future costs of Type 1 and Type 2 diabetes in the UK, including direct health costs and indirect societal and productivity costs. *Diabet Med* 2012; **29**: 855–62.

176 Luengo-Fernandez R, Leal J, Gray A, Sullivan R. Economic burden of cancer across the European Union: a population-based cost analysis. *Lancet Oncol* 2013; **14**: 1165–74.

177 Office for National Statistics. Mid-2009 Population Estimates: estimated resident population by single year of age and sex. London, 2012 http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.

uk/ons/rel/pop-estimate/population-estimates-for-uk--england-and-wales--scotland-and-northern-ireland/2009/index.html.

178 OFX. Yearly Average Rates. OFX. 2016. https://www.ofx.com/en-gb/forex-news/historical-exchange-rates/yearly-average-rates/ (accessed June 22, 2017).

179 Curtis L, Burns A. Unit Costs of Health and Social Care 2015. Canterbury, 2015 http://www.pssru.ac.uk/project-pages/unit-costs/2015/.

180 Appleby J, Harrison T, Foot C, Smith A, Gilmour S. Explaining Variations in Primary Care Trusts' Spending on Cancer Services. London, 2011 https://www.kingsfund.org.uk/publications/articles/explaining-variations-primary-care-trusts-spending-cancer-services.

181 Claxton K, Martin S, Soares M, *et al.* Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technol Assess* 2015; **19**: 1–503, v–vi.

182 Department of Health. Reference Costs 2015-16. London, 2016 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/577083/Reference_Costs_2015-16.pdf.

183 Dolan P, Torgerson DJ, Wolstenholme J. Costs of breast cancer treatment in the United Kingdom. *Breast* 1999; **8**: 205–7.

184 Bhatnagar P, Wickramasinghe K, Williams J, Rayner M, Townsend N. The epidemiology of cardiovascular disease in the UK 2014. *Heart* 2015; **101**: 1182–9.

185 World Health Organization. DisMod II. 2014. http://www.epigear.com/index_files/dismod_ii.html.

186 Scarborough P, Smolina K, Mizdrak A, Cobiac L, Briggs A. Assessing the external validity of model-based estimates of the incidence of heart attack in England: a modelling study. *BMC Public Health* 2016; **16**: 1135.

187 Currie CJ, Kraus D, Morgan CL, Gill L, Stott NCH, Peters JR. NHS acute sector expenditure for diabetes: the present, future, and excess in-patient cost of care. *Diabet Med* 1997; **14**: 686–92.

188 Farmer AJ, Wade AN, French DP, *et al.* Blood glucose self-monitoring in type 2 diabetes: a randomised controlled trial. *Health Technol Assess* 2009; **13**: iii–iv, ix–xi, 1-50.

189 York Health Economics Consortium. Bowel Cancer Services: Costs and Benefits. York, 2007 https://www.sheffield.ac.uk/polopoly_fs/1.44049!/file/FinalBowelCancerSummaryReport-Apr07.pdf.

190 O'Leary BA, Olynyk JK, Neville AM, Platell CF. Cost-effectiveness of colorectal cancer screening: comparison of community-based flexible sigmoidoscopy with fecal occult blood testing and colonoscopy. *J Gastroenterol Hepatol* 2004; **19**: 38–47.

191 O'Neill CB, Atoria CL, O'Reilly EM, LaFemina J, Henman MC, Elkin EB. Costs and trends in pancreatic cancer treatment. *Cancer* 2012; **118**: 5132–9.

192 National Institute for Health and Care Excellence. Assessing Cost Effectiveness. In: The Guidelines Manual. London: National Institute for Health and Care Excellence, 2012: 107–22.

193 Gray AM, Clarke PM, Wolstenholm JL, Wordsworth S. Measuring, valuing, and

analysing health outcomes. In: Applied Methods of Cost-effectiveness Analysis in Health Care. Oxford: Oxford University Press, 2011: 83–118.

194 Papaioannou D, Brazier J, Paisley S. NICE DSU technical support document 9: The identification, review and synthesis of health state utility values from the literature. Sheffield, 2010.

195 EuroQol Research Foundation. About EQ-5D. 2016. https://euroqol.org/eq-5d-instruments/ (accessed June 22, 2017).

196 Macran S. The Relationship between Body Mass Index and Health-Related Quality of Life. York, 2004 https://www.york.ac.uk/che/pdf/DP190.pdf.

197 Lewis JD, Brown A, Localio AR, Schwartz JS. Initial evaluation of rectal bleeding in young persons: a cost-effectiveness analysis. *Ann Intern Med* 2002; **136**: 99–110.

198 Ko CY, Maggard M, Livingston EH. Evaluating health utility in patients with melanoma, breast cancer, colon cancer, and lung cancer: a nationwide, population-based assessment. *J Surg Res* 2003; **114**: 1–5.

199 U.K. Prospective Diabetes Study Group. Quality of life in type 2 diabetic patients is affected by complications but not by intensive policies to improve blood glucose or blood pressure control (UKPDS 37). *Diabetes Care* 1999; **22**: 1125–36.

200 Sullivan PW, Slejko JF, Sculpher MJ, Ghushchyan V. Catalogue of EQ-5D scores for the United Kingdom. *Med Decis Making* 2011; **31**: 800–4.

201 Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997; **35**: 1095–108.

202 ICD10Data.com. ICD10Data.com. 2016. http://www.icd10data.com/ (accessed June 22, 2017).

203 Sullivan PW, Ghushchyan V. Preference-Based EQ-5D index scores for chronic conditions in the United States. *Med Decis Making* 2006; **26**: 410–20.

204 Higgins JP, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. The Cochrane Collaboration, 2011 www.handbook.cochrane.org.

205 Prescott RJ, Kunkler IH, Williams LJ, *et al.* A randomised controlled trial of postoperative radiotherapy following breast-conserving surgery in a minimum-risk older population. The PRIME trial. *Health Technol Assess* 2007; **11**: 1–149, iii–iv.

206 Hall PS, Hamilton P, Hulme CT, *et al.* Costs of cancer care for use in economic evaluation: a UK analysis of patient-level routine health system data. *Br J Cancer* 2015; **112**: 948–56.

207 Haines TP, Sinnamon P, Wetzig NG, *et al.* Multimodal exercise improves quality of life of women being treated for breast cancer, but at what cost? Randomized trial with economic evaluation. *Breast Cancer Res Treat* 2010; **124**: 163–75.

208 Vrettos I, Kamposioras K, Kontodimopoulos N, *et al.* Comparing health-related quality of life of cancer patients under chemotherapy and of their caregivers. *ScientificWorldJournal* 2012; **2012**: 135283.

209 Brown M, Marsh T, Rtveladze K, *et al.* Managing overweight and obesity among children. Report on Economic Modelling and Cost Consquence Analysis. 2013 https://www.nice.org.uk/guidance/ph47/evidence/economic-modelling-report-430357501.

210 Romanus D, Kindler HL, Archer L, *et al.* Does health-related quality of life improve for advanced pancreatic cancer patients who respond to gemcitabine? Analysis of a randomized phase III trial of the cancer and leukemia group B (CALGB 80303). *J Pain Symptom Manage* 2012; **43**: 205–17.

211 Ara R, Brazier JE. Using Health State Utility Values from the General Population to Approximate Baselines in Decision Analytic Models when Condition-Specific Data are Not Available. *Value Health* 2011; **14**: 539–45.

212 FSA in Northern Ireland. Salt Targets. 2016. http://www.food.gov.uk/northern-ireland/nutritionni/salt-ni/salt_targets (accessed June 22, 2017).

213 Scientific Advisory Committee on Nutrition. Salt and Health. London, 2003 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/338782/SACN_Salt_and_Health_report.pdf.

214 Food Standards Agency. Timeline of salt reduction progress. 2016. https://www.food.gov.uk/northern-ireland/nutritionni/salt-ni/salt_targets/salt-timeline (accessed June 22, 2017).

215 Wyness LA, Butriss JL, Stanner SA. Reducing the population's sodium intake: the UK Food Standards Agency's salt reduction programme. *Public Health Nutr* 2017; **15**: 254–61.

216 Public Health England. National Diet and Nutrition Survey: assessment of dietary sodium. Adults (19 to 64 years) in England, 2014. London, 2016 https://www.gov.uk/government/statistics/national-diet-and-nutrition-survey-assessment-of-dietary-sodium-in-adults-in-england-2014.

217 Department of Health. Public Health Responsibility Deal. 2015. https://responsibilitydeal.dh.gov.uk/ (accessed June 22, 2017).

218 NatCen Social Research, MRC Human Nutrition Research, University College London Medical School. National Diet and Nutrition Survey Years 1-4, 2008/09-2011/12. [data collection]. 7th Edition. 2015 http://dx.doi.org/10.5255/UKDA-SN-6533-6.

219 Collins M, Mason H, O'Flaherty M, Guzman-Castillo M, Critchley J, Capewell S. An economic evaluation of salt reduction policies to reduce coronary heart disease in England: a policy modeling study. *Value Health* 2014; **17**: 517–24.

220 Food Standards Agency. Impact Assessment of the Revised Salt Reduction Targets - Annex 1. 2009 https://www.food.gov.uk/sites/default/files/multimedia/pdfs/annex1parasreleasedmay09.pdf.

221 Public Health England. Public Health England sodium monitoring costs. Personal Communication. 2016.

222 Chisholm D, Mendis S, Abegunde D, Martinez RX. WHO costing tool for prevention and control of non-communicable diseases. Geneva, 2011.

223 He FJ, Li J, MacGregor GA. Effect of longer term modest salt reduction on blood pressure: Cochrane systematic review and meta-analysis of randomised trials. *BMJ* 2013; **346**: f1325.

224 Lewington S, Clarke R, Qizilbash N, Peto R, Collins R, Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* 2002; **360**: 1903–13.

225 R Core Team. R: A language and environment for statistical computing. 2016.

226 Harper H, Hallsworth M. Counting Calories. How under-reporting can explain the apparent fall in calorie intake. London, 2016 http://www.behaviouralinsights.co.uk/publications/counting-calories-how-under-reporting-can-explain-the-apparent-fall-in-calorie-intake/.

227 Department for Environment Food and Rural Affairs. Family Food 2014. London, 2015 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/485982/familyfood-2014report-17dec15.pdf.

228 Mente A, O'Donnell M, Rangarajan S, *et al.* Associations of urinary sodium excretion with cardiovascular events in individuals with and without hypertension: a pooled analysis of data from four studies. *Lancet* 2016; **388**: 465–75.

229 Cogswell ME, Mugavero K, Bowman BA, Frieden TR. Dietary Sodium and Cardiovascular Disease Risk — Measurement Matters. *N Engl J Med* 2016; **375**: 580–6.

230 Cobb LK, Anderson CAM, Elliott P, *et al.* Methodological Issues in Cohort Studies That Relate Sodium Intake to Cardiovascular Disease Outcomes: A Science Advisory From the American Heart Association. *Circulation* 2014; **129**: 1173–86.

231 Cook NR, Appel LJ, Whelton PK. Lower Levels of Sodium Intake and Reduced Cardiovascular Risk. *Circulation* 2014; **129**: 981–9.

232 World Health Organization. Guideline: Sodium intake for adults and children. Geneva, 2012 http://www.who.int/nutrition/publications/guidelines/sodium_intake_printversion.pdf.

233 Institute of Medicine. Sodium Intake in Populations. Assessment of Evidence. Washington, D.C.: National Academies Press, 2013 DOI:10.17226/18311.

234 Be Active Birmingham. Be Active Birmingham. 2016. http://beactivebirmingham.co.uk/ (accessed June 22, 2017).

235 Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. *Can J Appl Sport Sci* 1985; **10**: 141–6.

236 Ainsworth B, Haskell W, Herrmann S, *et al.* 2011 Compendium of Physical Activities. *Med Sci Sport Exerc* 2011; **43**: 1575–81.

237 Department of Health. Start Active, Stay Active: A report on physical activity from the four home countries' Chief Medical Officers. London, 2011 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216370/dh_128210.pdf.

238 Lyon A, Frew EJ, Pallan M, Park S. Evaluation of the Birmingham Be Active Programme (draft). Birmingham, 2011.

239 Office for National Statistics. Population estimates for UK, England and Wales, Scotland and Northern Ireland, revised: mid-2012, mid-2013 and mid-2014. London, 2016 https://www.ons.gov.uk/releases/populationestimatesforukenglandandwalesscotlandandnorthernirelandrevisedmid2012mid2013andmid2014.

240 Sport England. Active People Survey, 2010-2011. 2013 http://dx.doi.org/10.5255/UKDA-SN-7138-3.

241 Kyu HH, Bachman VF, Alexander LT, *et al.* Physical activity and risk of breast cancer, colon cancer, diabetes, ischemic heart disease, and ischemic stroke events: systematic review and dose-response meta-analysis for the Global Burden of Disease Study 2013. *BMJ* 2016; **354**: i3857.

242 Wahid A, Manek N, Nichols M, *et al.* Quantifying the association between physical activity and cancer: a systematic review and meta-analysis. .

243 Wahid A, Manek N, Nichols M, *et al.* Quantifying the Association Between Physical Activity and Cardiovascular Disease and Diabetes: A Systematic Review and Meta-Analysis. *J Am Heart Assoc* 2016; **5**: e002495.

244 Mytton OT, Tainio M, Ogilvie D, Panter J, Cobiac L, Woodcock J. The modelled impact of increases in physical activity: the effect of both increased survival and reduced incidence of disease. *Eur J Epidemiol* 2017; **32**: 235–50.

245 Forouzanfar MH, Afshin A, Alexander LT, *et al.* Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016; **388**: 1659–724.

246 StataCorp. Stata statistical Software: Release 14. 2015.

247 Health and Social Care Information Centre. Health Survey for England 2012. London, 2013 http://content.digital.nhs.uk/catalogue/PUB13218/HSE2012-Sum-bklet.pdf.

248 Prince SA, Adamo KB, Hamel ME, Hardt J, Connor Gorber S, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act* 2008; **5**: 56.

249 Hobbs N, Godfrey A, Lara J, *et al.* Are behavioral interventions effective in increasing physical activity at 12 to 36 months in adults aged 55 to 70 years? A systematic review and meta-analysis. *BMC Med* 2013; **11**: 75.

250 Sport England. Active People Interactive. 2017. http://activepeople.sportengland.org/ (accessed June 22, 2017).

251 Moore SC, Lee I-M, Weiderpass E, *et al.* Association of Leisure-Time Physical Activity With Risk of 26 Types of Cancer in 1.44 Million Adults. *JAMA Intern Med* 2016; **176**: 816–25.

252 Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Value Health* 2012; **15**: 843–50.

253 Briggs A, Scarborough P, Smith A. Modelling in Public Health. In: Regmi K, Gee I, eds. Public Health Intelligence: Issues of Measure and Method. Springer, 2016: 67–90.

254 Vemer P, Corro Ramos I, van Voorn GAK, Al MJ, Feenstra TL. AdViSHE: A Validation-Assessment Tool of Health-Economic Models for Decision Makers and Model Users. *Pharmacoeconomics* 2016; **34**: 349–61.

255 Cobiac LJ, Tam K, Veerman L, Blakely T. Taxes and Subsidies for Improving Diet and Population Health in Australia: A Cost-Effectiveness Modelling Study. *PLOS Med* 2017; **14**: e1002232.

256 Cobiac LJ, Vos T, Veerman JL. Cost-effectiveness of interventions to reduce dietary salt intake. *Heart* 2010; **96**: 1920–5.

257 Smith-Spangler CM, Juusola JL, Enns EA, Owens DK, Garber AM. Population strategies to decrease sodium intake and the burden of cardiovascular disease: a cost-effectiveness analysis. *Ann Intern Med* 2010; **152**: 481–7, W170-3.

258 Mason H, Shoaibi A, Ghandour R, *et al.* A cost effectiveness analysis of salt reduction policies to reduce coronary heart disease in four Eastern Mediterranean countries. *PLoS One* 2014; **9**: e84445.

259 Health Economists' Study Group. Health Economists' Study Group. 2017. http://hesg.org.uk/ (accessed June 22, 2017).

260 The University of Otago. Burden of Disease Epidemiology, Equity & Cost-Effectiveness Programme (BODE^3). 2017. http://www.otago.ac.nz/wellington/departments/publichealth/research/bode3/index.html (accessed June 22, 2017).

261 Scarborough P, Bhatnagar P, Wickramasinghe K, Smolina K, Mitchell C, Rayner M. Coronary Heart Disease Statistics 2010. London, 2010 https://www.bhf.org.uk/publications/statistics/coronary-heart-disease-statistics-2010.

262 Allender S, Peto V, Scarborough P, Boxer A, Rayner M. Coronary heart disease statistics 2006 edition. London, 2006 https://www.bhf.org.uk/publications/statistics/coronary-heart-disease-statistics-book-2006.

263 Sadler K, Nicholson S, Steer T, *et al.* National Diet and Nutrition Survey - Assessment of dietary sodium levels among adults (aged 19-64) in England, 2011. London, 2012 https://www.gov.uk/government/publications/assessment-of-dietary-sodium-levels-among-adults-aged-19-64-in-england-2011.

264 Cappuccio FP, Markandu ND, Carney C, Sagnella GA, MacGregor GA. Double-blind randomised trial of modest salt restriction in older people. *Lancet* 1997; **350**: 850–4.

265 Department of Health Vascular Programme Team. Treatment of Heart Attack National Guidance. Final Report of the National Infarct Angioplasty Project (NIAP). London, 2008.

266 Yusuf S, Hawken S, Ounpuu S, *et al.* Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* 1985; **364**: 937–52.

267 Lim SS, Vos T, Flaxman AD, *et al.* A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; **380**: 2224–60.

268 Salomon JA, Haagsma JA, Davis A, *et al.* Disability weights for the Global Burden of Disease 2013 study. *Lancet Glob Heal* 2015; **3**: e712-23.

269 Mozaffarian D, Fahimi S, Singh GM, *et al.* Global Sodium Consumption and Death from Cardiovascular Causes. *N Engl J Med* 2014; **371**: 624–34.

270 He FJ, Pombo-Rodrigues S, MacGregor GA. Salt reduction in England from 2003 to 2011: its relationship to blood pressure, stroke and ischaemic heart disease mortality. *BMJ Open* 2014; **4**: e004549.

271 UK Biobank. UK Biobank. 2017. www.ukbiobank.ac.uk (accessed June 22, 2017).

272 AgMIP. The Agricultural Model Intercomparison and Improvement Project. 2014.

http://www.agmip.org/ (accessed June 22, 2017).

273 CISNET. National Cancer Institute Cancer Intervention and Surveillance Modeling Network. 2017. https://cisnet.cancer.gov/ (accessed June 22, 2017).

274 Ferguson B. Investing in prevention: the need to make the case now. Public Health Matters Blog, Public Health England. 2016. https://publichealthmatters.blog.gov.uk/2016/02/22/investing-in-prevention-the-need-to-make-the-case-now/ (accessed June 22, 2017).

275 World Health Organization. From burden to 'best buys': reducing the economic impact of non-communicable diseases in low- and middle-income countries. Geneva, 2011 http://www.who.int/nmh/publications/best_buys_summary/en/.

276 Mallender J, Bertranou E, Owen L, Lester-George A, Jhita T, Roberts S. Physical Activity Return on Investment Tool. London, 2013 https://www.nice.org.uk/about/what-we-do/into-practice/return-on-investment-tools/physical-activity-return-on-investment-tool.

277 The University of Queensland School of Population Health. Assessing Cost Effectiveness (ACE) in Prevention Study. 2016. https://public-health.uq.edu.au/research/centres/past-centres/assessing-cost-effectiveness-ace-prevention-study (accessed June 22, 2017).

278 Salomon JA, Vos T, Hogan DR, *et al.* Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. *Lancet* 2012; **380**: 2129–43.

279 Cobiac L, Vos T, Doran C, Wallace A. Cost-effectiveness of interventions to prevent alcohol-related disease and injury in Australia. *Addiction* 2009; **104**: 1646–55.

280 Australian Institute of Health and Welfare. AIHW Disease costs and impacts study data. Canberra, 2001.

281 Kelly FJ, Kelly J. London air quality: a real world experiment in progress. *Biomarkers* 2009; **14 Suppl 1**: 5–11.

282 Marsh K, Ganz ML, Hsu J, Strandberg-Larsen M, Gonzalez RP, Lund N. Expanding Health Technology Assessments to Include Effects on the Environment. *Value Health* 2016; **19**: 249–54.

283 National Institute for Health and Care Excellence. How NICE measures value for money in relation to public health interventions. Local government briefing [LGB10]. London, 2013 https://www.nice.org.uk/Media/Default/guidance/LGB10-Briefing-20150126.pdf.

284 Sen A. Capability and Well-being. In: Nussbaum M, Sen A, eds. The Quality of Life. Oxford: Clarendon Press, 1993.

285 Wilson E, Sussex J, Macleod C, Fordham R. Prioritizing health technologies in a Primary Care Trust. *J Health Serv Res Policy* 2007; **12**: 80–5.

286 Maciosek M V., Edwards NM, Coffield AB, *et al.* Priorities Among Effective Clinical Preventive Services. *Am J Prev Med* 2006; **31**: 90–6.

287 Banke-Thomas AO, Madaj B, Charles A, van den Broek N. Social Return on Investment (SROI) methodology to account for value for money of public health interventions: a systematic review. *BMC Public Health* 2015; **15**: 582.

288 The Cabinet Office. A guide to Social Return on Investment. London, 2009 https://www.bond.org.uk/data/files/Cabinet_office_A_guide_to_Social_Return_

on_Investment.pdf.

289 Murray CJL, Richards MA, Newton JN, *et al.* UK health performance: findings of the Global Burden of Disease Study 2010. *Lancet* 2013; **381**: 997–1020.

290 Townsend N, Wickramasinghe K, Bhatnagar P, *et al.* Coronary heart disease statistics 2012 edition. London, 2012 https://www.bhf.org.uk/publications/statistics/coronary-heart-disease-statistics-2012.

291 Luengo-Fernández R, Leal J, Gray A, Petersen S, Rayner M. Cost of cardiovascular diseases in the United Kingdom. *Heart* 2006; **92**: 1384–9.

292 Barton P, Andronis L, Briggs A, McPherson K, Capewell S. Effectiveness and cost effectiveness of cardiovascular disease prevention in whole populations: modelling study. *BMJ* 2011; **343**.

293 National Institute for Health and Care Excellence. NICE public health guidance 25: Prevention of cardiovascular disease. Manchester, 2010 https://www.nice.org.uk/guidance/ph25.

294 Andronis L, Barton P, Bayliss S, Hyde C. PH25 prevention of cardiovascular disease: reviews and primary studies - 4 cost effectiveness. London, 2010 https://www.nice.org.uk/guidance/ph25/evidence/reviews-and-primary-studies-4-cost-effectiveness-374845933.

295 McPherson K, Marsh T, Brown M. Foresight. Tackling Obesities: Future Choices – Modelling Future Trends in Obesity and the Impact on Health. 2nd Edition. London, 2007.

296 Scarborough P, Nnoaham K, Clarke D, Capewell S, Rayner M. Modelling the impact of a healthy diet on cardiovascular disease and cancer mortality. *J Epidemiol Community Health* 2012; **66**: 420–6.

297 Smolina K, Wright FL, Rayner M, Goldacre MJ. Incidence and 30-day case fatality for acute myocardial infarction in England in 2010: national-linked database study. *Eur J Public Health* 2012; **22**: 848–53.

298 Rothwell PM, Coull AJ, Silver LE, *et al.* Population-based study of event-rate, incidence, case fatality, and mortality for all acute vascular events in all arterial territories (Oxford Vascular Study). *Lancet* 2005; **366**: 1773–83.

299 Office for National Statistics. Mortality in England and Wales. London, 2015 https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/deathregistrationssummarytablesenglandandwalesreferencetables.

300 Lee S, Shafe ACE, Cowie MR. UK stroke incidence, mortality and cardiovascular risk management 1999-2008: time-trend analysis from the General Practice Research Database. *BMJ Open* 2011; **1**: e000269.

301 Holden SH, Barnett AH, Peters JR, *et al.* The incidence of type 2 diabetes in the United Kingdom from 1991 to 2010. *Diabetes Obes Metab* 2013; **15**: 844–52.

302 Health and Social Care Information Centre. National Diabetes Audit 2011-12. London, 2013 http://content.digital.nhs.uk/catalogue/PUB12421.

303 Ratib S, West J, Crooks CJ, Fleming KM. Diagnosis of liver cirrhosis in England, a cohort study, 1998-2009: a comparison with cancer. *Am J Gastroenterol* 2014; **109**: 190–8.

304 Office for National Statistics. Cancer Registration Statistics, England: 2012. London, 2014 https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2014-06-19.

305 Salomon JA, Murray CJL. The Epidemiologic Transition Revisited: Compositional Models for Causes of Death by Age and Sex. *Popul Dev Rev* 2002; **28**: 205–28.

306 Mathers CD, Loncar D. Projections of Global Mortality and Burden of Disease from 2002 to 2030. *PLoS Med* 2006; **3**: e442.

307 Dauchet L, Amouyel P, Hercberg S, Dallongeville J. Fruit and vegetable consumption and risk of coronary heart disease: a meta-analysis of cohort studies. *J Nutr* 2006; **136**: 2588–93.

308 Dauchet L, Amouyel P, Dallongeville J. Fruit and vegetable consumption and risk of stroke: a meta-analysis of cohort studies. *Neurology* 2005; **65**: 1193–7.

309 Vieira AR, Abar L, Vingeliene S, *et al.* Fruits, vegetables and lung cancer risk: a systematic review and meta-analysis. *Ann Oncol* 2016; **27**: 81–96.

310 Threapleton DE, Greenwood DC, Evans CEL, *et al.* Dietary fibre intake and risk of cardiovascular disease: systematic review and meta-analysis. *BMJ* 2013; **347**: f6879.

311 Aune D, Chan DSM, Greenwood DC, *et al.* Dietary fiber and breast cancer risk: a systematic review and meta-analysis of prospective studies. *Ann Oncol* 2012; **23**: 1394–402.

312 Aune D, Chan DSM, Lau R, *et al.* Dietary fibre, whole grains, and risk of colorectal cancer: systematic review and dose-response meta-analysis of prospective studies. *Br Med J* 2011; **343**: d6617.

313 Zhang Z, Xu G, Ma M, Yang J, Liu X. Dietary fiber intake reduces risk for gastric cancer: a meta-analysis. *Gastroenterology* 2013; **145**: 113–120.e3.

314 Aune D, Chan DSM, Vieira AR, *et al.* Red and processed meat intake and risk of colorectal adenomas: a systematic review and meta-analysis of epidemiological studies. *Cancer Causes Control* 2013; **24**: 611–27.

315 Song P, Lu M, Yin Q, *et al.* Red meat consumption and stomach cancer risk: a meta-analysis. *J Cancer Res Clin Oncol* 2014; **140**: 979–92.

316 Aune D, Ursin G, Veierød MB. Meat consumption and the risk of type 2 diabetes: a systematic review and meta-analysis of cohort studies. *Diabetologia* 2009; **52**: 2277–87.

317 Prospective Studies Collaboration, Lewington S, Whitlock G, *et al.* Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. *Lancet* 2007; **370**: 1829–39.

318 Prospective Studies Collaboration, Whitlock G, Lewington S, *et al.* Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet* 2009; **373**: 1083–96.

319 Norat T, Aune D, Vieira AR, Chan D, Navarro Rosenblatt D, Vieira R. The Associations between Food, Nutrition and Physical Activity and the Risk of Pancreatic Cancer. WCRF/AICR Systematic Literature Review Continuous Update Project Report. London, 2011 http://www.wcrf.org/sites/default/files/SLR_pancreatic_cancer_2011.pdf.

320 Clarke R, Frost C, Collins R, Appleby F, Peto R. Dietary lipids and blood cholesterol: quantitative meta-analysis of metabolic ward studies. *BMJ* 1997; **314**: 112–7.

321 Te Morenga LA, Howatson AJ, Jones RM, Mann J. Dietary sugars and cardiometabolic risk: systematic review and meta-analyses of randomized controlled trials of the effects on blood pressure and lipids. *Am J Clin Nutr* 2014; **100**: 65–79.

322 Christiansen E, Garby L. Prediction of body weight changes caused by changes in energy balance. *Eur J Clin Invest* 2002; **32**: 826–30.

323 Peters SAE, Huxley RR, Woodward M. 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**: 1542–51.

324 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. *Lancet* 2014; **383**: 1973–80.

325 Danaei G, Ding EL, Mozaffarian D, *et al.* The preventable causes of death in the United States: comparative risk assessment of dietary, lifestyle, and metabolic risk factors. *PLoS Med* 2009; **6**: e1000058.

326 GBD 2013 Risk Factors Collaborators, Forouzanfar MH, Alexander L, *et al.* Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015; **386**: 2287–323.

327 University of California B (USA), Max Planck Institute for Demographic Research (Germany). Human Mortality Database. http://www.mortality.org/ (accessed June 22, 2017).

328 Peasgood T, Ward SE, Brazier J. Health-state utility values in breast cancer. *Expert Rev Pharmacoecon Outcomes Res* 2010; **10**: 553–66.

Appendices

Appendix 1 Protocol for stakeholder involvement

Protocol for stakeholder involvement

Version 1.2. 9th April, 2015

Title

Estimating and comparing the cost effectiveness of primary prevention policies affecting diet and physical activity in England

DPhil student researcher

Dr Adam Briggs

BHF Centre for Population Approaches on Non-Communicable Disease Prevention,

Nuffield Department of Population Health,

Old Road Campus,

Headington,

Oxford.

OX3 7LF

+44 (0)1865 617785

adam.briggs@dph.ox.ac.uk

Supervisors

Dr Peter Scarborough

BHF Centre for Population Approaches on Non-Communicable Disease Prevention,

Nuffield Department of Population Health,

Old Road Campus,

Headington,

Oxford.

OX3 7LF

+44 (0)1865 289248

peter.scarborough@dph.ox.ac.uk

Dr Jane Wolstenholme

Health Economics Research Centre,

Nuffield Department of Population Health,

Old Road Campus,

Headington,

Oxford.

OX3 7LF

+44 (0)1865 289269

jane.wolstenholme@dph.ox.ac.uk

Funders

Wellcome Trust Research Training Fellowship

Nuffield Department of Population Health, University of Oxford

Lay summary

In England, there are many government policies aimed at preventing heart disease and stroke. These range from national campaigns such as "5-A-DAY" which promotes increasing fruit and vegetable consumption, to individuals receiving prescriptions to visit the gym. In

an increasingly resource constrained health service, informing policy makers of how the costs and benefits of different health policies compare is key to making better decisions.

This research project will compare the impact on major diseases such as heart disease and cancer in England, and the associated intervention and healthcare costs and savings, of policies that affect diet and physical activity. The project may also be extended to include prescription drugs. To do this, a computer model will be developed. The results will estimate the health impact of different policies for each gender, for different age groups, and for different income groups. These outcomes will be investigated over time meaning that it will be possible to estimate the impact of the policy five, ten, or 20 years following its implementation. The model will be designed so that in the future it can be further developed to assess a wider range of policies and disease outcomes, as well as be used directly by decision makers.

Scientific summary

Importance of the research

CHD and stroke are the major causes of mortality in the UK[289,290] with diet and physical inactivity contributing to more disability adjusted life years (DALYs) in 2010 than any other risk factor.[289] CHD and stroke are estimated to cost the UK in excess of £30 billion per year[291] and it has been suggested that prevention strategies are likely to be cost-effective.[10,292]

The National Institute for Health and Care Excellence (NICE) has outlined in a public health guidance document how premature deaths from CHD and stroke can be prevented through changing population diet and physical activity.[293] However, the estimates of cost-effectiveness of the policies assessed in the guidance are from a range of sources using different baseline datasets making it not possible for decision makers to directly compare

the policies.[294] Current models investigating this area have problems such as not using UK specific relative risk estimates[117,118] and not accounting for historical disease trends.[117,118,292,295] Internationally, more sophisticated country specific models are being developed; for example, the Assessing Cost Effectiveness (ACE) Prevention study in Australia,[277] and by the BODE[3] research group in New Zealand.[260]

Aims and objectives

This project aims to compare the cost effectiveness of primary coronary heart disease (CHD) and stroke prevention policies targeting diet and physical activity in England.

A novel computer model will be developed to assess the comparative cost effectiveness of primary prevention policies in England. The model will estimate the impact of policies affecting diet and physical activity on CHD and stroke morbidity and mortality by age, sex, and socio-economic group, and estimate their cost-effectiveness. Initially, the specific policies which will be investigated have been identified using an approach similar to the ACE Prevention study. The selected policies are:

1. Free access to leisure centres and swimming pools
2. Food producers reformulating food to be healthier (reducing salt, sugar, and unhealthy fat content)

The primary outcomes for each intervention are:

1. Change in number of CHD and stroke events
2. Change in number of quality-adjusted life years (QALYs)
3. Cost-effectiveness (£ per QALY gained)

The model will be designed so that following completion of this project, it can be further developed into an increasingly sophisticated interactive tool to assess more interventions (for example changing drug prescribing practices) for a wider range of disease outcomes (such as including cancer).

Methods

Phase 1

A systematic review of the literature will identify which models have been previously used to estimate the health impact and cost effectiveness of primary disease prevention policies at the population and the individual level. This will inform the ideal design of the England specific longitudinal model that, as with the ACE Prevention work and BODE^3, is anticipated to be based on Markov decision modelling and microsimulation. Markov modelling allows for the assessment of incremental cost-effectiveness ratios by simulating a population travelling through life with each member existing in a single disease state, whilst microsimulation allows for quantification of the stochastic uncertainty associated with effect-size estimates.

The model will use state transition parameters to predict the probability of an individual of a given age, sex, and socio-economic group moving into a new disease state from their current state. Disease states modelled will be incident CHD and stroke.

Stakeholders (for the purpose of this protocol, these are the potential users of the model – healthcare and public health decision makers) and will be consulted at this stage to inform the model's structure, disease inputs, and model outputs and maximise its usefulness for decision making

Phase 2

State transition parameters will be incidence and case fatality rates by age, sex, and socio-economic group, with trend data from linked Hospital Episode Statistics and death certificates, held by the Nuffield Department of Population Health, Oxford, being an ideal source. These data will be used to project future CHD and stroke incidence and case fatality rates based on previous trends, and to derive the state transition parameters.

Phase 3

Each disease state has an associated utility and cost, and each intervention has a cost, identifiable from the Cost-Effectiveness Analysis (CEA) registry held by Tufts Medical Centre and the academic literature for utility values, and using bottom-up assessments of healthcare spending using Personal Social Services Research Unit (PSSRU) unit cost information, NHS reference costs, and the British National Formulary. Utilities are readily converted to quality adjusted life years (QALYs) by multiplying the number of years each person spends in a given disease state by the associated utility. Total costs can be calculated in the same way. Future costs and utilities will be discounted at a rate of 3.5% per year (as recommended by NICE).

At this stage stakeholders will again be contacted to offer feedback on model validity, structure, and outputs.

Phase 4

In order to calculate costs and QALYs associated with the counterfactual scenarios (post policy interventions), the state-transition parameters will change based on the effect of the policy being modelled. A likely method is using population impact fractions (PIFs) derived from a comparative risk assessment model called PRIME (previously called DIETRON), developed by Dr Scarborough[296] and can already be used to derive PIFs from food and physical activity.

For proposed interventions, the relationship between the intervention and the change in diet and physical activity levels will be parameterised following a review of the evidence from trials or observational studies, using age, sex, and socio-economic group specific parameters if available. Costs and QALYs associated with baseline and counterfactual scenarios will be used to calculate incremental cost-effectiveness ratios (ICERS) of the different interventions.

Phase 5

Both probabilistic and deterministic sensitivity analyses will be conducted. Sensitivity analyses will assess the sensitivity of the results to key modelling assumptions such as the discount rate, unit costs, utilities, and effectiveness parameters. Internal validation will be achieved by modelling extreme scenarios and identifying unrealistic results. The model will be externally validated by comparison of the results with prospective cohort data (including submitting a proposal to the UK BioBank for access to diet and health outcome data), and with relevant trial data.

The model will be built using Microsoft Excel and will be developed to conform to the International Society for Pharmcoeconomics and Outcomes Research's (ISPOR) good research practices for state-transition modelling,[12] and NICE guidelines for assessing cost effectiveness.[13]

Stakeholders will be contacted again at this stage to inform how the model can be best presented and disseminated to inform decision making.

Rationale for stakeholder involvement

In order to make the computer model as useful and interpretable to the end user as possible, it is important to ensure that decision makers in health and public health are consulted early in the model's development. These stakeholders will be asked to give feedback on the face validity of the model – do the steps in the model make sense and are there any crucial

steps that are missing. They will also be asked whether the disease processes modelled and the outcome produced are useful or what could be done to improve these. Finally, they will be asked if they would be likely to use the model or its results, and if not, why not. This feedback will directly influence the model's development.

Participant recruitment

Participant inclusion criteria

Local and national decision makers in healthcare or public health, chief executives of relevant major charitable organisations, leaders of national medical and public health professional bodies. (See appendix A1.1)

Exclusion criteria

If a participant leaves their job during the study period, the participant will be excluded from the study with data collected up to that point retained.

Method of recruitment

Relevant stakeholders will be identified using a stakeholder analysis (see appendix 1.1). Professional email addresses of potential participants will be identified through internet searches. Participants will be emailed an invitation to participate, with patient information form attached and a link to online consent form and study questionnaire, at three time points during the study: before the model is developed (t=0), once during model development (t=1 year), and at the end of the model's development (t=2 years). Depending on the progress of the project, participants will have between two and four weeks to complete the questionnaire. Between 20 and 30 participants are expected to be invited.

Training of research staff

Dr Adam Briggs will have up to date Good Clinical Practice training and will be the only researcher inviting participants to take part in the study.

Information provided to participants

Information provided can be seen on the participant information sheet (appendix A1.2). It will include the minimum requried by the CUREC guidance on participant information (www.admin.ox.ac.uk/curec/guidance/index.shtml).

Consent of participants

Informed consent will be obtained through the checking of boxes asking the relevant participant declarations online before proceding to the questionnaire (see appendix A1.3 for consent form). Participants will be given a minimum of two weeks and a maximum of four weeks from receipt of the invitation email and the participant information sheet to complete the online questionnaire.

Reward to participants

There will be no reward or incentive offered to participants to take part in the study.

Potential risks to participants

There are no direct risks involved with taking part in the study.

Data collection and use of data

Participants will be invited to complete up to three online questionnaires, each taking between ten and twenty minutes to complete, over the course of two years.

At t=0 participants will be asked about what model outcomes they would find most useful. This includes the population groups, disease processes, and time frames modelled. Participants will also be asked whether they think that the steps within the model make

sense - are they valid representations of real life? (see appendix 1.4 for sample text in invitation email, and appendix 1.5 for draft of the questionnaire at T=0).

Participant responses will be divided into one of five categories:

i. National governmental organisation (i.e. NICE, DoH, National Government, PHE)
ii. Local governmental organisation (i.e. Local Government Association, Public Health England Centre, Local Authority/Council)
iii. Charitable organisation
iv. Professional organisation (i.e. AoMRC, RCP, FPH, ADPH)
v. Other (please state)

Respondents from each category will be equally weighted, with the top three ranked responses to questions two, three, and four (questions concerning what outputs, disease processes, and time frame the model should include) either included in the model or specifically justified to stakeholders and within the thesis as to why they are not included.

Question five is more likely to be included at t=2 rather than t=0. The responses will inform how results of the model are communicated - if there is a demand for using the best available evidence (irrespective of quality) then this may lead to the reporting of two different model outcomes. Each with either greater or fewer disease processes included. This will also aid assessment of the models' structural validity.

Each response to question six, concerning the model's face validity, will be individually addressed.

The most frequent response to question seven will dictate how uncertainty will be presented, and this may also result in uncertainty being presented in multiple ways in order to be of value to different audiences.

The responses to whether participants would use the model and how it could be made more useful will directly inform the model's development and presentation of results.

At t=1 year participants will be contacted to inform them of how the model is developing and to share information on where questionnaire responses made at t=0 have not been possible to include in the model due to problems such as lack of data availability or technical difficulty. A second questionnaire will be used to get feedback on the validity of the model, including whether it is something that the participants would find useful.

How responses will be handled will depend on the wording of the questions, but broadly speaking responses will be handled in the same way as at t=0.

At t=2 years (the project's completion) a shorter questionnaire will be sent to participants to assess how valid the particpants think the model is, whether they would use it or its results, and what could be done to make it more useful. These data will be used to finalise the model's presentation of results and to guide future development of the model.

Participants will be free to withdraw from the study at any time without penalty by informing the study team with anonymised data collected up to that point retained.

Data security and confidentiality

The questionnaire will be operated using Bristol Online Surveys and each questionnaire will be available to participants for between two and four weeks (depending on the stage of the overall project), the data collection period. Participant identifiable responses will therefore be available to the study team by accessing Bristol Online Surveys during the data collection period after which anonymised responses will be downloaded and stored on a password protected computer drive, with identifiable responses deleted from Bristol Online Surveys.

The identity of each participant is collected in the online questionnaire in order to know whether the participant wishes to remain in the study or not. This single piece of information will be collected in a separate file to the data collected from the survey, with no linked identifier between the two files. Therefore no participant identifiable data will be stored in this study.

Dissemination plan

Results from the questionnaires will be used to inform the development of the model, and anonymised responses will be published both in the thesis and in relevant academic publications describing the model's development.

Ethics approval

Ethics approval has been granted from the Central University Research Ethics Committee, University of Oxford, reference number: MS-IDREC-C1-2015-052.

Appendices to appendix 1

Appendix 1.1. Stakeholder analysis

List of stakeholders

Manage closely

- National government
- Local Government Association
- Department of Health
- Public Health England
- National Institute for Health and Care Excellence (NICE)
- Responsibility Deal
- Association of Directors of Public Health
- Faculty of Public Health

Keep informed

- Wellcome Trust*
- Academy of Medical Royal Colleges
- British Heart Foundation

Keep satisfied

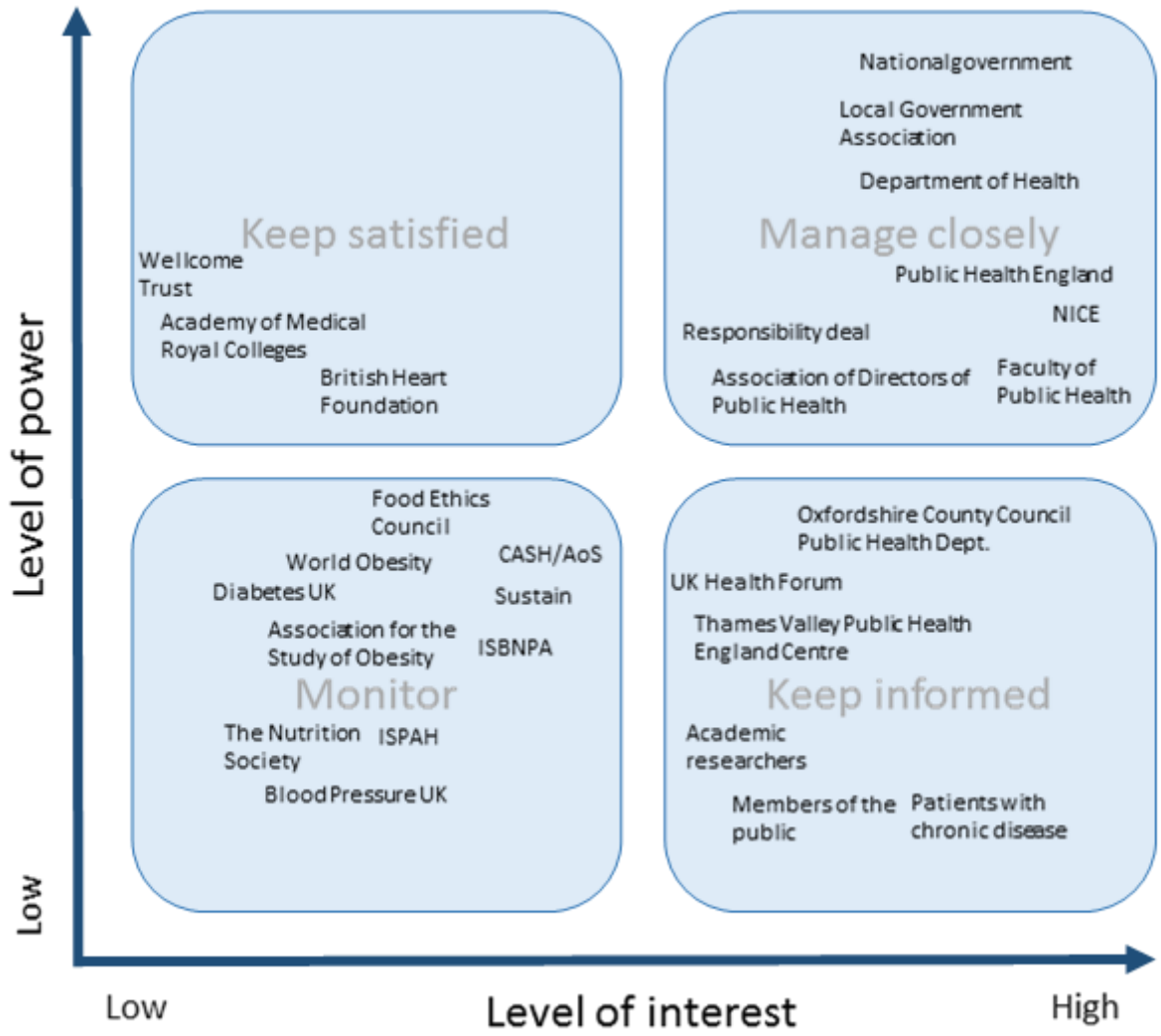
- Oxford County Council Public Health Department
- UK Health Forum
- Thames Valley Public Health England Centre
- Academic Researchers*
- Members of the public*
- Patients with chronic disease*

Monitor

- Food Ethics Council
- Consensus Action on Salt and Health/Action on Sugar (CASH/AoS)
- World Obesity
- International Society of Behavioural Nutrition and Physical Activity (ISBNPA)
- International Society for Physical Activity and Health (ISPAH)
- Diabetes UK
- Sustain
- Association for the Study of Obesity
- The Nutrition Society
- Blood Pressure UK

*Will not be contacted as part of this protocol.

Stakeholder analysis



Appendix 1.2. Participant information sheet

Estimating and comparing the cost effectiveness of primary prevention policies affecting diet and physical activity in England

Participant Information Sheet

V1.2 09/04/15

(to be emailed to the participants and to be available online)

Who is the study team?

This study is a doctorate research project being led by Dr Adam Briggs, a DPhil candidate funded by the Wellcome Trust at the University of Oxford, who can be contacted using the details below:

BHF Centre for Population Approaches on Non-Communicable Disease Prevention,

Nuffield Department of Population Health,

University of Oxford,

New Richards Building,

Old Road Campus,

Headington,

Oxford.

OX3 7LF

01865 617785

adam.briggs@dph.ox.ac.uk.

The supervisors are Dr Peter Scarborough and Dr Jane Wolstenholme, both senior researchers at the Nuffield Department of Population Health, University of Oxford.

What is the study aiming to do?

In England there are many government policies aimed at preventing heart disease and stroke. These range from national campaigns such as "5-A-DAY" which promotes increasing fruit and vegetable consumption, to individuals receiving prescriptions to visit the gym. In an increasingly resource constrained health service, informing policy makers of how the costs and benefits of different health policies compare is key to making better decisions.

This research project will compare the impact of policies that affect diet and physical activity on major diseases such as heart disease and cancer in England, and the associated implementation and healthcare costs and savings. The project may also be extended to include prescription drugs. To do this, a computer model will be developed. The results will estimate the health impact of different policies for each gender, for different age groups, and for different income groups.

Why am I being invited to take part?

You are being invited because you have a major role within a health related organisation. In order to ensure that the computer model is relevant to decision makers and potential users of the model such as you, we are seeking your view on what outcomes you would find most useful. This includes the population groups, disease processes, and time frames that

are modelled. We would also value your opinion on whether the steps within the model make sense, and whether it would be a useful model in your day-to-day work.

What will the study involve?

The study involves answering up to three on-line questionnaires over two years. Each questionnaire will take between ten and twenty minutes to complete.

Does it matter if I move job or organisations during the study?

If your job changes during the study then we will stop contacting you; we will retain and use any data collected up to that point.

Can I ask questions before I agree to take part?

Yes, you can contact Dr Adam Briggs using the phone number, email address, or postal address above to ask any questions before agreeing to take part.

Can I choose not to take part and leave the study at any time?

Yes, you can choose not to take part, and can withdraw from the study at any time without penalty by advising us of your decision or by not responding to requests to complete the online questionnaire within the relevant time frame (two to four weeks from receipt of the invitation email). We will retain and use any anonymised data collected to that point (meaning any personally identifying comments have been removed).

Who has reviewed this project?

This project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee – reference number MS-IDREC-C1-2015-052.

Who will have access to the data and where will it be stored?

The data will be collected through the online questionnaire powered by Bristol Online Surveys, to which the research team will have access whilst the questionnaire is open (the data collection period). Anonymous responses will be downloaded and stored electronically on a secure password protected computer drive at the end of each data collection period (two to four weeks from receipt of invitation). Participant identifiable responses on Bristol Online Surveys will be deleted at the end of each data collection period.

How will the data be used?

The information collected will be used to help develop the computer model.

The University of Oxford is committed to the dissemination of its research for the benefit of society and the economy and, in support of this commitment, has established an online archive of research materials. This archive includes digital copies of student theses successfully submitted as part of a University of Oxford postgraduate degree programme. Holding the archive online gives easy access for researchers to the full text of freely available theses, thereby increasing the likely impact and use of that research.

If you agree to participate in this project, the research, including anonymised responses (meaning any personally identifying comments have been removed) will be written up as a thesis. On successful submission of the thesis, it will be deposited both in print and online in the University archives, to facilitate its use in future research. The thesis will be published with open access, meaning it is available to every internet user. Anonymised responses will also be published as open access files in the scientific literature when reporting on the development of the model.

What will happen to the data at the end of the project?

No individually identifiable data will be stored. Individually identifiable online responses will be deleted at the end of each data collection period (the two to four weeks that the questionnaire is open). Anonymised data will be made publicly available in the thesis and academic publications arising from this work.

How will you benefit from taking part?

Your input to the model's development will hopefully lead to a more relevant and useful model that in the future will help with policy decision making and lead to better health outcomes.

What are the risks involved?

There are no direct risks involved with participating.

How do I raise concerns or make a complaint?

If you have a concern about any aspect of this project, please speak to Dr Adam Briggs' supervisor (Dr Peter Scarborough, peter.scarborough@dph.ox.ac.uk, 01865 289248) who will do his best to answer your query. Dr Scarborough should acknowledge your concern within 10 working days and give you an indication of how he intends to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the chair of the Research Ethics Committee at the University of Oxford (Chair, Medical Sciences Inter-Divisional Research Ethics Committee; Email: ethics@medsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD). The chair will seek to resolve the matter in a reasonably expeditious manner.

Appendix 1.3. Consent form

Estimating and comparing the cost effectiveness of primary prevention policies affecting diet and physical activity in England

Consent form

V1.2 09/04/15

(to be available online)

Who is the study team?

This study is a doctorate research project being led by Dr Adam Briggs, a DPhil candidate funded by the Wellcome Trust at the University of Oxford, who can be contacted using the details on the participant information sheet. The supervisors are Dr Peter Scarborough and Dr Jane Wolstenholme, both senior researchers at the Nuffield Department of Population Health, University of Oxford.

What is the study aiming to do?

In England, there are many government policies aimed at preventing heart disease and stroke. These range from national campaigns such as "5-A-DAY" which promotes increasing fruit and vegetable consumption, to individuals receiving prescriptions to visit the gym. In an increasingly resource constrained health service, informing policy makers of how the costs and benefits of different health policies compare is key to making better decisions.

This research project will compare the impact of policies that affect diet and physical activity on major diseases such as heart disease and cancer in England, and the associated

implementation and healthcare costs and savings. It may also be extended to include prescription drugs. To do this, a computer model will be developed. The results will estimate the health impact of different policies for each gender, for different age groups, and for different income groups. You are being invited to take part in order to help make the model as relevant as possible.

Participant declarations

Have you read the participant information sheet, V1.2 09/04/15?

Have you had the opportunity to ask questions about the study and received satisfactory answers to questions, and any additional details requested?

Do you understand that you may withdraw from the study without penalty at any time by advising the researchers of this decision?

Do you understands that this project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee?

Do you understand who will have access to personal data provided, how the data will be stored; and what will happen to the data at the end of the project?

Do you understand that how anonymised personal data collected will be published and stored both as a thesis and in the academic literature?

Do you agree to participate in this study?

Do you understand how to raise a concern and make a complaint?

Appendix 1.4. Invitation email

Dear x,

I am studying for a doctorate funded by the Wellcome Trust at the Nuffield Department of Population Health, University of Oxford. My research aims to estimate how cost-effective different public health policies are in England. To do this, I am developing a computer model to allow different policies affecting diet, physical activity, and the prescribing of medications to be compared.

As a decision maker in health and a potential user of this model, I want to make sure it is as relevant to you as possible. With your consent, I would really appreciate your input at the development stage by answering a short on-line questionnaire which should take between 10 and 20 minutes to complete.

Attached is the participant information sheet. The same information can be found using the link below, once you have read this, you will be asked for your consent to take part before you access the questionnaire. Finally you will be asked if you wish to be contacted again regarding this study in the future.

https://www.survey.bris.ac.uk/oxford/comparingpolicies

The questionnaire is open until the 1st May.

If you have any questions or want to know about how I plan to use this information then please see the patient information sheet using the link above or get in touch by email: adam.briggs@dph.ox.ac.uk, or by phone: 01865 617785. This study has research ethic approval from the University of Oxford Central University Research Ethics Committee, number MS-IDREC-C1-2015-052.

Many thanks,

Adam

Dr Adam Briggs

Wellcome Trust Research Training Fellow

BHF Centre for Population Approaches on Non-Communicable Disease Prevention,

Nuffield Department of Population Health,

University of Oxford,

New Richards Building,

Old Road Campus,

Headington,

Oxford.

OX3 7LF

01865 617785

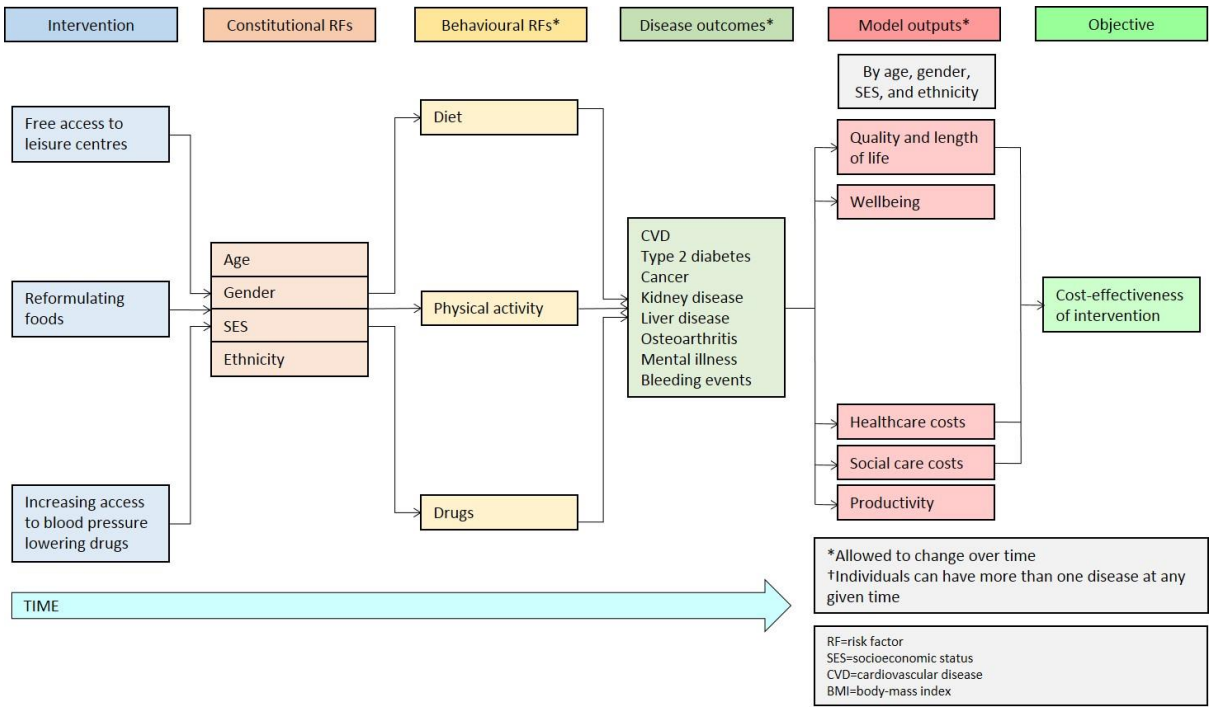
adam.briggs@dph.ox.ac.uk

Appendix 1.5. Draft of questionnaire, T=0

Draft of proposed questions for initial questionnaire:

1. Which group best describes your organisation?
    a. National governmental organisation (i.e. NICE, DoH, National Government, PHE)
    b. Local governmental organisation (i.e. Local Government Association, Public Health England Centre, Local Authority/Council)
    c. Charitable organisation
    d. Professional organisation (i.e. AoMRC, RCP, FPH, ADPH)
    e. Other (please state)

This is a simplified diagram of the model I plan to build. Policies affecting diet, physical activity, or the use of preventative medication affect people differently depending on their age, gender, socio-economic status, and ethnicity. The change in these risk factors then leads to changes to disease outcomes in England and the associated costs to the health service, social service, and the individual.



2. The outputs of the model will be:
    - change in disease burden,
    - change in deaths from disease, and
    - change in health sector, social sector, and personal costs.

    These can be used to calculate cost-effectiveness. What outputs would be most useful to you (please rank)?

    a. Cost-effectiveness from the health service perspective
    b. Cost-effectiveness from the health and social care perspective
    c. Change in costs to healthcare
    d. Change in costs to social care
    e. Change in costs to individual (loss of earnings, loss of productivity, transport costs etc).
    f. Outcomes reported by age group
    g. Outcomes reported by socio-economic group
    h. Outcomes reported by ethnicity
    i. Outcomes reported by gender
    j. Change in measure of quality and length of life
    k. Change in measure of population wellbeing
    l. Change in disease burden and deaths reported by disease
    m. Other (please specify)
3. What disease processes are most relevant to you (please rank)?
    a. Heart disease
    b. Stroke
    c. Cancer
    d. Mental health
    e. Musculoskeletal disease
    f. Kidney disease

g. Liver disease
h. Diabetes
i. Other (please specify)

4. Over what time frame would you be most interested in the results being presented (please rank)?
    a. 1 year
    b. 5 years
    c. 10 years
    d. 20 years
    e. 50 years
    f. Lifetime of the current population
    g. Other (please specify)
5. (Likely to be included at t=2 rather than t=0) Within the model, the degree to which changing a risk factor leads to a change in a disease outcome is derived from empirical research. The quality of research available varies, and if only very high quality evidence is used, the model will more accurately represent outcomes but for fewer diseases than if lower quality evidence (or best available evidence) is used. In this situation, more disease outcomes can be modelled but with less confidence in the results.

    What standard of evidence is acceptable to you for the model to use?
    a. Gold-standard evidence (i.e. meta-analyses of randomised controlled trials or cohort studies), thereby restricting the number of disease outcomes to where the evidence is strong.
    b. The best evidence available, where evidence exists. This allows for more outcomes but with more uncertainty.
    c. Other
        i. If other, please specify

6. If you were presented with the model diagram above and the results of the comparison of different public health policies, does the logic of the model's flow seem valid and plausible to you?
    a. Yes
    b. No
        i. If no, what aspect is not believable and what should be added or changed?
7. Is it important to you to have a measure of the uncertainty in the main results being presented?
    a. Yes
    b. No
        i. If yes, how do you prefer uncertainty to be communicated?
            1. 95% confidence intervals
            2. Best and worst case scenarios
            3. Range of possible scenario outcomes under different assumptions
            4. Other (please specify)
8. Is this model, or its results, something you would potentially use?
    a. Yes
        i. If yes, what would you use the model or its results for?
    b. No
        i. If no, what could be done to make it more useful to you?
9. Would you be happy to be contacted in the future with information on how the project is progressing and to be asked to provide optional additional input (it is anticipated this will be on two further occasions over the coming two years)?
    a. Yes
    b. No

i. Please provide your name so that I can either maintain you or remove you from the list of participants based on your answer above (your responses will remain anonymous either way)

10. Any other comments or suggestions?

Appendix 2 Central University Research Ethics Committee approval letter

MS IDREC
Research Services, University of Oxford

University Offices, Wellington Square, Oxford, OX1 2JD
Tel: +44(0)1865 616577 Fax: +44(0)1865 280467
ethics@medsci.ox.ac.uk http://www.admin.ox.ac.uk/curec/

UNIVERSITY OF OXFORD

CONFIDENTIAL

Ref: MS-IDREC-C1-2015-052

Dr Adam Briggs
Nuffield Department of Population Health
University of Oxford
Old Road Campus
Oxford

9th April 2015

Dear Dr Briggs

CUREC checklist

I am writing to acknowledge receipt of your CUREC-1 form for your project: **Estimating and comparing the cost effectiveness of primary prevention policies affecting diet and physical activity in England**

On the basis of the information you have provided this has now been approved by the Medical Sciences IDREC **subject to:**

a) **your following the BPS guidelines for online research;**

b) **it is your responsibility to comply with the requirements for administering any tests or questionnaires and if in doubt to contact the publisher of those tests or questionnaires.**

The reference number for this project is **MS-IDREC-C1-2015-052** and is valid for a period of **36 months** from the CUREC 1 approval date, **30th March 2015**. Please may I remind you that your project may be reviewed at some stage during an annual audit of projects.

Amendments

Should you at some stage alter some of the techniques or procedures then you should first undertake a checklist (CUREC-1) to see whether these changes alter the ethics of the research. If these remain the same then the committee will require notification of the changes to lodge with the project. If they do not remain the same then you may need to complete a CUREC-2 form and undergo further scrutiny by the committee.

Please do not hesitate to contact me if you have any queries about this.

Yours sincerely

Gill Halstead

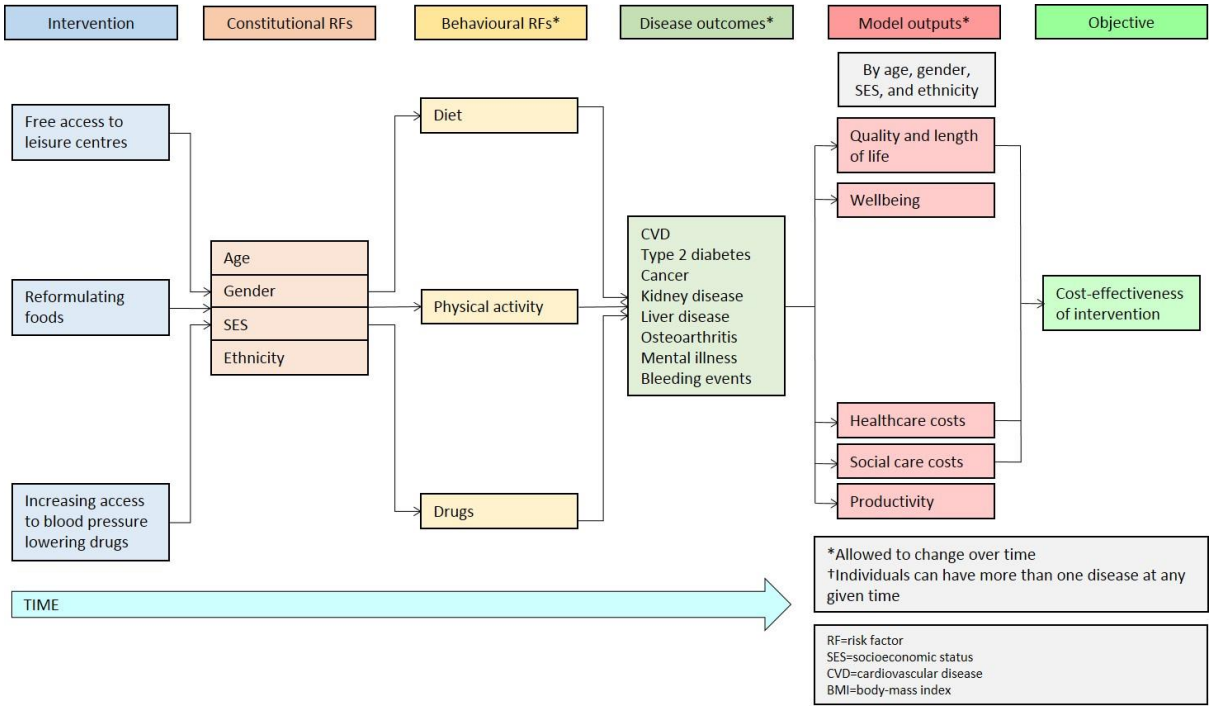
Gill Halstead
Research Ethics Co-ordinator, Medical Sciences

Appendix 3 Stakeholder questionnaire and summary of responses

Below is a copy of the questionnaire and a summary of stakeholder responses. Twelve of 29 respondents completed the questionnaire, three of the 12 did not complete every question. Two respondents answered all but one question, and one respondent did not answer four questions but provided extensive free text responses.

Questions one to eight are the consent questions, question nine asks the respondent to classify their organisation into one of five categories. Free-text responses have been removed in order to ensure respondents' confidentiality.

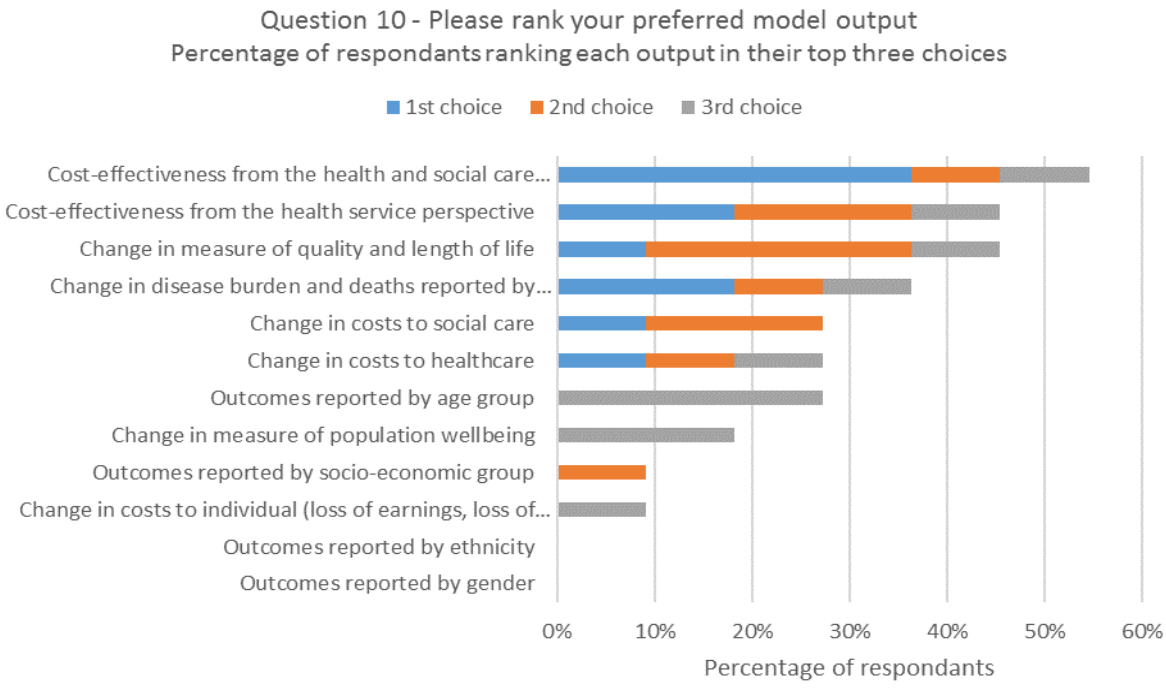
This is a simplified diagram of the model I plan to build. Policies affecting diet, physical activity, or the use of preventative medication affect people differently depending on their age, gender, socio-economic status, and ethnicity. The change in these risk factors then leads to changes to disease outcomes in England and the associated costs to the health service, social service, and the individual.



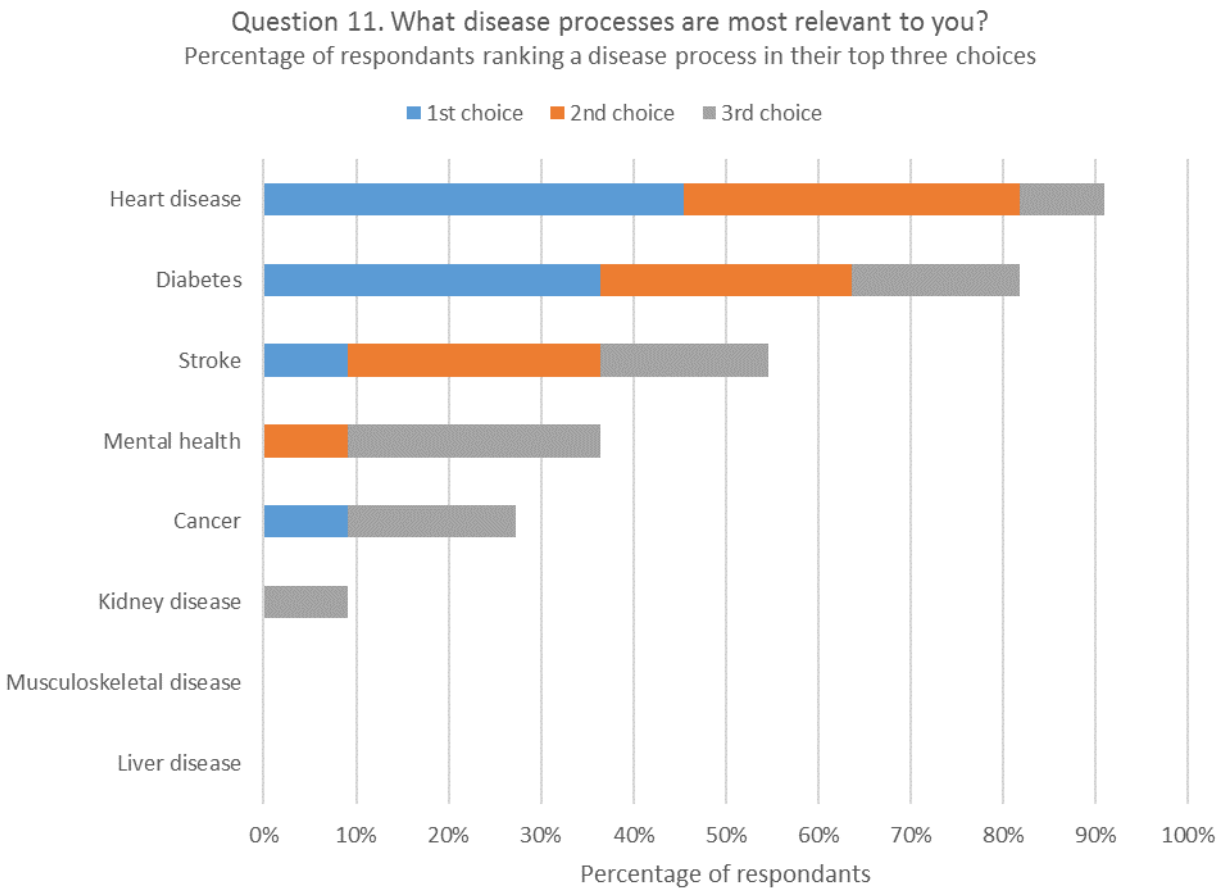
10. The outputs of the model will be:
    - change in disease burden,
    - change in deaths from disease, and
    - change in health sector, social sector, and personal costs.

    These can be used to calculate cost-effectiveness. What outputs would be most useful to you (please rank)?

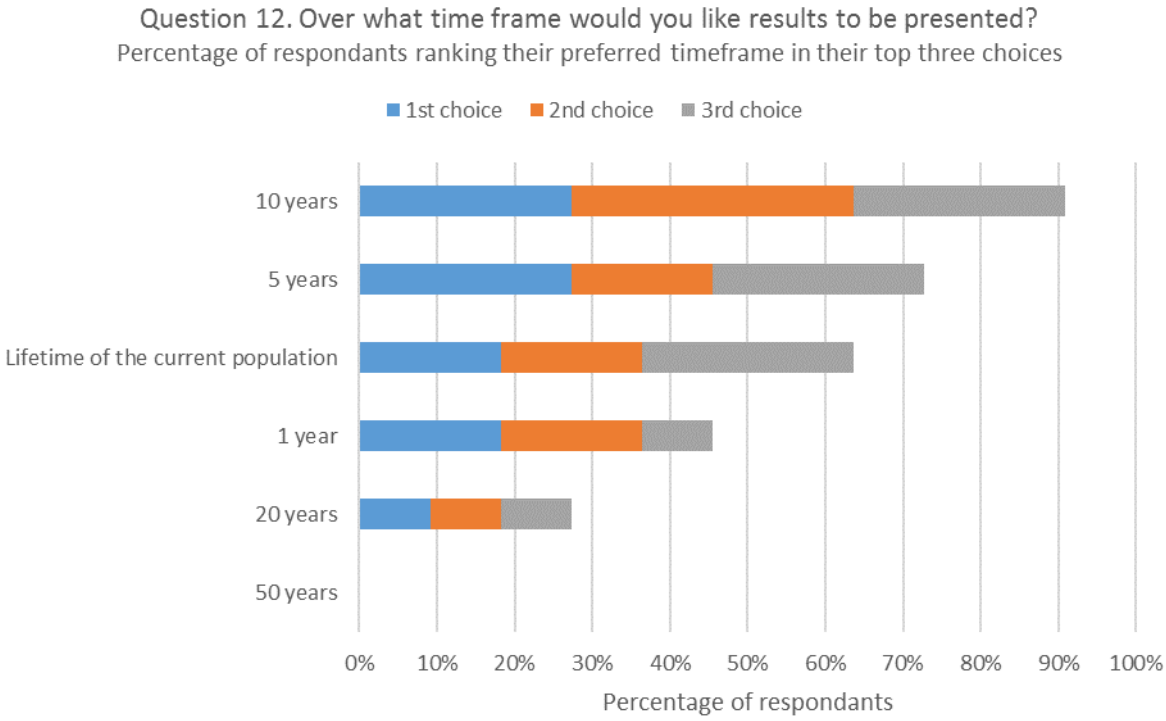
    a. Cost-effectiveness from the health service perspective
    b. Cost-effectiveness from the health and social care perspective
    c. Change in costs to healthcare
    d. Change in costs to social care
    e. Change in costs to individual (loss of earnings, loss of productivity, transport costs etc).
    f. Outcomes reported by age group
    g. Outcomes reported by socio-economic group
    h. Outcomes reported by ethnicity
    i. Outcomes reported by gender
    j. Change in measure of quality and length of life
    k. Change in measure of population wellbeing
    l. Change in disease burden and deaths reported by disease
    m. Other (please specify)



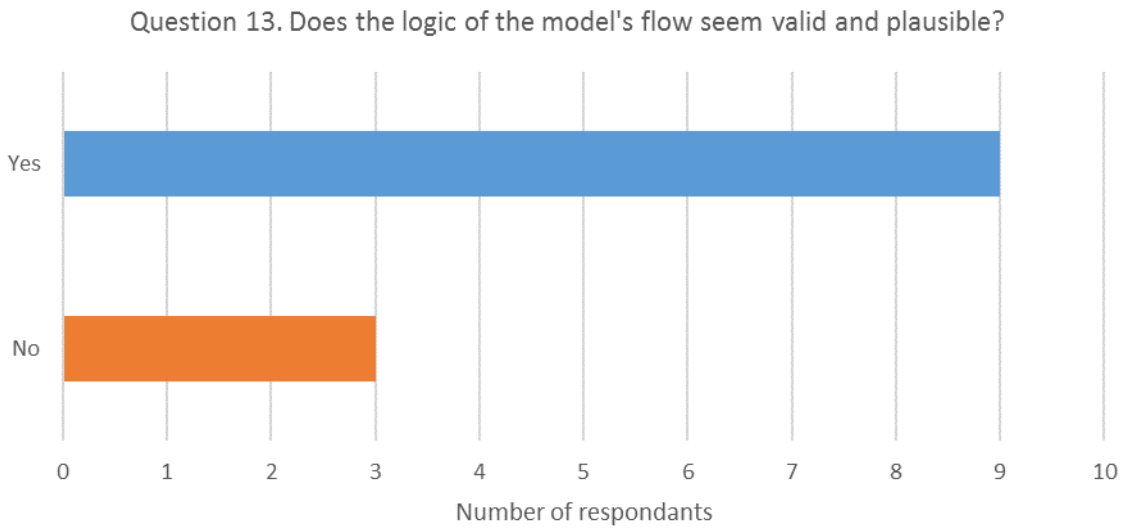
11. What disease processes are most relevant to you (please rank)?
    a. Heart disease
    b. Stroke
    c. Cancer
    d. Mental health
    e. Musculoskeletal disease
    f. Kidney disease
    g. Liver disease
    h. Diabetes
    i. Other (please specify)



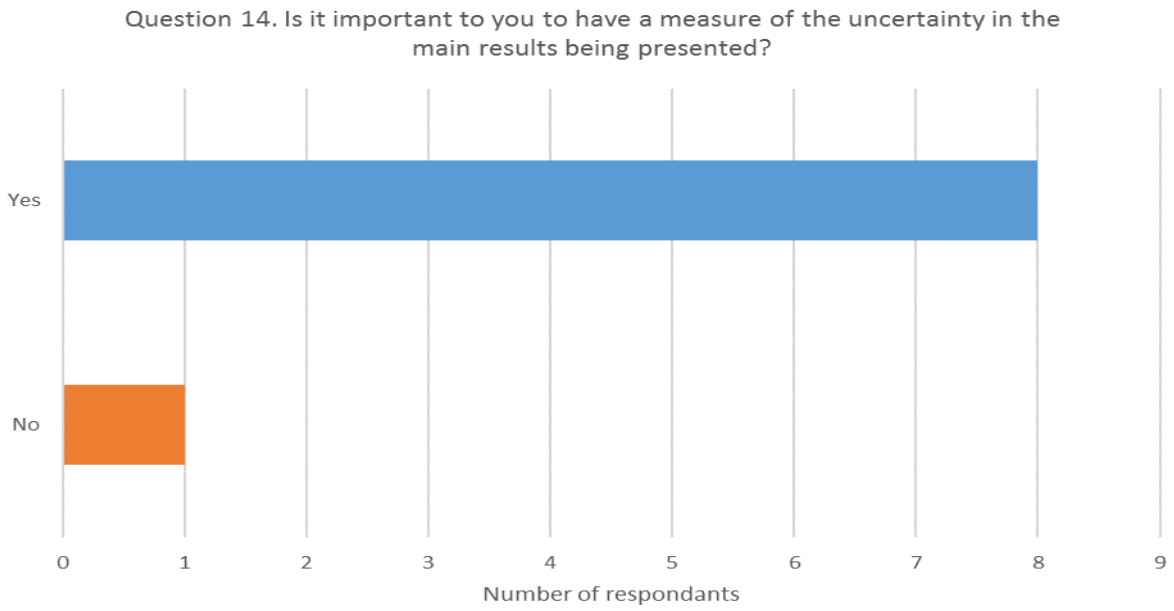
12. Over what time frame would you be most interested in the results being presented (please rank)?
    a. 1 year
    b. 5 years
    c. 10 years
    d. 20 years
    e. 50 years
    f. Lifetime of the current population
    g. Other (please specify)



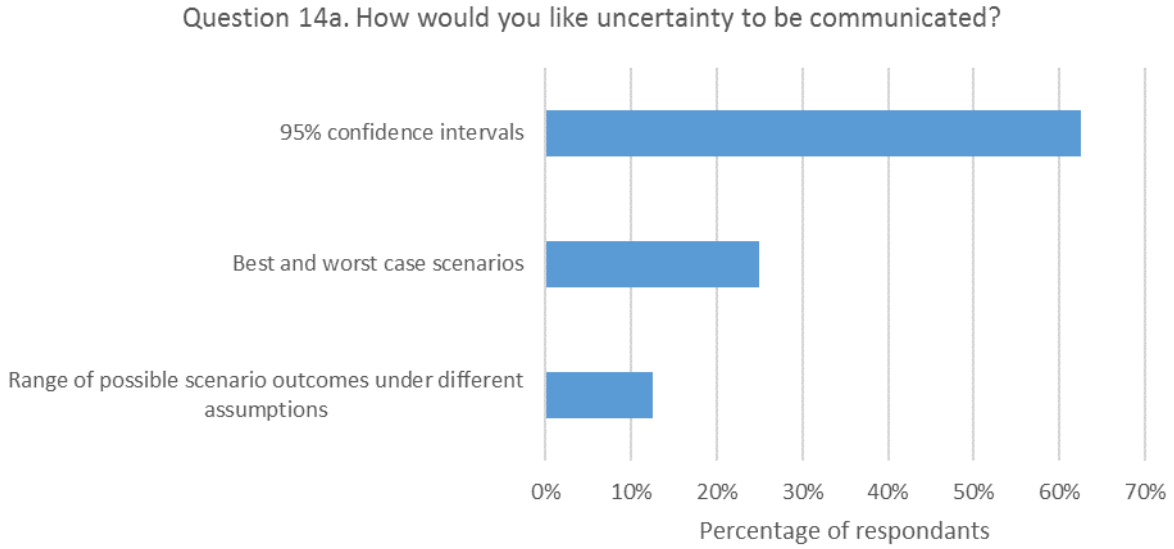
13. If you were presented with the model diagram above and the results of the comparison of different public health policies, does the logic of the model's flow seem valid and plausible to you?
    a. Yes
    b. No
        i. If no, what aspect is not believable and what should be added or changed?



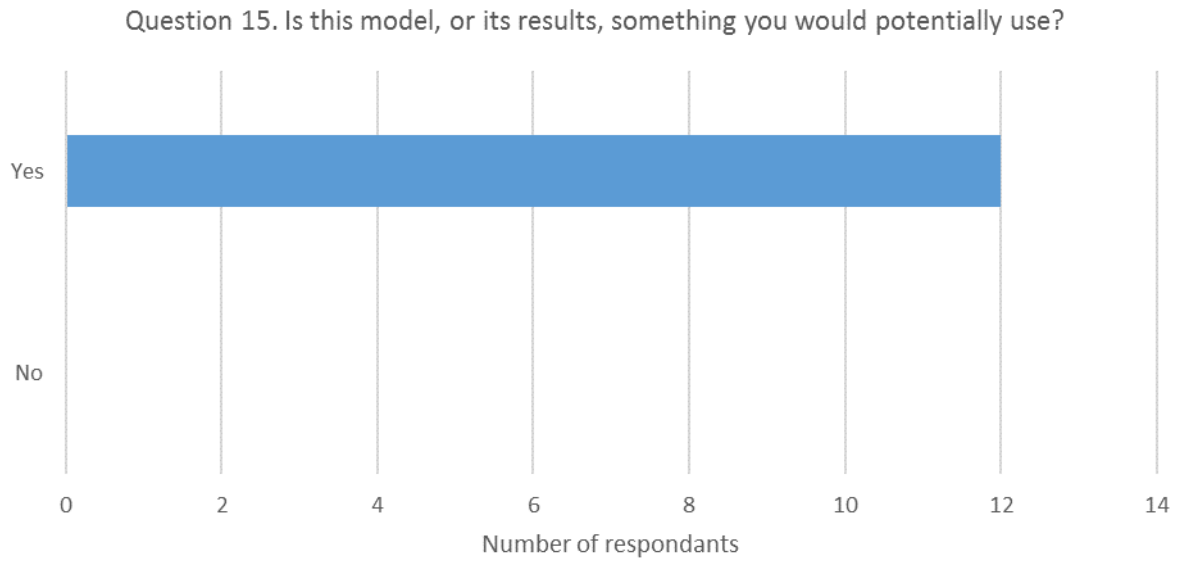
14. Is it important to you to have a measure of the uncertainty in the main results being presented?
    a. Yes
    b. No



    i. If yes, how do you prefer uncertainty to be communicated?
        1. 95% confidence intervals
        2. Best and worst case scenarios
        3. Range of possible scenario outcomes under different assumptions
        4. Other (please specify)



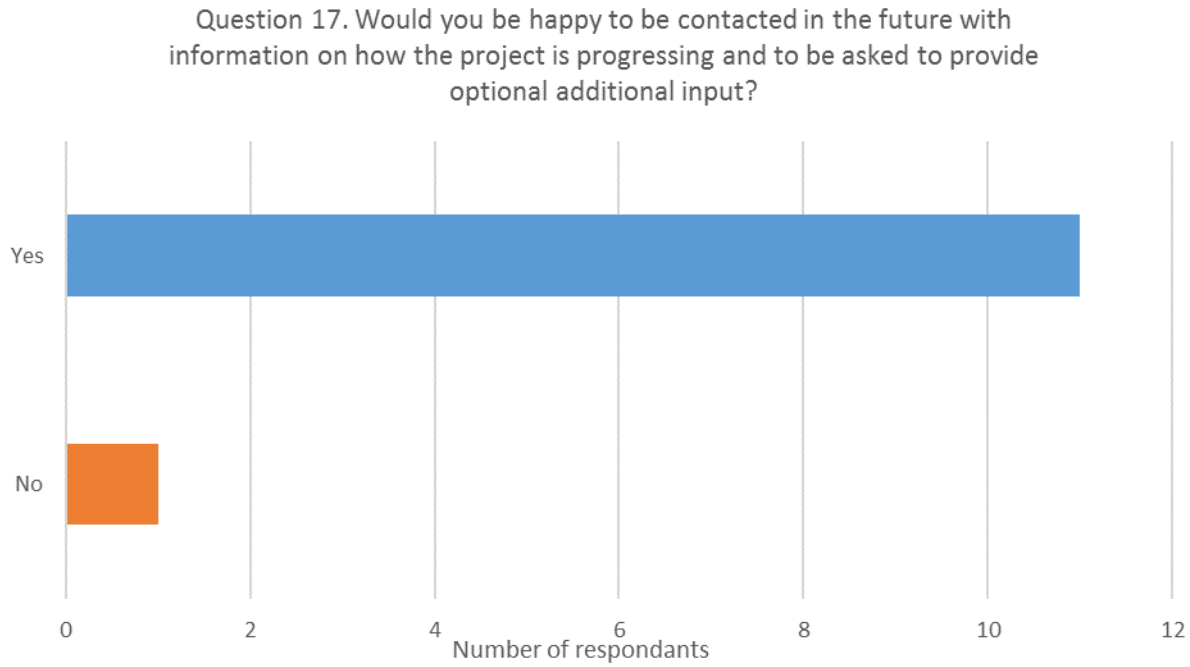
15. Is this model, or its results, something you would potentially use?
    a. Yes
        i. If yes, what would you use the model or its results for?
    b. No
        i. If no, what could be done to make it more useful to you?



16. Any other comments or suggestions?

Responses removed to maintain confidentiality.

17. Would you be happy to be contacted in the future with information on how the project is progressing and to be asked to provide optional additional input (it is anticipated this will be on two further occasions over the coming two years)?



Appendix 4 Data sources and input parameters for PRIMEtime and PRIMEtime CE

Table A4.1 Disease specific data inputs for PRIMEtime CE (all age and sex specific), reproduction of table 1 in supplementary file of Cobiac et al. (with permission) [131]

Disease	Data and methods
Coronary heart disease (CHD)	Incidence of CHD estimated from incidence rates of first acute myocardial infarction (derived from Hospital Episode Statistics[297]), adjusted using the proportion of unstable angina among all coronary events in the OXVASC study.[298] Mortality rates from the Office of National Statistics cause-specific death registrations (number of deaths where myocardial infarction was mentioned on the death certificate).[299] Case fatality rates and baseline prevalence derived using DISMOD II.[185]
Stroke	Incidence of first stroke estimated from the OXVASC study[298] and data from the General Practice Research Database.[300] Mortality rates from the Office of National Statistics cause-specific death registrations.[299] Case fatality rates and baseline prevalence derived using DISMOD II.
Type two diabetes	Incidence rates from the UK Clinical Practice Research Datalink.[301] Type two diabetes mortality rate ratios and prevalence estimated from the National Diabetes Audit 2011/12.[302] Case fatality rates derived using DISMOD II.[185]
Cirrhosis	Incidence rates from a population-based cohort study linking the Clinical Practice Research Datalink and Hospital Episode Statistics.[303] Mortality rates from the Office of National Statistics cause-specific death registrations.[299] Case fatality rates and baseline prevalence derived using DISMOD II.[185]
Cancers	Incidence rates from Cancer Registrations Statistics, England, 2012.[304] Mortality rates from the Office of National Statistics cause-specific death registrations.[299]

Case fatality rates and baseline prevalence derived using DISMOD II.[185]
NB Backg~~round~~ UK disease trends derived by Cobiac et al. using methodology from the global burden of disease project [131,305,306]

Table A4.2 PRIMEtime CE risk factors, intermediate variables, exposure parameters, outcomes, and modelled uncertainty distributions.

Adapted and updated from Cobiac et al.[131] Dietary relative risks adjusted for energy intake where possible

Parameter	Exposure parameters	Outcomes	Unit of change	Distribution	Relative risk (SE)	Notes
RISK FACTORS						
Fruit	Mean (SD) g/day for consumers and % consuming <1 fruit portion daily from the National Diet and Nutrition Survey (NDNS)[218]	CHD[307]	Per 106g/day fruit	Lognormal	0.93 (0.019)	
		Stroke[308]	Per 106g/day fruit	Lognormal	0.89 (0.023)	
		Lung cancer[309]	Per 100g/day fruit	Lognormal	0.94 (0.02)	
Vegetables	Mean (SD) g/day for consumers and % consuming <1 vegetable portion daily from NDNS[218]	CHD[307]	Per 106g/day vegetables	Lognormal	0.89 (0.034)	
		Lung cancer[309]	Per 100g/day vegetables	Lognormal	0.94 (0.025)	
Fibre (cereal only)	Mean (SD) g/day from NDNS[218]	CHD[310]	Per 10g/day fibre (cereal)	Lognormal	0.91 (0.02)	Cereals only to avoid double counting for fruit and vegetables
Fibre	Mean (SD) g/day from NDNS[218]	Breast cancer[311]	Per 10g/day fibre	Lognormal	0.93 (0.027)	
		Colorectal cancer[312]	Per 10g/day fibre	Lognormal	0.9 (0.034)	
		Stomach cancer[313]	Per 10g/day fibre	Lognormal	0.56 (0.12)	
Red meat intake	Mean (SD) g/day from NDNS[218]	Colorectal cancer[314]	Per 100g/day red meat	Lognormal	1.3 (0.06)	
		Stomach cancer[315]	Per 100g/day red meat	Lognormal	1.13 (0.036)	
		Type two diabetes[316]	Per 100g/day red meat	Lognormal	1.2 (0.072)	
Processed meat intake	Mean (SD) g/day from NDNS[218]	Colorectal cancer[314]	Per 50g/day processed meat	Lognormal	1.38 (0.07)	
		Type two diabetes[316]	Per 50g/day processed meat	Lognormal	1.57 (0.1)	
Serum cholesterol	Mean (SD) mmol/L from NDNS[218]	CHD[317]	Per -1mmol/l total cholesterol	Lognormal	<49 years: 0.44 (0.034) 50-59 years: 0.58 (0.034)	

---					60-69 years: 0.72 (0.018) 70-79 years: 0.82 (0.015) 80+ years: 0.85 (0.21)	
		Stroke[317]	Per -1mmol/l total cholesterol	Lognormal	<59 years: 0.90 (0.037) 60-69 years: 1.02 (0.027) 70-79 years: 1.04 (0.025) 80+ years: 1.06 (0.031)	
Systolic blood pressure	Mean (SD) mmHg from NDNS[218]	CHD[224]	Per -20mmHg systolic blood pressure	Lognormal	<49 years: 0.49 (0.042) 50-59 years: 0.50 (0.015) 60-69 years: 0.54 (0.0094) 70-79 years: 0.60 (0.013) 80+ years: 0.67 (0.023)	
		Stroke[224]	Per -20mmHg systolic blood pressure	Lognormal	<49 years: 0.36 (0.057) 50-59 years: 0.38 (0.034) 60-69 years: 0.43 (0.024) 70-79 years: 0.50 (0.02) 80+ years: 0.67 (0.03)	
Body mass index	Mean (SD) kg/m^2 from NDNS[218]	CHD[318]	Per 5kg/m^2 BMI	Lognormal	35-59 years: 1.5 (0.039) 60-69 years: 1.4 (0.031) 70-79 years: 1.31 (0.033) 80-89 years: 1.3 (0.055)	
		Stroke[318]	Per 5kg/m^2 BMI	Lognormal	35-59 years: 1.76 (0.075) 60-69 years: 1.49 (0.056) 70-79 years: 1.33 (0.056) 80-89 years: 1.1 (0.083)	
		Diabetes[318]	Per 5kg/m^2 BMI	Lognormal	BMI 15-25: 0.96 (0.25) BMI 25-50: 2.16 (0.067)	
		Pancreas cancer[319]	Per 5kg/m^2 BMI	Lognormal	1.1 (0.016)	
		Colorectal cancer[318]	Per 5kg/m^2 BMI	Lognormal	Men: 1.24 (0.016) Women: 1.09 (0.019)	
		Breast cancer[318]	Per 5kg/m^2 BMI	Lognormal	Women 60+ years: 1.12 (0.018)	
		Kidney cancer[318]	Per 5kg/m^2 BMI	Lognormal	Men: 1.24 (0.039) Women: 1.34 (0.034)	
		Liver cancer[318]	Per 5kg/m^2 BMI	Lognormal	1.47 (0.078)	
		Liver cirrhosis[318]	Per 5kg/m^2 BMI	Lognormal	BMI 15-25: 0.73 (0.016) BMI 25-50: 1.79 (0.077)	
Physical activity	Mean (SD) MET hours per week from Active People's Survey[240]	CHD[243]	11.25METhr/wk increase	Normal	-0.204 (0.027)	Parameter is the beta value of the 0.25 transformation of the

						relative risk, unadjusted for BMI. See chapter four.
---		Stroke[243]	11.25METhr/wk increase	Normal	-0.195 (0.041)	Parameter is the beta value of the 0.25 transformation of the relative risk, unadjusted for BMI. See chapter four.
		Diabetes[243]	11.25METhr/wk increase	Normal	-0.240 (0.023)	Parameter is the beta value of the 0.25 transformation of the relative risk, unadjusted for BMI. See chapter four.
		Colorectal cancer[242]	11.25METhr/wk increase	Normal	-0.080 (0.059)	Parameter is the beta value of the 0.25 transformation of the relative risk, unadjusted for BMI. See chapter four.
		Breast cancer[242]	11.25METhr/wk increase	Normal	-0.053 (0.023)	Parameter is the beta value of the 0.25 transformation of the relative risk, unadjusted for BMI. See chapter four.
RISK FACTORS OPERATING THROUGH INTERMEDIATE VARIABLES						
Total fat	% of total energy[218]	Total serum cholesterol (mmol/l)[320]	Per 1% energy total fat	Normal	0.020 (0.005)	
Saturated fat	% of total energy[218]	Total serum cholesterol (mmol/l)[320]	Per 1% energy saturated fat	Normal	0.052 (0.003)	
Monounsaturated fatty acids	% of total energy[218]	Total serum cholesterol (mmol/l)[320]	Per 1% energy MUFA	Normal	0.005 (0.003)	
Polyunsaturated fatty acids	% of total energy[218]	Total serum cholesterol (mmol/l)[320]	Per 1% energy PUFA	Normal	-0.026 (0.004)	
Dietary cholesterol	mg/day[218]	Total serum cholesterol (mmol/l)[320]	Per 1g/day dietary cholesterol	Normal	0.0007 (0.0001)	
Free sugars	% of total energy[218]	Per -18.9% energy added sugar[321]		Normal	-0.23 (0.056)	
Salt consumption	g/day[218]	Systolic blood pressure (mmHg)[223]	Per 100mmol/24hr urinary sodium	Normal	5.80 (1.71)	Grams of salt consumed per day converted into urinary sodium excretion.

Total energy	kJ/day[218]	BMI (kg/m^2)[322]				Details of equations describing the relationship between energy intake and body weight for men and women can be found in Christiansen and Garby.[322]
Diabetes	Prevalence	CHD[323]	If have diabetes compared to no diabetes	Lognormal	Men: 1.85 (0.063) Women: 2.63 (0.076)	
		Stroke[324]	If have diabetes compared to no diabetes	Lognormal	Men: 1.83 (0.067) Women: 2.28 (0.085)	
RISK FACTORS' THEORETICAL MINIMUM RISK						
Body mass index (kg/m^2)[325]	-	-	-	Normal	21 (1)	
Systolic blood pressure (mmHg)[325]	-	-	-	Normal	115 (6)	
Total cholesterol (mmol/L)[325]	-	-	-	Normal	3.8 (0.6)	
Vegetable intake (g/day)[267]	-	-	-	Normal	400 (30)	
Fruit intake (g/day)[267]	-	-	-	Normal	300 (30)	
Fibre intake (g/day)[267]	-	-	-	Normal	30 (3)	
Red meat intake (g/day)[267]	-	-	-	Normal	14.3 (1.43)	
Processed meat intake (g/day)[267]	-	-	-	Normal	0	
Physical activity (METhrs/wk)[326]	-	-	-	Normal	133 (13.3)	
MEDIATION FACTORS						
Ischaemic stroke mediation	BMI (kg/m^2)	Systolic blood pressure (mmHg)[326]		Normal	0.65 (0.04)	
	Fruit intake (g/day)	Systolic blood pressure (mmHg)[326]		Normal	0.42 (0.17	
	Vegetable intake (g/day)	Systolic blood pressure (mmHg)[326]		Normal	0.54 (0.2)	
	Fruit intake (g/day)	Total cholesterol (mmol/L)[326]		Normal	0.027 (0.017)	

	Vegetable intake (g/day)	Total cholesterol (mmol/L)[326]	Normal	0.047 (0.026)
Ischaemic heart disease mediation	BMI (kg/m^2)	Systolic blood pressure (mmHg)[326]	Normal	0.31 (0.016)
	Fruit intake (g/day)	Systolic blood pressure (mmHg)[326]	Normal	0.39 (0.15)
	Vegetable intake (g/day)	Systolic blood pressure (mmHg)[326]	Normal	0.47 (0.21)
	Fruit intake (g/day)	Total cholesterol (mmol/L)[326]	Normal	0.008 (0.0057)
	Vegetable intake (g/day)	Total cholesterol (mmol/L)[326]	Normal	0.012 (0.01)

SD, standard deviation; NDNS, National Diet and Nutrition Survey; CHD, coronary heart disease; BMI, body mass index; MET, Metabolic Equivalent of Task

Table A4.3 PRIMEtime CE sources and uncertainty distributions for baseline population data, costs, and utilities. All inputs are age and sex specific

Parameter	**Data and methods**
English population	From Office for National Statistics census data, no uncertainty estimated.[239]
Mortality rates	Extracted from the Human Mortality Database, no uncertainty estimated.[327]
Health sector costs	Disease specific costs derived from NHS England programme budgeting data[155] and unrelated disease costs estimated using NHS England cost curves.[165] Uncertainty estimated using a generic multiplication factor across all health sector costs with a gamma distribution based on a normal distribution (mean 1, SD 0.1). For more detail see chapter three.
Societal costs	Disease specific and unrelated productivity, social care, and wider societal costs estimated using a Department of Health tool published as a supplementary file in Claxton et al.[150] Uncertainty estimated using a generic multiplication factor across all societal costs with a gamma distribution based on a normal distribution (mean 1, SD 0.1). For more detail see chapter three.

Utilities	Baseline mean EQ-5D utility scores and disease specific decrements and their standard errors taken from Sullivan et al., with adjustments made for age and number of chronic conditions.[200] For more detail see chapter three.

Appendix 5 List of sensitivity analyses

Sensitivity analysis	**Explanation of what is changed compared to the primary analysis**	**Rationale**
Changing the time horizon.	Time horizon changing from 10 years in the primary analysis to one year, five years, 20 years, and 100 years (lifetime of the cohort).	The time horizon has the potential to significantly impact the cost effectiveness estimate. The tested time horizons were the options given to the project stakeholders. Shorter time horizons were chosen to align with local and national budget cycles, and longer time horizons allow for long term health benefits to be realised. A 50 year time horizon was not included as no stakeholder selected this as one of their top three choices (see appendix three).
Analysing results from an NHS perspective.	Estimating cost effectiveness using the change in NHS costs and intervention costs only (without any societal costs).	To make the results more directly relevant to NHS decision makers and to make them more consistent with other cost-effectiveness studies that estimate results from a healthcare perspective.[62,85,144,219]
Analysing results from a social care perspective.	Estimating cost effectiveness using the change in social care and intervention costs only (without any NHS costs).	To make the results more applicable to those in local authorities with responsibility for social care budgets rather than NHS budgets.
Including social care costs and productivity.	Adding an economic estimate of changes to productivity arising from the intervention.	To quantify the effects of the intervention on productivity so as to better represent the intervention's societal impact, the method is the same as that

		used by the UK Health Forum (see section 3.1.2.3).[148]
Including all wider societal costs.	Including an economic estimate of the intervention on all wider societal costs.	To quantify all wider societal costs using the tool produced by the Department of Health.[150] This provides results that are applicable to decision makers considering the broadest possible societal perspective (see section 3.1.2.3).
Using a discount rate of 3.5%.	Changing the discount rate for costs and outcomes from 1.5% to 3.5%.	The 3.5% discount rate is that advocated by NICE for non-public health interventions.[13]
No unrelated disease costs included.	Removing from the model any NHS and social care costs accrued from diseases unrelated to those explicitly modelled as a consequence of living longer following the intervention.	To better understand the impact of including costs arising from unrelated diseases on the final cost per QALY estimate (see section 3.1.2.2).
No cancer included in the model	Cancer removed from the model so that only IHD, stroke, type two diabetes, and liver cirrhosis are included.	The annual NHS costs of cancer per person estimated using the methods in section 3.1 were often lower than those reported in other studies. If estimated cancer costs are inaccurate, removing cancer from the model will demonstrate how much the cancer estimates affect the cost per QALY estimate.
Only including diseases directly related to the risk factor affected.	For the diet intervention, only IHD and stroke are included in the model, and for the physical activity intervention, IHD, stroke, type two diabetes, breast cancer, and colorectal cancer are included.	To understand the impact on the final cost per QALY of including diseases in the model that are not directly affected by the risk factor changed by the intervention.

Primary care costs based on hospital episode statistics data.	Attributing primary care costs to each disease based on the number of hospital admissions rather than the amount spent on primary care prescribing.	To understand how much the assumptions made when estimating primary care expenditure by disease affect the cost per QALY (see section 3.1.2.2).
Specialised services costs based on hospital episode statistics data.	Attributing specialised services costs to each disease based on the number of hospital admissions rather than the amount spent on 'other secondary care' in 2012/13.	To understand how much the assumptions made when estimating specialised services expenditure by disease affect the cost per QALY (see section 3.1.2.2).
Maternity costs estimated using the general and acute cost curve.	Attributing NHS England expenditure to each age and sex group based on the relative expenditure reported by the NHS England general and acute cost curve[165] rather than as a proportion of the number of births by age of the mother in England.	To understand how the assumptions made when attributing maternity costs by age and sex to unrelated disease costs affect the final cost per QALY (see chapter 3.1.2.2).
No delay in the introduction of the intervention (diet intervention only).	The effects and costs of the diet intervention come into full effect immediately rather than over a three year period.	To understand how the delay in the introduction of the diet intervention affects the cost per QALY.
No industry costs included (diet intervention only).	The cost to industry is not included in the intervention's costs.	To estimate the cost per QALY if industry were to not incur any costs from reformulation and instead reformulation occurs as part of the product cycle (a scenario tested by Collins et al., see section 4.1.1.2).[219]
Not including industry costs or government administrative	Only included costs from government monitoring.	Government administrative costs are relatively uncertain (see section 4.1.1.2). This sensitivity analysis is to understand how

costs (diet intervention only).		much removing monitoring costs would have on the cost per QALY, replicating a scenario by Collins et al.[219]
Fall in physical activity over time (physical activity intervention only).	Physical activity levels are not sustained following the intervention with 50% of individuals who move from active to recommended moving back to active one year after the intervention.	To understand how a fall in the sustainability of the intervention would affect the cost per QALY, replicating a scenario used by Frew et al.[144]

Appendix 6 Quality checklist for assessing studies identified in figure 3.2

Study screening questions

If answer no to one of the questions below, then disregard study.

1. Does the study include estimation of UK or English healthcare costs covering ICD-10 codes of interest?
2. Is the year of analysis 2001 or later?

Checklist questions

Item no.	Criteria	Yes/Partially/No (score 2/1/0)
1	Does the study state economic perspective?	
2	If estimated, are indirect and direct costs reported separately?	
3	Are methods described for cost and valuation estimates?	
4	Are all relevant healthcare costs included, or if not, is it clear as to which are and which are not?	
5	Are currency, price data, conversion rates, and medical care inflation calculations reported and appropriate?	
6	Is the time horizon stated, and are annual costs available or can they be derived?	
7	Are discount rates stated where appropriate?	
8	Is the relevant uncertainty around costs reported?	
9	What is the year of analysis (score 0 if 2001-2006, 1 if 2006-2011, 2 if >2011)?	

Appendix 7 Protocol for breast cancer utility values

This protocol describes the steps shown in figure 3.6 (section 3.2.2.2) and explains how breast cancer utility values were identified.[194]

1. Search strategy

    a. Inclusion criteria

- Paper describes empirically estimated original health state utility value or is a systematic review of utility values
- Utility value relates to health state of interest
- EQ-5D used to describe health state
- Patient population used to measure change in HRQoL
- Description given of technique used to value health state and of population used to value change to HRQoL
- Study population aged > 17 years of age and applicable to the English population suffering from disease (applicable to the English population is defined as studies with patients from Europe, United States of America, Canada, Australia, and New Zealand)
- Published in English

    b. Databases

- MEDLINE
- EMBASE
- PsychINFO
- Science Citation Index (ISI)
- Social Sciences Citation Index
- Conference Proceedings Citation Index – Science (CPCI-S)
- Cochrane library (CDSR, CENTRAL, DARE, CMR, HTAD, NHSEED, ABOUT)
- The EQ-5D website

- (The Cost-Effectiveness Analysis Registry at Tufts Medical Centre not searched because Oxford University does not have premium access thereby limiting search options and results)

c. Supplementary searching

Search of reference lists (including of non-included reviews)

d. Search terms

Search terms for identifying breast cancer utility values developed using those found in Papaioannou et al. 2010 and Peasgood et al. 2010 [194,328].

An example of search terms used is shown below for Web of Science:

Indexes=SCI-EXPANDED, SSCI, CPCI-S Timespan=1945-2015

1. TS=(euroqol OR euro qol OR eq5d OR eq 5d)
2. TS=((breast NEAR/2 cancer*) OR (breast NEAR/2 neoplasm*) OR (breast NEAR/2 carcinoma*) OR (breast NEAR/2 neoplasm*) OR (breast NEAR/2 tumo*) OR (malignan* NEAR/2 breast) OR (mammary NEAR/2 cancer*) OR (mammary NEAR/2 neoplasm*) OR (mammary NEAR/2 carcinoma*) OR (mammary NEAR/2 neoplasm*) OR (mammary NEAR/2 tumo*) OR (malignan* NEAR/2 mammary))
3. #1 AND #2

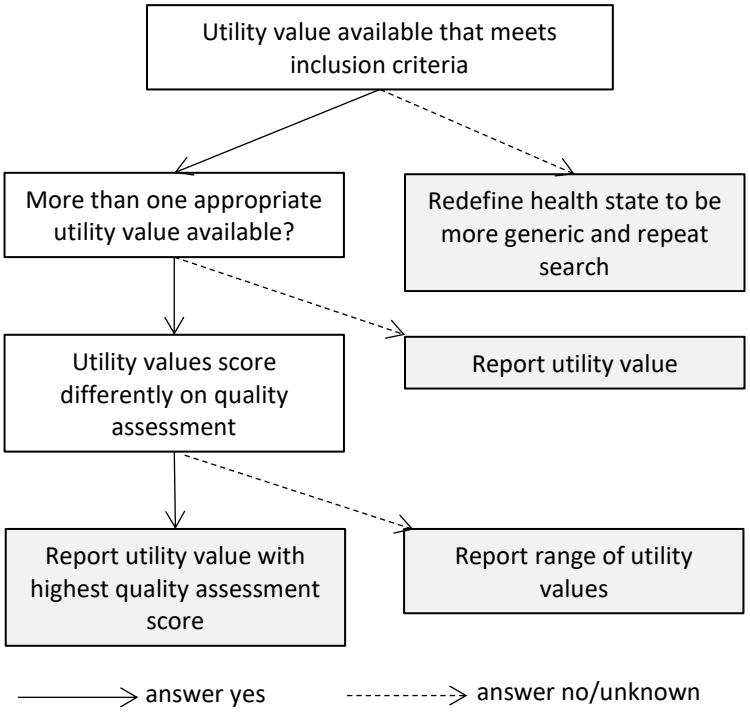
e. Refining the search

The search strategy can iteratively develop to balance the sensitivity and specificity of the search.

2. Deciding which utility values to report

    a. Flow chart for deciding which utility values to report is shown in figure A7.1

Figure A7.1 Flow chart for deciding which utility values to report



    b. Study quality assessment

If a systematic review needs to be quality assessed, use adapted CASP checklist (box A3.2.i).[162] For the systematic review to be considered good quality, the answer to questions 1-5 and to question 9 in box A7.1 should all be *yes*.

Box A7.1. Adapted CASP-style checklist for appraising the quality of a systematic review of utility values[162,194]

For the following questions, answer Yes, No, Partly, or Can't tell:

1. Did the review ask a clearly-focused question?

Consider if the question is focused in terms of:

- Population describing the health states (ideally patients)
- Population valuing the change in HRQoL (ideally public)

- Method of elicitation (ideally choice-based method e.g. TTO)

2. Did the review include the right type of study?

Consider if the included studies:

- address the review's question.
- are appropriate studies.

3. Did the reviewers try to identify all relevant studies?

Consider as a minimum:

- Were a number of electronic databases searched? (ideally clinical and specific health economic)
- Were reference lists scrutinised for retrieved references?

Ideally, but not mandatory, consider that the search methods should involve:

- personal contact with experts.
- searches for unpublished studies.
- citation and author searching.

4. Did the reviewers assess the quality of the included studies?

Consider the:

- Sample size.
- Respondent selection and recruitment.
- Inclusion/exclusion criteria.
- Response rates to instrument used.
- Numbers (%) lost to follow-up.
- Are reasons provided for any loss to follow-up?
- How is missing data from the instruments used to describe the health states dealt with? Is the method rigorous?
- Any other problems with the study.

5. Did the reviewers assess the relevance of the included studies to the review question?

- Population describing the health states (ideally patients).
- Population valuing the HRQoL (ideally public).
- Method of elicitation (ideally choice-based method e.g. TTO).

6. If the results of the studies have been combined, was it reasonable to do so?

- Are the results of each study are clearly displayed?
- Are the results were similar from study to study (look for tests of heterogeneity)?
- Are the reasons for any variations in results are discussed?

7. How are the results presented and what is the main result?

- Is there a full account of why studies were excluded?
- Is there are full justification of why studies were included?
- How are the results expressed (descriptive statistics or coefficients of a model)?

8. How precise are these results?

Consider:

- if a confidence interval was reported, would your decision about whether or not to use this intervention be the same at the upper confidence limit as at the lower confidence limit?
- if a p-value is reported where confidence intervals are unavailable.

9. Can the health state utility values be used in the health states in your decision model?

Consider:

- how relevant the population describing the health state is to the health states in the decision model.
- have all subgroups been considered e.g. age, disease severity, setting?
- do the utility values match the NICE reference case?
- how do the results need to be modified for the decision model?

If assessing empirical studies, the criteria in table A7.1 will be used to assess study quality and applicability to the English setting, adapted from Papaioannou et al. 2010.[194] No study will be excluded on the basis of the score obtained, instead the score will be used to help decide which utility values to report (see figure A7.1). If comparing a review with an empirical study, a judgement based on the two separate quality assessment tools will be used to decide which utility value to report.

Table A7.1 Assessing quality and applicability of empirical studies estimating utility values

Criteria	**What to consider**	**Yes/partially/no (score 2/1/0)**
Uncertainty	Does the estimated utility value include the variance reflecting the sample size?	
Respondents	Is the study population comparable to that in the model (consider inclusion criteria)?	
Response rates	Is response rate reported and large enough for it not to bias results?	
Loss to follow-up and missing data	Are losses to follow-up and missing data small and unlikely to lead to bias?	
Instrument used	Is EQ-5D used to describe health states?	
Population using instrument	Did the relevant patient population describe the change in health state and is the patient population applicable to the English setting (score 2 if UK patient population, score 1 if	

	non-UK patient population but applicable to UK*)?
Population valuing change in HRQoL	Was the valuation of change in health state completed by the UK (score 2) or UK applicable (score 1) general population?
Technique of valuing health state	Was TTO used (score 2), or a different choice based method (score 1)?
Any other problems	Deduct 0, 1, or 2 points depending on number and scale of problems

*countries applicable to UK are defined as countries in Europe as well as the United States of America, Canada, Australia, and New Zealand.

3. Data extraction form

The data extraction form has two roles, the principal role is to collate all the relevant information from a study that may be used to inform the model. The second role of the data extraction form is to allow for further development of the inclusion criteria through direct comparisons of different studies – figure 3.6 (section 3.2.2.2) illustrates this iterative process with arrows in both directions; table A7.2 is the data extraction form.

Table A7.2. Data extraction form

Study ID	Full study reference	Inclusion criteria met	Study design	Country

Health state (including time horizon)	Instrument used to describe health state	Health state valuation technique	Population characteristics measuring HRQoL change	Population characteristics valuing HRQoL change

Respondent selection and recruitment	Response rate	Loss to follow up and reasons	Other problems/biases

Utility values (per subgroup and health state)				
Sample size	Mean (SD)	Median	Range	95% CIs