USING COST-EFFECTIVENESS ANALYSIS TO ADDRESS HEALTH INEQUALITY CONCERNS

James Shen Kit Love-Koh

Ph.D.

University of York

Economics

September 2017

Abstract

Little quantitative assessment of health inequality impacts occurs in the economic evaluation of health care. Distributional cost-effectiveness analysis (DCEA) offers an extension to cost-effectiveness analysis, estimating health inequality changes alongside population health.

This thesis addresses two important methodological and empirical challenges for DCEAs. First, in order to measure inequality change, a baseline level distribution of health needs to be estimated. In chapter 2 we do this for England using health survey and national mortality data to predict lifetime health. We estimate a gap between the most and least healthy fifths of 10.97 QALYs.

Second, we estimate how the health effects of health care budget changes in England are allocated between social groups. We estimate the socioeconomic distribution of health care utilisation by disease, age and gender, which are used to disaggregate results from a previous study that estimated effects of expenditure by disease area. We find a substantial gradient in health effects, with 27% and 13% incurred by the most and least deprived fifths of the population, respectively.

We apply the findings of the previous chapters to two different types of DCEA. In chapter 4, we propose a simplified version of DCEA, in which intervention health benefits follow the gender and socioeconomic patterns of health care utilisation. We apply this approach to 27 technology appraisals conducted by NICE, and find five interventions increase population health and worsen inequality; all still increase social welfare even when inequality aversion is high.

In Chapter 5 we conduct a full DCEA to evaluate smoking cessation interventions. We adapt a decision model to incorporate a wide range of key model inputs varied by socioeconomic status. As effectiveness and uptake are greater in the least deprived groups, all interventions increase health inequality, despite the greater number of smokers in the more deprived groups.

Table of Contents

List of Tables

List of Figures

Acknowledgements

Most of all I would like to thank my wife, Alpha, for the patient support that she displayed each day I spent on this research.

I am also immensely grateful for the guidance, encouragement and support of my supervisors Susan Griffin and Richard Cookson, who have helped to make conducting this research such a worthwhile and productive experience.

Thanks must also go to my other co-authors, Karl Claxton, Miqdad Asaria, Thomas Patton, Nils Gutacker, Rebekah Pennington and Lesley Owen. Particular thanks are due to Miqdad, whose sage advice was generously given whenever it was sought.

Thanks are also owed to Marta Soares, Mark Sculpher, Aki Tsuchiya and Anita Charlesworth, for the comments and advice they provided on various drafts.

Declaration

I confirm that the work presented in this thesis is my own. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Funding for this Ph.D. was provided by the Department of Health's Policy Research Unit for Economic Evaluation. The work from which chapter 5 was developed was commissioned and funded by NICE's public health programme as part of the institute's Economic and Methodological Unit. All of this research was conducted in conjunction with the research objectives of the respective units and was co-authored with the authors stated below.

Chapter 1 is solely my own work. I planned and conducted and wrote both parts of the literature review. Comments and suggestions on structure and additional papers of interest were suggested by advisory panel: Dr Susan Griffin, Professor Richard Cookson and Professor Karl Claxton.

Chapter 2 is written in co-authorship with Dr Miqdad Asaria, Dr Susan Griffin and Professor Richard Cookson. It was developed from my M.Sc. dissertation and published in a peer reviewed journal under the title: The Social Distribution of Health: Estimating Quality-Adjusted Life Expectancy in England; *Value in Health*, vol. 18, pp. 655–662, 2015. The original dissertation was developed substantially: large sections have been rewritten; new data were prepared and analysed; and multiple new analyses and technical improvements have been conducted. I am the lead author, wrote the initial draft and subsequent revisions and disseminated the paper. I also designed and conducted all of the quantitative analysis.

Chapter 3 is written in co-authorship with Professor Richard Cookson, Professor Karl Claxton and Dr Susan Griffin. Earlier versions of this paper were presented at the 2nd EuHEA PhD Student-Supervisor and Early Career Researcher Conference in 2015 at Universitè Paris-Dauphine and the 2016 Winter Health Economists' Study Group Meeting in Manchester. I contributed to the original research idea and am

again the lead author. I prepared the data, designed and conducted the empirical analysis and wrote the drafts and revisions.

Chapter 4 is co-authored with Professor Richard Cookson, Dr Susan Griffin, Dr Thomas Patton and Dr Nils Gutacker. I am the lead author for this work, and carried out the empirical analysis, wrote the initial draft and subsequent revisions. The data on National Institute of Health Care Excellence (NICE) technology appraisals was extracted and prepared by Dr Thomas Patton and Dr Nils Gutacker. The section of the chapter describing this extraction is based on an earlier version prepared by them both. I conducted the remaining data preparation and analysis, and contributed, along with Professor Karl Claxton and the other co-authors, to the original research idea.

Chapter 5 is co-authored with Dr Susan Griffin, Becky Pennington and Dr Lesley Owen. It was developed from a previous piece of work entitled "Frameworks for considering health inequality and their suitability for use within NICE public health guidelines". The decision analytic model used in the analysis was developed by Alexandra Filby and Dr Matthew Taylor of the York Health Economics Consortium for the NICE public health guideline on smoking cessation interventions and services. I contributed to the analysis plan and was the lead author, writing the draft and revising subsequent versions in conjunction with Dr Susan Griffin. I contributed to the data preparation and conducted the model adaptation and the quantitative analysis.

Introduction

Governments and health care organizations throughout the world are faced with inequalities in health in their populations. There is a strongly held view amongst the organisations faced with these problems that inequalities should be reduced in the interests of fairness (DH 2013; PHAA 2012; Raphael 2000). The UK is no different: health inequalities have been on the public agenda for decades (DHSS 1980; Lowdell et al. 1999), although they have, by some measures at least, remained relatively constant or increased (ONS 2013; Asthana et al. 2013). Health inequalities can be tackled by addressing their 'social determinants', such as income and education level, which are associated with better health production capabilities and healthier lifestyle choices (Grossman 1972; Currie 2009). Decision-makers can and do influence these determinants through policies in the education sector, the labour market, or through a redistributive tax system (Sen 2002; Graham and Kelly 2004; Marmot et al. 2010). However, the extent to which they can wield each policy is limited politically. An alternative lever that can be used by social decision-makers for reducing health inequalities is through the provision of health care itself. This thesis aims to provide methodological and empirical foundations for the assessment of the health inequality impacts of interventions that directly target health as a primary outcome (including health care and public health), so that decision-makers can be routinely presented with quantitative evidence on this key social objective.

Over recent decades, a large number of health technology assessment (HTA) bodies have been established to assist decision-makers on health-related funding decisions through evidence-based techniques (Mathes et al. 2013). The body responsible for this function in the England is the National Institute for Health and Care Excellence (NICE), which in addition to evaluating medical technologies, pharmaceuticals and devices, also develops guidelines for clinical practice and public health. The principal quantitative analysis that is routinely conducted within the assessment process adopted by NICE has been cost-effectiveness analysis. The objective underpinning this analytical method, which prioritizes those technologies with the lowest cost-per-unit of health gain, is to maximise total population health. HTA bodies now define cost-effectiveness as being synonymous with health maximisation

and explicitly adopt it as their primary evaluative criterion. However, many studies have suggested that both decision-makers and the general public feel that other important criteria should also be considered, including the impacts on the distribution of health (Nord et al. 1995; Dolan et al. 2005; Dolan et al. 2008; Dolan & Tsuchiya 2005; Guindo et al. 2012; Mirelman et al. 2012; Tanios et al. 2013). This has led to repeated calls for an alternative decision framework that quantitatively incorporates concerns for health inequalities (Wagstaff 1991; Powers & Faden 2000; Sassi et al. 2001; Cookson et al. 2009).

The objective of reducing inequalities can sometimes conflict with maximizing health, creating an equity-efficiency trade-off problem (Wagstaff 1991; Williams & Cookson 2006). For example, a policy intervention may maximise total health but increase health inequality, or reduce health inequality but not maximise health. In such cases, a simple decision rule like health maximisation will be insufficient. A recently developed framework that can incorporate equity-efficiency trade-offs is distributional cost-effectiveness analysis (DCEA) (Asaria et al. 2015). DCEA is an extension of the existing process of cost-effectiveness analysis in which the expected health benefits and opportunity costs are estimated for equity-relevant subgroups and modelled onto a distribution of lifetime health. Pre- and post-intervention distributions are evaluated using a health-related social welfare function that is 'equity-regarding' (Adler 2013), meaning they embody concern for the distribution of health as well as the sum total. These allow for the explicit analysis of trade-offs between equity and efficiency.

In a typical cost-effectiveness analysis, a decision model is developed to estimate the incremental costs and benefits of a new treatment compared to current practice. A superficial distributional analysis may go beyond the average patient by conducting subgroup analysis for equity-relevant groups such as socioeconomic status or ethnicity, and calculating cost-effectiveness statistics for each group. A truly robust analysis of the health inequality impacts of a new intervention requires additional data and stages of analysis, however. One essential component is to account for how the health opportunity costs of funding a new intervention are distributed between groups; only then can the net health impact of an intervention be estimated. Since the funding for a new intervention will be generated from displacing services across the

health sector, the distribution of health opportunity costs should also relate to the entire health care budget.

A second key consideration relates to the measurement of health inequality impacts. Intuitively, the inequality impact will be determined by whether an intervention increases or reduces the level of health inequality in the population. Regardless of whether the measure of inequality is absolute or relative, a baseline distribution of health is required upon which to model the distributional impacts of the intervention (Cookson et al. 2017). Given that opportunity costs fall across the population, this baseline distribution also needs to be estimated for the whole population.

The objective of this thesis is to conduct methodical and empirical work that addresses both of these challenges. Methods are presented for estimating the baseline health and the distribution of health opportunity costs, and are applied to the case of the NHS in England. Two different approaches to conducting DCEA are then presented that utilise these results.

Chapter 1 is made up of two distinct literature reviews. The first section describes the current policy environment in the English NHS with respect to health inequality. Through this, we explore the implicit and explicit attitudes towards inequality present in the decision making processes of the NHS and NICE. The second section reviews the existing alternative and supplementary approaches to cost-effectiveness analysis that have been proposed in the literature. The feasibility and methodological validity of each approach is then critically appraised.

Chapter 2 demonstrates a technique for estimating quality-adjusted life expectancy by age, gender and socioeconomic status, which can be used to construct a baseline distribution of health. In adopting the quality-adjusted life year metric used in cost-effectiveness analysis, health impacts from these analyses can be modelled directly onto the distribution in order to measure health inequality impacts. This study contributes to the literature by:

- Being the first to estimate health inequality in England in terms of quality-adjusted life years;

- Providing an estimate of inequality that is sensitive to morbidity differences but goes beyond binary measures of ill-health (such as disability-free life expectancy)
- Using whole-population mortality data and health-related quality of life data from a random sample of over 25,000 people and capturing the uncertainty in these estimates through Monte Carlo simulation;
- Estimating a univariate distribution of quality-adjusted life expectancy at birth based on inequalities by gender and deprivation groups, in which the whole population is ranked from least to most healthy.

Chapter 3 investigates how the health effects from marginal changes in health system expenditure are distributed across social groups of interest. This chapter uses recent data published on the statistical relationship between health and expenditure in the English NHS by Claxton et al. (2015), which estimates how much and what type of health is lost when health care expenditure is reduced at the margin. We estimate how these health losses are distributed across society in the case of a single payer health care system with a fixed budget: the English NHS. Health care utilisation statistics by age, gender and disease are used to determine the distribution of health effects. This study provides three main contributions to existing knowledge on this topic:

- It is the first to provide an empirical estimate of the social distribution of health effects from marginal NHS budget changes that accounts for length and quality of life;
- As reductions in expenditure are analogous to the reallocation of existing funds to a newly approved intervention, the results also provide an estimate of the distribution of health opportunity costs;
- As the distribution also estimates how marginal increases to the budget are distributed, the findings represent the opportunity cost of alternative interventions when new funds are allocated to the health sector.

Chapters 4 and 5 demonstrate how the results of chapters 2 and 3 can be implemented in a distributional cost-effectiveness analysis framework. In chapter 4, a simplified version of DCEA is proposed to estimate and evaluate the health

inequality impacts of 27 health technologies approved by NICE between 2012 and 2014. This simplified method takes into account social variation in condition-specific health care utilisation, for which data are readily available, but not social variation in technology-specific uptake and efficacy, for which relevant data are often harder to find. For example, the approach allows for the inequality-reducing impact of hepatitis C therapies that relate to the higher proportion of patients in more deprived areas, but not the countervailing impact of therapies that relate to the better adherence and health outcomes of patients in less deprived areas. Health benefits are estimated using the distribution of hospital utilisation for the relevant condition, and combined with the health opportunity cost estimates from chapter 3 to obtain the 'net' health impact for each equity-relevant subgroup. These are modelled onto the baseline distribution from chapter 2. Inequality and social welfare measures are used to estimate the social impacts of each technology. This work makes the following contributions to the literature:

- Offers a new and simplified technique for estimating health inequality impacts that utilises widely available data on health care utilisation;
- Demonstrates the impact of accounting for the distribution of opportunity costs when analysing changes to health inequality;
- Provides approximate estimates of the health inequality impacts of a wide range of health technologies approved by NICE.

In chapter 5 we conduct a full DCEA to evaluate the health inequality impacts of 21 smoking cessation interventions, including both behavioural interventions and pharmacotherapies. A decision model developed to support a NICE public health guideline is retrospectively adapted to estimate cost-effectiveness results by socioeconomic group, the results of which are then used to conduct the DCEA. The objective is to demonstrate that by using a traditional decision model, a robust DCEA can be conducted with limited additional resources through pragmatic literature reviews that identify data on how decision model inputs vary by social group. This analysis also allows us to explore the impact of including additional sources of socioeconomic variation by comparing our results to those yielded by the simplified DCEA approach outlined in chapter 4. These results are estimated using the average estimates of incremental costs and health and the socioeconomic and gender

distribution of smokers in the population. This chapter makes the following contributions to the literature:

- Outlines how decision analytic models can be adapted to estimate the incremental costs and health effects of new interventions by equity-relevant subgroups;
- Estimates to what extent the simplified approach to DCEA overestimates health inequality impacts by not accounting for differential efficacy and uptake;
- Provides quantitative evidence on how funding smoking cessation interventions might affect health inequalities in England.

Chapter 1: Literature review

Allocating resources in health care in an equitable way is an immensely complex process. Potential recipients all have competing moral and clinical claims to resources, yet finite resources mean it is impossible to treat them all. Williams and Cookson (2006) identify the lack of clarity around the concept of equity as one of the principal reasons that HTA has failed to formally incorporate criteria other than efficiency. Johri and Norheim (2012) support this claim, citing the multifaceted nature of equity as the principal constraint on methodological progress. Equity can be defined in terms of the characteristics of patients and their illnesses. For example, many studies have argued that greater priority should be given to those with more severe illnesses (Dolan 1998; Ubel 1999; Nord 1993; Nord 2005; Brazier et al. 2013). Another concern has been that a health maximisation approach may discriminate against those who have relatively limited capacity to benefit from treatment, such as the elderly, the co-morbid and the disabled, since treatments directed at them are likely to be comparatively less effective (Nord et al. 1999; Kamm 2013).

The focus of this thesis is on a different aspect of equity - inequalities in lifetime health. An equitable distribution of health does not imply equality of health, however. Differences in health are the consequence of a complex web of causes that occur throughout the life course, some of which can be judged fair, others unfair. However, discussions on the specific set of social value judgements that determine if causes are fair or unfair will be left to others (Fleurbaey & Schokkaert 2009). Instead, this thesis provides methodological and empirical contributions to an evaluative framework that formally incorporates concerns for health inequality, howsoever it be defined. The methods adopted in the following chapters can therefore be adapted to whatever set of social characteristics are deemed unfair and are not specific to those that are used in the applications presented (which are gender and socioeconomic status).

This chapter reviews the existing approach in the UK health system to addressing health inequalities and the efforts made in the academic literature to develop evaluative frameworks that incorporate concerns for them.

1.1 Current institutional arrangements

1.1.1 Health inequalities and the NHS

The NHS has a clear mandate to provide equitable health care and to reduce health inequalities (Department of Health 2013). Highly publicized government-commissioned reports such as the Black Report (DHSS 1980) and the Acheson Report (Lowdell et al. 1999) have advocated for impacts on inequalities to be considered in the provision of health care and in public health interventions. In another NHS report, Graham and Kelly (2004) highlight some of the major determinants of social inequalities in health and make several recommendations with respect to monitoring and reducing them. In 2009, a parliamentary report lamented the magnitude of inequalities and the lack of rigorous evaluation studies looking at the health inequality impacts of policies, concluding that more needs to be done in health care (with specific reference to the NICE model) and public health (House of Commons Health Committee 2009, p.48).

Since 2002, the NHS has used its resource allocation procedures to reduce health inequalities by funnelling additional funds to regions with higher levels of 'unmet need.' The resource allocation formula estimates health care utilisation as a function of patient population characteristics that affect regional need. It is used to guide the budget apportioned to regional commissioning bodies, which up until 2012 were Primary Care Trusts (PCTs) and are now Clinical Commissioning Groups (CCGs). Following the introduction of the AREA formula (Sutton et al. 2002) and its replacement, the CARAN formula (Morris et al. 2007), the deprivation level of commissioning bodies has had a much greater effect on its budget allocation due to the inclusion of variables designed to represent levels of 'unmet need'.

Unmet need reflects the belief that those in groups with higher levels of deprivation are more likely to have undiagnosed disease, less likely to utilise preventive health

care services and less likely to secure high quality co-ordinated care for long-term conditions for a variety of complex reasons, described in more detail in Cookson et al. (2016). In terms of the allocation formula, this creates the effect of biasing the impact of deprivation downwards because of a correlation between deprivation levels and unobserved health variables (Morris et al. 2007, p.17). By creating morbidity indices that reflect the unmet needs of the population and controlling for them in the analysis, the authors argue that effects of deprivation are more adequately captured.

The inclusion of unmet need in the resource allocation formula has increased the resources provided to more deprived areas of the UK, meaning that significant portions of the NHS budget are dedicated to reducing these avoidable health inequalities. However some, including Asthana et al. (2013), have pointed out that health inequalities have not been noticeably reduced and have actually grown larger in some respects. The reason for this lies in the prevailing view that social and economic determinants of health, such as income and education, have a far greater impact than health care (Shiell 2009; Marmot et al. 2008; Powers & Faden 2000) and have consequently overridden any beneficial effects of resource allocation policy. In this sense, policies from outside of the health sector that focus on these social determinants can be just as valuable as the efforts made within it towards reducing health inequalities (Barr et al. 2017). However, cross-sectoral evaluation is currently a Pandora's box of methodological and empirical issues made even more complicated when considering distributional impacts. We therefore leave the discussion of its use to others (for example, see Claxton et al. (2010) and Adler (2011)).

1.1.2 Health inequalities and NICE

Increasing the level of resources to more deprived areas does not automatically help to reduce health inequalities, since CCGs may not be funding health care towards those with the worst health or greatest need. The body tasked with performing economic evaluations and informing the purchasing decisions of CCGs is NICE. However, the formal evaluation model used by NICE, which will be reviewed in this section, does not estimate distributional impacts quantitatively, meaning that we

ultimately do not know whether the resource allocation policies are improving health inequalities.

A typical quantitative analysis employed by NICE examines the costs of the intervention that fall on the NHS budget and the health benefits created, measured in quality-adjusted life years (QALYs) (NICE 2013). Since the budget is fixed, all of the costs used to fund the new technology displace other services currently provided somewhere within the NHS. Costs can therefore, in principle, be converted into health opportunity costs using an estimate of how much health is displaced per monetary unit, which are then compared with the health benefits to see if there would be a positive net health effect of approving the technology. Net health benefit (NHB) can be written:

$$\mathrm{NHB} = \Delta \mathrm{H} - {}^{\Delta \mathrm{C}}/_{\mathrm{k}}$$

Where $\Delta \mathrm{H}$ is the change in health over the best available comparator; $\Delta \mathrm{C}$ is the change in costs; and k represents the health opportunity cost: the cost-per-QALY of the displaced services. Recent empirical work has estimated k at around £13,000 per QALY (Claxton et al. 2015). Whilst k is conceived of as health opportunity cost in principle (Drummond et al. 2005; McCabe et al. 2008), NICE itself uses a cost-effectiveness 'threshold' range of between £20,000 and £30,000. This means that technologies will be approved when the net benefit is positive for all $\mathrm{k} \leq 20{,}000$, and rejected when the net benefit is negative for all $\mathrm{k} \geq 30{,}000$, unless there is (i) strong evidence of significant additional non-health benefit or (ii) the they treat patients at the end of life, both of which are expanded upon below. Rather than to reflect health opportunity costs, NICE adopts this threshold range to achieve a "reasonable compromise between ensuring everyone has fair and equitable access to the NHS and enabling access to new and innovative treatments" (Dillon 2015). McCabe et al. (2008) argue that whilst these other attributes of benefit should not be excluded from the decision-making process, the threshold should, however, reflect health opportunity cost and that cost-effectiveness should be a necessary but not sufficient condition for approval.

The principle of health maximisation that follows from cost-effectiveness analysis is a special form of utilitarianism that uses QALYs as the maximand rather than individual utility (Wagstaff 1991; Bleichrodt 1997). The distributional principle underlying health maximization is that every QALY is equal and no recipient is more or less deserving than any other (Drummond 1989; Wagstaff 1991). This principle of 'a QALY is a QALY is a QALY' (Dolan & Tsuchiya 2006) is problematic on ethical and empirical grounds. Health maximisation is not the sole professed goal of decision-makers in the UK and elsewhere, as evidenced by the threshold range adopted by NICE and the continued desire to improve the health of the worst-off and dedicate more resources to them (NICE 2008a). Numerous studies of decision-makers and the general public have also demonstrated that there are many other attributes of benefit considered socially important (see Guindo et al. (2012) for a review). It is for these reasons that the continued exclusive use of cost-effectiveness analysis as the quantitative evaluative tool in HTA has motivated the development of a number of alternative methodologies that are reviewed later in this chapter.

Despite concerns for inequality not being represented in the quantitative aspects of HTA, they do appear in other stages of the process. This first occurs when assessors define the remit of the appraisal ('scoping') where issues of interest are identified. These include 'the disease or health condition and the population(s) affected' (NICE 2013, p.16) and 'issues relating to advancing equality of opportunity' (NICE 2013, p.17). Inequality concerns are also deemed relevant in the assessment process when technologies have an incremental cost-effectiveness ratio above £20,000. Rawlins et al. (2010) note that in these circumstances, where stronger evidence is required for approval, the impact of technologies on disadvantaged populations likely to be in worse health (such as those on low incomes or ethnic minorities) must be taken into account by the Technology Appraisal Committees (TACs).

At the same time, TACs must also consider the principles outlined in NICE Social Value Guidance (NICE 2008a) when making an approval decision. One such principle is that recommendations that restrict the recipient population in terms of their characteristics, including age, race, gender, disability or socioeconomic status, can only be made in special circumstances (NICE 2008a, p.25). Thus, NICE committees must look more favourably upon group-level interventions benefitting

the disadvantaged without discriminating against them in the delivery of individual services. Another aspect that is missing from the current approach is how health inequality impacts are reflected in the health opportunity costs of approval decisions. TACs are instructed to look kindly on interventions that benefit those likely to be in poor health without consideration or knowledge of how those populations are affected by the displacement of other services. Without knowledge of how the opportunity costs are distributed, the net effect of interventions on reducing health inequalities will remain unknown.

This somewhat incomplete and informal approach toward incorporating inequality concerns in NICE processes has been labelled unsatisfactory by commentators (Lindholm & Rosen 1998; Williams 1997; Cookson et al. 2009). Greater focus has generally been placed on other equity characteristics related to clinical need such as access to care, severity of illness (Shah 2009; Green 2009; Brazier et al. 2013) and treatments at the end of life (NICE 2009; Collins & Latimer 2013; Shah et al. 2014). Only the latter are given formal weighting during NICE appraisals: if a treatment increases mean overall survival by more than three months to patients with mean life expectancy of less than six months, a higher cost-effectiveness threshold of £50,000 per QALY is used. Consideration of other equity aspects may therefore be ad hoc, inconsistent and opaque (Baltussen & Niessen 2006; Goetghebeur et al. 2008). This is contrary to its commitment to the 'Accountability for Reasonableness' framework (Daniels 2000) that demands complete transparency for legitimacy, leading Shah to conclude that there is little guidance to TACs about how to approach potential trade-offs between equity and efficiency (Shah et al. 2011, p.11), despite the clear importance of the former.

1.1.3 Public Health

In addition to the provision of health care, NICE has since 2005 also played a role in public health by providing guidance on cost-effective interventions (NICE 2012b). These include programmes intended to change lifestyle factors, such as smoking cessation and increased physical activity, which come closer to the more 'upstream' factors that can potentially have a significant role in reducing health inequalities (Weatherly et al. 2009; Whitehead & Ali 2010; Shah et al. 2011). This is reflected in

public health agencies across the world placing great emphasis on this objective (Raphael 2000; Norwegian Ministry of Health and Care Services 2007; PHAA 2012), with the Department of Health in the UK making it a key indicator in its Public Health Outcomes Framework (DH 2013). As with health care however, no efforts are made to precisely estimate the potential impacts on the level of health inequality. Weatherly et al. (2009) find that empirical literature on public health interventions rarely mentions inequality impacts, whilst health inequality audits (DH 2006) and impact assessments (Taylor & Blair-Stevens 2002), recommended by NICE in its inequality guidance to local authorities (NICE 2012a), are not quantitative, making it difficult to ascertain the consequences of policies.

The methodological challenges for the evaluation of public health interventions are different to those for health care, since the public health budgets are allocated to local authorities rather than CCGs. This raises issues similar to education or tax policies that impact on health inequality, which are beyond the scope of the current methodology and indeed this review. Regardless, public health is an important tool in tackling unfair health inequality and more research is needed on how it can be best utilised.

1.1.4 Discussion

This section summarises how the UK approaches the problem of unfair health inequality. Whilst the topic of this thesis is how this can be done through health care provision, it is acknowledged that policies implemented from outside the health sector that target education or income can be equally, if not more effective at reducing disparities. Public health is another avenue that could be explored, but both require different methodological approaches due to how they are financed.

Given that NICE guidance could potentially ensure the provision of inequality-reducing technologies, economic evaluation remains an under-utilised resource for combating inequality. This is not a phenomenon limited to the UK, with HTA guidance in both Australia and Canada avoiding any formal implementation of inequality concern in the evaluation process (Department of Health and Ageing 2009; CADTH 2006). The reluctance to do so likely stems from contentious issues

related to distributive justice. Introducing a framework that is explicit and transparent with regard to inequality concerns requires a commitment to certain social value judgements about what sources of inequality are unfair.

The effect of these issues is that the distributional impacts of health technologies, despite their importance to decision-makers, remain unclear. In order for the additional resources allocated to more deprived areas to have the desired effect, decision-makers require an alternative evaluative process that informs them as accurately as possible as to how treatments affect the distribution of health.

1.2 Frameworks for evaluating inequality effects in the health sector

Inequality concerns, along with other aspects of equity, have been at the heart of the development of alternative methodologies in health economic evaluation. This section reviews the frameworks that have been proposed to address the deficiencies of CEA along with limitations and potential for practical use. These are (i) dashboard approaches; (ii) equity weighting; and (iii) social welfare analysis.

These frameworks can be applied using linear programming techniques as well as through decision analytic modelling. Linear programming employs optimization algorithms to maximise an objective function, which under a health maximisation approach would be the sum of QALYs subject to a budget constraint. Equity can be incorporated either as a constraint on this function or incorporated in to it directly. Examples of constraints include providing equal treatment for similar patients (Epstein et al. 2007) or providing a minimum health benefit to all patients (Cleary et al. 2010; Earnshaw et al. 2002).

1.2.1 'Dashboard' approaches

Dashboard approaches to evaluation reflect the reality that numerous criteria are important in the decision-making process. The three methods that fall under this banner differ in the methods they use to estimate each criterion and whether they are synthesized into an overall score.

1.2.1.1 Cost-consequence analysis

Cost-consequence analysis (CCA) is the methodologically simplest framework. Coast (2004) argues that this is a strength of CCA when compared to the complex 'black box' methods used in cost-effectiveness analysis. In a CCA, the effects of a new intervention on a range of qualitative and quantitative criteria are estimated by researchers and presented to the decision-maker, who implicitly weights each criterion. Mauskopf et al. (1998) acknowledge that this process is potentially vulnerable to personal biases due to the imprecise nature of the criteria and opacity over how criteria are weighted.

Distributional impacts can potentially be incorporated into CCA, and are put forward as a criterion by Trueman and Anokye (2013). Burger et al. (2010) do so in their evaluation of outreach strategies for pregnant women, noting the programmes targeted and improved outcomes for those with low incomes or in ethnic minorities.

However, reducing the complexity of the process inherently reduces the strength of the evidence for or against the adoption of a technology. Furthermore, given the qualitative nature of any equity criteria and the lack of opportunity costs in the analyses, net distributional impacts are likely to be imprecise and uncertain. However, when a decision-maker wishes to consider additional aspects of benefit, and data to support them are limited, a CCA may present as the most feasible option.

1.2.1.2 Multicriteria decision analysis

Multicriteria decision analysis (MCDA) is an extension of the CCA framework that utilises multi-attribute utility theory to analyse trade-offs between different measures and objectives. The first stage involves the selection of a set of relevant, comprehensive and independent criteria, using either the theoretical literature or focus groups. Next, interventions are modelled on the criteria, and in the third stage they are scored to create a 'performance matrix', then combined into an overall intervention score using a function that applies elicited weights to each criterion (for a more thorough explanation of MCDA processes, see Baltussen and Niessen (2006), Goetghebeur (2008) and Thokala and Duenas (2012)).

Johri and Norheim (2012) recommend MCDA as the most promising avenue for incorporating equity concerns into economic evaluation. As with CCA, a criterion can be included to reflect how the intervention impacts upon disadvantaged populations, as recommended by Baltussen et al. (2013) and included by Peacock et al. (2007). Thus far however, measurement of distributional impacts has been simplistic or qualitative.

MCDA is presently a framework in development. Many issues remain over the selection of criteria, as a recent review by Guindo et al. (2012) highlights. Thokala and Duenas (2012) point to several more unresolved problems. It is not clear whether a consistent set of criteria ought to be applied across interventions, and there are no accepted methods for scoring criteria on appropriate scales. Even the most widely applied MCDA framework, EVIDEM (Goetghebeur et al. 2010; Tony et al. 2011; Goetghebeur et al. 2012; Miot et al. 2012), uses a simple stated preference method to elicit weights and scores. This increases the likelihood of personal biases when compared to more robust experimental methods such as discrete choice experiments. Lastly, opportunity costs would need to be estimated for every criterion when applied to health systems with fixed budgets – a challenge not yet undertaken by any study.

1.2.1.3 Extended cost-effectiveness

The final 'dashboard' approach is extended cost-effectiveness analysis (ECEA). Formulated by Verguet et al. (2013), ECEA is specifically intended for evaluating health interventions for developing countries. Interventions are evaluated not only in terms of health outcomes (i.e. number of deaths) but also on the level of financial risk protection afforded and the household expenditures avoided. Furthermore, the impact on each criterion is presented by wealth quintile, providing clear and explicit information on the distributional impacts. For example, in an analysis of providing UHC for tuberculosis in India, the authors find that 80% of deaths averted are in the bottom two income quintiles (Verguet et al. 2014).

As ECEA is applied to developing countries where there is no fixed budgetary constraint, consideration of opportunity costs is not restricted to health. Whilst the funding of new interventions will displace some public and private health care, opportunity costs will also fall on other public services and private consumption.

Each of these facets of opportunity cost will have distributional consequences for health and income that should be taken into account. For example, the analysis of tobacco taxation policy in Lebanon by Salti et al. (2016) quantifies the tax revenue by income quintile, but does not calculate the net effect by estimating the distribution of benefit (health or otherwise) from any consequent public service provision. Another potential problem is that ECEA does not offer any way in which to trade off between criteria. Nevertheless, given its recent creation, ECEA is a promising framework that uses sophisticated techniques to model distributional impacts.

1.2.2 Equity weighting

The idea of weighting the benefits of an intervention according to the characteristics of the recipients is the most intuitive way to reflect the principle that the social value of those benefits differs between individuals. It is an approach that has long been suggested in relation to both cost-benefit analysis (Harberger 1978) and cost-effectiveness analysis (Bleichrodt 1997; Bleichrodt et al. 2004), and weights have been applied a range of characteristics. The UK Treasury, for example, advocate weights on technical grounds, to reflect diminishing marginal utility (HM Treasury 2011). Other studies, including Nord (1993), Nord et al. (1999) and Brazier et al. (2013) have formulated weights to reflect severity of disease. A thorough discussion of all the arguments and principles that justify weighting certain characteristics, as well as the empirical evidence that supports them, is beyond the scope of this review but can be found in Sassi et al. (2001).

Two types of weights can potentially reduce unfair inequalities, each resting on different ethical arguments. The first is to differentiate between social groups such as socioeconomic classes, as suggested by Williams (1997), or occupational groups as advocated by Lindholm et al. (1998). Giving greater weight to the benefits accruing to the least-healthy groups will then reduce unfair health inequalities, given that such differences are unjust to start with (see Marmot (1997)). The second type is weighting benefits by age. This requires that the health variable we are trying to reduce inequality in is lifetime health, and has come to be known as the ‘fair innings’ argument (Williams 1997). Williams (2001), supported by Dolan and Olsen (2001) and Tsuchiya (2000), argues that priority should be given to the young so that they

are given the best opportunity to have a ‘fair innings’ (for example, 70 QALYs). Implementation of age weights would thereby reduce these inequalities in lifetime health. Bleichrodt et al. (2004) advocate a different ethical position, proposing a system of weights that depend only on the rank of the individual in the distribution of health.

The implementation of equity weights is far from simple. First, as Johri and Norheim (2012) acknowledge, studies that attempt to elicit weights from a sample of the population, such as those by Nord (1999), Bleichrodt et al. (2005) and Maestad and Norheim (2009) can be subject to a number of experimental biases. Nord notes that the complexity of questions posed in the studies may also lead to inauthentic responses. Morton (2014), when analysing the equity impacts of treatment strategies for depression, avoids specifying a particular set of weights by conducting sensitivity analysis using a range of different weights using mathematical programming techniques. By running an optimisation over different combinations of group weights, he determines how frequently an intervention is the simulated optimal solution.

Numerous other problems are also identified by Wailoo et al. (2009). The authors point out that equity characteristics can change over the time horizon used in cost-effectiveness modelling, complicating matters significantly. Another issue relates to the multitude of equity characteristics and how weights can reflect interactions between them. Most importantly, equity weights are not reflected in the health services displaced when funding interventions under a fixed budget, since the characteristics of patients using the displaced services are unknown. These criticisms are echoed by Bobinac et al. (2012) and Paulden et al. (2014), the latter with respect to the weights applied by NICE to end-of-life treatments. Recent work by Claxton et al. (2015) did account for opportunity costs by estimating a series of equity-weighted thresholds based on the burden of illness and wider social benefits of diseases that are related to health services displaced at the margin. These weights were not, however, adopted by NICE, which maintains that a QALY has the same weight for all population groups (NICE 2014).

1.2.3 The social welfare function approach and distributional cost-effectiveness analysis

The idea of using social welfare functions, usually applied in the income distribution literature, to health distributions was first proposed by Wagstaff (1991) and later developed by Dolan (1998) and Bleichrodt (2004). Health-related social welfare functions (HRSWFs), as they came to be known, are a particularly relevant tool for analysing health inequalities due to their ability to trade off between equity and efficiency concerns through a parameter that reflects the level of inequality aversion. As Johri and Norheim (2012) state, HRSWFs evaluate health distributions by multiplying average health by a measure of inequality, such as the Atkinson index (Atkinson 1970) advocated by Lindholm and Rosen (1998) and Norheim (2013).

An example of the HRSWFs being applied to economic evaluation in health care was performed by Robberstad and Norheim (2011). Their study looked at two competing interventions: vaccinations for infants and hypertensive drugs for adults with reduced life expectancy. The impacts on lifetime health expectancy were modelled for each treatment option and the distributions were evaluated using HRSWF with inequality aversion.

Distributional cost-effectiveness analysis (DCEA), a framework proposed by Asaria et al. (2015), further developed this idea of modelling health distributions and evaluating them using HRSWFs. The evaluation process in DCEA involves five steps. First, a baseline distribution of health that reflects unfair inequalities in health is estimated using a metric consistent with economic evaluation such as the QALY or disability-adjusted life years (DALYs). Second, the effects of the intervention are modelled for each of the population groups used to construct the baseline distribution. The opportunity costs of the intervention are then estimated for each of the relevant population groups. Fourth, the effects and opportunity costs are modelled onto the baseline distribution to generate a post-intervention distribution that reflects the net health effects. Lastly, the post-intervention distributions can be evaluated using a HRSWF.

A problematic aspect of DCEA is the determination of the level of inequality aversion. As with the elicitation of equity weights and criteria weights in MCDA, normative judgements or empirical evidence could form the basis of the social value judgements required to trade off between equity and efficiency. The latter approach is exemplified by Robson et al. (2016), who conducted face-to-face and online elicitation exercises on the UK general population and found that health gains for the poorest fifth of the population were weighted between six and seven times more highly than the richest fifth. However, despite re-weighting the samples to reflect the makeup of the general population, studies of this type may not accurately reflect the true population preference. First, the level of inequality aversion expressed by individuals may be influenced by the way in which the exercise is conducted or the phrasing of the questions. Ali et al. (2017) conclude that these 'framing' effects can considerably alter the level of inequality aversion but do not eliminate it altogether. Second, selection bias may still be present if there is a systematic difference between participants and non-participants with otherwise similar characteristics.

At present, DCEA provides a more robust methodological basis for incorporating inequality concerns than the alternative frameworks included in this review. Utilising heath expectancy techniques to estimate a baseline distribution of health could provide a quantitative, evidence-based picture of health inequalities and a means of evaluating changes in inequality expected from funding an intervention. To enable these analyses, DCEAs do require a lot of data on the variation of model inputs by equity-relevant variables that may be not be widely available if patient-level data are not used. However, unlike some of the other methodologies looked at in this review, it also provides a mechanism for incorporating the opportunity costs of any purchasing decision, an essential component of ascertaining the net distributional effect.

1.2.4 Discussion

This review identifies three principal approaches to incorporating inequality concerns into health economic evaluations: dashboard approaches, equity weighting and social welfare analysis. Whilst these have been presented independently, many aspects of the approaches are interlinked and rely on similar assumptions.

Multi-criteria decision analysis represents a decision-making framework that encompasses most of the approaches described. ECEA, for instance, is a form of MCDA in which the separate criterion are not weighted and formally traded-off. DCEA is also a form of MCDA in which a distributional criterion is added to cost-effectiveness results, and where the weights between these criteria are determined by the inequality aversion parameter.

DCEA and equity weighting are similarly related. Whilst the former has the additional steps of modelling net health impacts onto health distributions, the way in which these distributions are evaluated is determined by applying differential weights to individuals. The principal difference is that the social welfare functions used in DCEA, such as the Atkinson or Kolm indices, weight individuals on their rank in the health distribution rather than on another socially relevant criterion.

Two key considerations emerge across the discussions of all frameworks. The first is the role of health opportunity costs, which are not dealt with in any methodological rigour in any of the frameworks, with the exception of DCEA. This is a crucial shortcoming that must be addressed in order for a framework to adequately evaluate whether interventions will improve or exacerbate health inequality. This importance can be shown in a simple example in which society has two groups, healthy and unhealthy. An intervention imposes health opportunity costs of 50 and 100 QALYs and health benefits of 100 and 125 QALYs to the healthy and unhealthy groups, respectively. This intervention, despite having a pro-poor distribution of benefit, would still increase health inequality. Thus, the distribution of opportunity costs is a necessary requirement to determining health inequality impacts.

The second common consideration is that each framework, with the exception of ECEA (where inequality and total gain are not formally traded off), includes at least one normative parameter that cannot be scientifically determined. In DCEA this is the inequality aversion parameter used to conduct social welfare analysis, whilst in MCDA it is the weights assigned to the competing criteria. This reflects the nature of incorporating an aspect of equity into a quantitative framework. The defensibility of multiple reasonable positions on issues of equity have so far inhibited the progress of

alternative frameworks to CEA emerging in health technology assessment. However, the act of making them explicit is also an appealing argument for adopting such frameworks, as they make the value judgements underpinning decisions more transparent than the ad hoc deliberations currently in place.

Chapter 2: Estimating the social distribution of health in England

2.1 Introduction

The prevalence of many diseases and illnesses differ by income, gender and race. While some of these will shorten lifespan, others will limit our ability to function and flourish in life. Both aspects are central to our health experience, and it is therefore crucial that when we estimate inequalities of health between social groups, the differences in both quality and quantity of life are accounted for. There are various ways of summarising a population's overall lifetime experience of health by combining information on both mortality and morbidity into a single figure, generating an estimate of 'health expectancy'. Perhaps the best known metric is disability-free life expectancy (DFLE), which subtracts years from life expectancy using a binary indicator of ill-health or disability.[1]

Quality-adjusted life expectancy (QALE) is a more recent approach to estimating health expectancy that uses a continuous ratio scale variable to measure morbidity, thus enabling it to incorporate detailed multi-attribute data on health-related quality of life. The rising popularity of the quality-adjusted life year (QALY) metric, through its use in health technology assessment, has led to the inclusion of preference-based heath-related quality of life (HRQL) questionnaires in national health surveys, affording researchers the opportunity to estimate QALY weights for a wide range of population subgroups using large, nationally representative datasets. However, implementation of the QALE metric in health inequality research has been limited to regional analyses (Collins 2013; Collins 2017), despite widespread application of other health expectancy indicators to inequality measurement (Bajekal 2005; Wood et al. 2006).

[1] The term health-adjusted life expectancy (HALE) is also widely used in the literature but has been applied to DFLE (Collins 2013), disability-adjusted life expectancy (Mathers et al. 2001) and QALE (Manuel & Schultz 2004). To avoid such ambiguities, this paper will not use the term HALE in describing health expectancy measures.

As well as health inequality measurement, the baseline social distribution of health that is estimated in this chapter can be used to model the distributional impacts of new health care and public health interventions in distributional cost-effectiveness analysis (Asaria, Griffin, Cookson, et al. 2015; Asaria, Griffin & Cookson 2015; Cookson et al. 2017). Doing so will enable decision-makers to (i) gauge how health inequality is expected to change if the intervention was funded; (ii) determine whether a trade-off might exist between maximising population health gain and reducing health inequality; and (iii) evaluate, if a trade-off does occur and there is a degree of inequality aversion, the joint impacts through social welfare analysis.

The aim of this study is to generate predictions of QALE for age, gender and socioeconomic groups using nationally representative survey data and mortality rates. By combining these with their respective population estimates, we then create a rank ordering of the population by QALE that reflects social inequalities in health, from which we can calculate social welfare indices. The merits of this analysis are three-fold. First, a QALE distribution will allow for the effects of health care and public health interventions on population health to be modelled directly using methods and metrics consistent with cost-effectiveness analysis. Second, using the QALY in population health measurement provides a richer measure of inequality by reflecting differences in both mortality and morbidity, which are both shown to be worse in lower socioeconomic groups (Cookson, Asaria, et al. 2016). Third, the social welfare results indicate, for a given level of inequality aversion, the amount of average health that society would be willing to give up in order to obtain an equal distribution.

2.2 Overview of health expectancy

Health expectancy metrics all follow a similar methodology and differ in terms of the way that morbidity is incorporated. They can be divided into five principal types, which are listed in Table 2.1.

The first stage common to all health expectancy metrics is to construct life tables using national age-specific mortality rates to generate predictions of life expectancy (Chiang 1972). The life tables are then adjusted by whatever measure of morbidity is

employed, using a method originally proposed by Sullivan (1971). For DFLE, morbidity is a binary measure indicating the presence or absence of disability or limiting long-standing illness, estimated at the population level through prevalence rates estimated from survey or administrative data.

Developments in HRQL measurement have provided more sensitive instruments to quantify individuals' health status. Disability-adjusted life years (DALY), developed by Murray (1994), attach pre-determined weights to each life year associated with a particular disability; combining the weights of each disease and their prevalence rates with life tables using the Sullivan's method creates an estimate of disability-adjusted life expectancy (DALE). In doing so, each life year lived with a disability or chronic disease is adjusted for severity rather than being excluded entirely as with DFLE.

Table 2.1: Taxonomy of health expectancy metrics

Metric	**Morbidity measure**	**Example**
Life expectancy	None	White & Butt (2015)
Disability-free life expectancy	Disabled/limiting illness – yes/no (2 states)	ONS (2013)
Healthy life expectancy	Self-assessed good health – yes/no (2 states)	Wood et al. (2006)
Disability-adjusted life expectancy	Prevalence of a disease or disability and associated disability weights to generate DALYs (220 states + sequelae)	Salomon et al. (2012)
Quality-adjusted life-expectancy	Generic multi-dimensional self-reported health questionnaire and associated quality weights to generate QALYs (EQ-5D-3L – 245 states, SF-6D – 18,000, HUI3 – 972,000[2])	Collins (2013)

In this study we adopt an alternative morbidity measure that has been developed for the economic evaluation of health technologies and public health interventions: the quality-adjusted life year (NICE 2014). Health-related quality of life weights are constructed from patients presenting illness and receiving treatment, who complete multi-item questionnaires on their own health at different points in time. The questionnaires cover a range of dimensions important to individual health, each with a range of severity levels, and are known as generic instruments as they are not specific to any particular clinical area, enabling the effectiveness of treatments in

[2] See Rabin & de Charrao (2001) for more information on EQ-5D-3L, Brazier et al. (2002) for SF-6D and Feeny et al. (2002) for the HUI3.

different clinical areas to be compared (Drummond et al. 2005). Each permutation of responses represents a health state and is assigned a pre-determined utility score from a tariff elicited from the general population, which are anchored at zero (for health states as bad as death) with an upper limit of 1 (for full health with no problems in any dimension of health). Using a variation on Sullivan's method outlined in Roset & Gaminde (1999), health-related quality of life (HRQL) weights are used to adjust life years for morbidity.

The application of health expectancy measures to socioeconomic inequalities in health has been highlighted as potentially one of its most promising and fruitful uses (Robine et al. 1999; Roset & Gaminde 1999). In Valkonen et al. (1997), Finnish life tables are constructed by level of education by linking death records to census data. Using several different definitions of disability and illness, the authors estimate gaps ranging from 7.3 to 13.1 healthy life years between high and low education groups. Bajekal (2005) compares healthy life expectancy (see Table 2.1) and DFLE between socioeconomic groups using mortality rates and deprivation scores for 8,595 electoral wards in the UK. The disparities between the least and most deprived decile groups were estimated at 16.9 healthy years (based on prevalence of self-assessed good health) and 12.4 disability-free years (based on prevalence of disability/long-standing illness) for males, and 16.8 and 9.9 for females, respectively. Wood et al. (2006) also look at both DFLE and healthy life expectancy between deprivation groups, this time in Scotland, using self-reported long-standing illness statistics from census data. They estimate a difference of 13.6 healthy years of life between the least and most deprived groups.

The extension of these methods to QALE is scarce. Jia et al. (2011) analyse gender and racial differences in QALE at age 18 in the US. White women had the highest QALE of 54.1 QALYs, compared with black men at 46.1 QALYs. Only Collins (2013) has focused specifically on QALE differences between socioeconomic groups. Using EQ-5D data from a special one-off regional survey, the author estimates QALE for the least and most deprived areas of the Wirral area in England (judged by national area-level deprivation score) and compares them with HLE estimates based on a binary self-assessed health variable. The study emphasizes the importance of evaluating morbidity and mortality together using a robust morbidity

indicator. A gap of 12.7 QALYs between those falling in the most and least deprived national quintile groups is estimated, compared with 8.1 life years and 14.1 healthy years, which under-estimate and over-estimate morbidity effects respectively.

2.3 Methods

Our analysis consists of five stages. First, using data in the Health Survey for England (HSE), we predict HRQL weights as a function of age, gender and socioeconomic status. Second, predictions of life expectancy are generated from national mortality data for age, gender and socioeconomic groups using life tables. The life tables are then adjusted using the HRQL estimates to create respective predictions of QALE. After the population estimates for each group are combined with their respective QALE to create the social distribution of health, we compute indices that evaluate social welfare in terms of both total health and health inequality. A worked example demonstrating stages one to four is shown in Figure 2.1.

2.3.1 Data and variables

The analysis uses pooled data from the three recent rounds of the Health Survey for England (HSE) in 2010, 2011 and 2012, with a combined sample size of 35,062. The HSE is an annual series that monitors a range of health conditions and risk factors for the non-institutionalised population. It uses a multi-stage stratified probability sampling design with a sampling frame of Postcode Address File that tries to ensure every member of the population has an equal chance of being selected. Details of the sampling methodology are in Boniface et al. (2012).

Health status is measured using the EQ-5D questionnaire (Rabin et al. 2011), a generic instrument used in HTA around the world to assess the treatment effects of interventions for a wide range of different health conditions (NICE 2004; Rabin & de Charrao 2001). The EQ-5D is a questionnaire that asks respondents to rate their own health in five dimensions: pain, mobility, anxiety/depression, self-care and usual activities. In the original EQ-5D-3L version used in this study, subjects rate their health on each dimension using one of three possible levels: no problems, some problems, or severe problems. This generates a possible 245 health states when

including the two additional states 'unconscious' and 'dead'. A single index figure is then given to each health state based on a country-specific tariff. The standard UK value set estimated by Dolan et al. (1995) was applied to our data.[3] This analysis is restricted to adults aged 16 and over, leaving a sample size of 25,320. This is due to the fact that EQ-5D is not responsive to the HRQL for children under this age, for whom there are other more appropriate instruments (Wille et al. 2010).

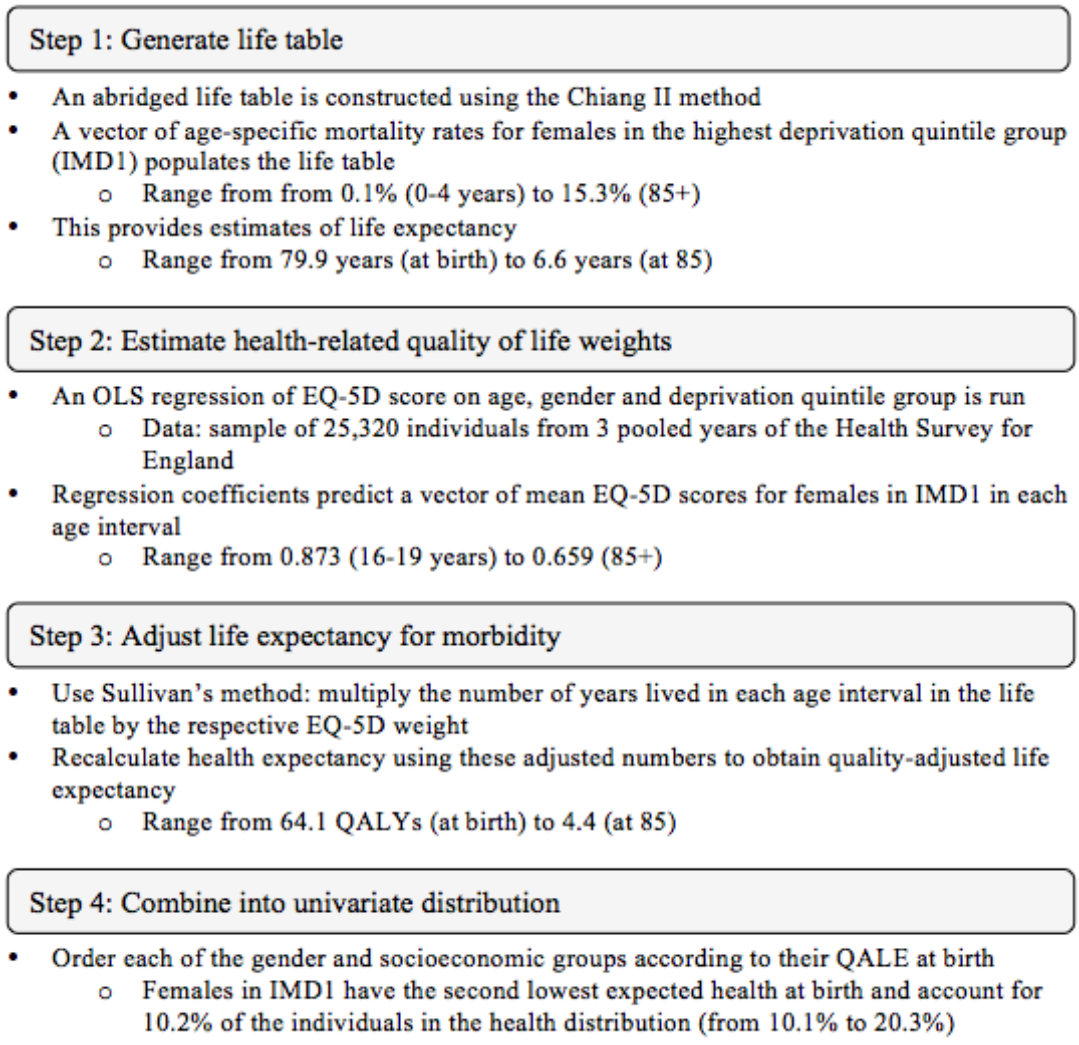
The socioeconomic variable we use is the Index of Multiple Deprivation (IMD) from 2010. This is a weighted area deprivation index of 38 variables covering seven dimensions of deprivation (employment, income, education, health, crime, living environment and housing/services) that is given to each of the 32,482 Lower Layer Super Output Areas (LSOAs) in England[4]. In 2010, the median LSOA population was 1,551 with an inter-quartile range of 1,429 to 1,708 and 99% had fewer than 2,731 residents. More information on the methods used to construct the IMD can be found in McLennan et al. (2011). The raw IMD score is not reported in the HSE, thus the variable used in the regression analysis is population IMD quintile group, with the first quintile group representing the most deprived and those in the fifth having the lowest deprivation.

We focus on age, gender and socioeconomic status as covariates, as these are often of interest in public health campaigns and are associated with large inequalities in population health. An additional advantage of using this set of variables is that they can potentially be collected in clinical trials. Whilst age and gender are routinely collected in most studies and surveys, there is now increasing interest in collecting equity-relevant data such as socioeconomic status (Mbuagbaw et al. 2017; Jull et al. 2017). Doing so will improve the feasibility of DCEA by enabling more routine estimation of equity-informative cost-effectiveness evidence.

[3] This value set, although nearly twenty years old, remains the largest and most representative source of UK survey data for estimating EQ-5D weights. The authors used the time trade-off method on a random sample of 3395 people from the adult UK population to elicit the value weights for 45 states and using econometric modelling to predict the rest.

[4] Assignment of domain weights were "driven by theoretical considerations and responses to the consultation processes" (McLennan et al. 2011, p.17).

Figure 2.1: Calculation of quality-adjusted life expectancy of females in the highest deprivation group



Note: IMD = Index of Multiple Deprivation; OLS = ordinary least squares; QALYs = quality-adjusted life years; QALE = quality-adjusted life expectancy

2.3.2 Regression Analysis

The distribution of EQ-5D utility score is heavily skewed: the proportion of individuals reporting severe problems on any of the dimensions is rare, ranging from 0.18% for mobility to 4.29% for pain; whilst the number reporting perfect health is over half, at 52.72%. Additionally, the utility data have an upper ceiling of 1. Whilst these properties suggest that a linear regression model may not be appropriate, we employ Ordinary Least Squares (OLS) as our estimator for two principal reasons. First, previous studies have shown OLS to perform well in comparison with other

types of estimator when used to model HRQL, particularly when using large sample sizes like those in the HSE (Petrou & Kupek 2008; Maheswaran et al. 2013; Vogl et al. 2012). Second, the principal diagnostic instrument for judging accurate HRQL prediction is accurate mean EQ-5D scores for age-sex-socioeconomic groups, since it is these that are used to adjust the life tables in the QALE process described below. This means that any potential imprecision of individual predicted scores caused by applying a linear model is not a cause for concern, so long as the mean group scores exhibit good fit.

Using OLS also has an additional benefit, in that the estimated coefficients can be directly interpreted and utilised to predict EQ-5D (and therefore QALE) for different populations than the one used in this study. Alternative models are run to determine the model specification, one with age-socioeconomic status and age-gender interaction terms and a second with a quadratic age term to test for non-linearity with respect to utility. We also perform sensitivity analysis using alternative two-part and Tobit models, described in section 2.3.7.

All statistical analyses are performed in Stata 12. Standard survey data analysis tools are used to (i) incorporate the probability weights supplied in the data that reflect imbalances between the sample population and the general population; and (ii) to account for the fact that scores within households – the primary sampling unit – can be correlated. Not accounting for the latter can distort statistical inference by reducing the standard errors.

Another issue was missing HRQL data, with significant item non-response occurring within the sample. A total of 3,177 (12.6%) observations were missing a utility score, with these individuals on average tending to be older, male, non-white and living in more deprived areas than the complete cases. A logit model regressing the probability of missingness on our variables of interest was used to determine whether the data are Missing Completely at Random (MCAR) – that missingness is not systematic or related to individual characteristics (Little & Rubin, 2002). This found that age, race, gender and deprivation level are all statistically significant predictors of missingness ($P<0.01$), correctly predicting 86% of cases with missing values. These results are summarised in Table 2.2.

Table 2.2: Results from logistic regression estimating the effect of social characteristics on the probability of EQ-5D score being missing

Variables	Logit coeff miss_eq5d	Odds ratio miss_eq5d
Age	0.0142***	1.014***
	(0.00106)	(0.00108)
Sex		
Male	Ref	Ref
Female	-0.0863**	0.917**
	(0.0388)	(0.0356)
Race		
Non-white		
White	-0.574***	0.563***
	(0.0581)	(0.0327)
IMD Quintile Group		
1 (most deprived)	Ref	Ref
2	-0.156***	0.856***
	(0.0576)	(0.0493)
3	-0.469***	0.626***
	(0.0608)	(0.0380)
4	-0.552***	0.576***
	(0.0616)	(0.0355)
5 (least deprived)	-0.644***	0.525***
	(0.0621)	(0.0326)
Constant	-1.779***	0.169***
	(0.0745)	(0.0126)
Observations	25,230	25,230

Notes:

1. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are in parentheses.
2. IMD = Index of Multiple Deprivation
3. Of the 35,062 in the original sample, 25,320 were 16 and over and could be included in the analysis.

2.3.3 Life tables

2.3.3.1 Life expectancy

Mortality data are acquired from a bespoke data extraction from the Office for National Statistics. This dataset contains, for 2012, population estimates and number of deaths by age (5 year bands), gender and IMD quintile group. Crude mortality rates for then calculated for each group, which are then used to construct ten abridged life tables using the Chiang II method (Chiang 1972) to obtain estimates of life expectancy for each of the 180 age-gender-socioeconomic groups (two genders,

five IMD groups, 18 age intervals). Crucially the mortality rates are used to calculate the probability that, conditional on them surviving up to the start of an age interval, an individual will die during it. This quantity, q_x, is given by the formula:

$$q_x = \frac{n_x m_x}{1 + (1 - a_x) n_x m_x}$$

Where n_x is the number of years in age interval x, m_x is the mortality rate and a_x is the proportion of the age interval that individuals who die during the interval are assumed to have survived. The latter is set at 0.5 for all intervals. Life expectancy at the start of an age interval is estimated by dividing the number of years lived in that and all successive intervals by the number of people alive at the beginning of the interval:

$$e_{xds} = \frac{\sum_x^z L_{xds}}{I_{xds}}$$

Where e_{xds} is life expectancy at the start of age interval x for deprivation quintile group d and gender s; z is the last age interval; L_{xds} is the total number of years lived by the surviving cohort in interval x; and I_{xds} is the surviving cohort at the start of the interval. The life tables provide a snapshot of health for a given year for each social group (the "period" approach) rather than a prediction of lifetime health experience (the "cohort" approach). This means that our results do not account for the fact that individuals' socioeconomic status may change over the life course.

2.3.3.2 HRQL adjustment

Life expectancy is then adjusted for HRQL using the predicted utility scores for each age-sex-socioeconomic group, via the Sullivan method (Sullivan 1971). Since we are unable to estimate HRQL for people aged 0-15, we assume that they experience the same average HRQL as those in the youngest age group for which they could be estimated (16-19 years). Obtaining the QALE estimate is nearly the same as for life expectancy; the difference being that we multiply the years lived in each age interval by the associated HRQL weight, u_{xds}:

$$QALE_{xds} = \frac{\sum_x^z L_{xds} * u_{xds}}{I_{xds}}$$

Additional life tables are constructed for further analyses: by IMD quintile group for each gender to enable comparisons with previous health expectancy studies, and by IMD quintile group combining genders to enable a non-gender-specific socioeconomic inequality estimate.

2.3.4 QALE

We analyse inequalities in QALE in three ways. First, bivariate distributions of QALE by socioeconomic status and gender are generated so that these inequalities can be analysed separately from the supplementary life tables just described. Second, an overall univariate distribution of health is constructed that reflects both types of disparities. This is done by first assuming that the prediction of QALE at birth is the same for all individuals within a gender-socioeconomic group (regardless of their age). We can then multiply each of the 20 gender-socioeconomic group QALE predictions by the number of people in the group, taken from ONS population estimates, and rank the whole population from lowest to highest QALE. The additional benefit of this distribution is that we account for the relative sizes of the gender-socioeconomic groups as well as the magnitude of the inequalities between them. Third, we compute the distribution of quality-adjusted age at death (QAD). This metric, as advocated by Gakidou et al. (2000), shows the distribution of expected lifetime QALYs experienced by the population. Assuming that individuals die at the midpoint of the age interval we calculate QAD for each age-gender-socioeconomic group by summing the QALYs in each age interval:

$$QAD_{xds} = \sum_{y=1}^{x-1} u_{yds}\, n_y + u_{xds} n_x a_x$$

2.3.5 Social welfare analysis

We can combine the amount of total population health and health inequality into a single index measure of health related social welfare by using a social welfare

function that describes the amount by which the least healthy should be prioritised for health improvement. An inequality aversion parameter assigns an implicit weight to individuals based on their health level relative to the mean. The level of inequality aversion determines the trade-off between improvements in total health and reductions in inequality, and allows the calculation of a social welfare index to summarise whether an intervention improves social welfare. We use the Atkinson and Kolm indices to measure social welfare using the univariate distribution of health. The Atkinson index, A_ε, measures inequality relatively and is given by:

$$A_\varepsilon = 1 - \left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{Q_i}{\bar{Q}}\right)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}$$

Where N is the total population, Q_i is the QALE estimate of the ith individual, $\bar{Q}$ is the mean QALE and ε is the inequality aversion parameter that quantifies the concern for relative inequality. Alternatively, the Kolm index, K_α, incorporates inequality on an absolute scale, where absolute inequality aversion is represented by the parameter α:

$$K_\alpha = \left(\frac{1}{\alpha}\right)\log\left(\frac{1}{N}\sum_{i=1}^{N} e^{\alpha[\bar{Q}-Q_i]}\right)$$

For both indices, a higher inequality aversion parameter indicates greater concern for the less healthy. Our analysis uses estimates of 10.95 for ε and 0.15 for α, based on a survey of the general public in England by Robson et al (2016). In this survey they asked the general public to choose between two interventions, where one increased total health more and the other offered less health but provided the greatest benefits to the poor. They varied the amount of difference in total health provided by the two interventions, and the extent by which the poor benefited relative to the rich, to identify the point where members of the public thought the interventions were equivalent in value. The results showed that the general public preferred an intervention with less total health but that reduced the health gap between the poor and the rich, allowing the inequality aversion to be estimated. Social welfare is

calculated by combining each index with the mean level of health in the distribution to obtain the 'equally distributed equivalent' (EDE) level of health:

$$EDE_{A,\varepsilon} = (1 - A_{\varepsilon})\bar{Q}$$
$$EDE_{K,\alpha} = (\bar{Q} - K_{\alpha})$$

Where $EDE_{A,\varepsilon}$ and $EDE_{K,\alpha}$ are the Atkinson and Kolm welfare scores, respectively. The equally distributed equivalent is the level of mean health (expressed in QALYs) in a completely equal distribution that yields an equivalent amount of social welfare to the distribution being evaluated. The difference between the EDE and $\bar{Q}$ therefore indicates the amount of average health that society would be willing to trade-off in order to obtain a perfectly equal distribution.

2.3.6 Uncertainty

A Monte Carlo simulation is performed to account for uncertainty over the two sets of parameters in the model: probability of death and utility scores. Standard errors for the former, $\sigma_{q_{xds}}$, are calculated using the following formula proposed by Chiang (1983):

$$\sigma_{q_{xds}} = \sqrt{(q_{xds}^2(1 - q_{xds})/D_{xds}}$$

Where is D_{xds} is the number of observed deaths in the respective age-gender-socioeconomic subgroup. A Cholesky decomposition is used to take correlated random draws from the regression coefficient distributions when simulating utility scores. 1,000 simulations are performed, from which standard errors and 95% CIs are constructed for QALE at birth for both genders. Subgroup distributions can also be estimated by reconfiguring the base year, from at birth to any point at life, depending on the population of interest. We re-estimate the distribution at 25, 40 and 65 to demonstrate this.

2.3.7 Sensitivity Analysis

Two alternative types of estimator are adopted to check the robustness of the OLS findings. These more complex model specifications allow for the skewed and bounded nature of the outcome variable. Two-part models (2PM) and tobit estimators (Tobin 1958) were selected as both have previously been advocated in favour of OLS when handling HRQL data (Austin et al. 2000; Basu & Manca 2012).

Both estimators calculate the expected value of HRQL using a similar underlying hypothesis. The tobit model assumes that the observed health status score, y, has been censored and that an underlying latent variable, y^*, is the true health status score, such that:

$$y = \begin{cases} 1 & if\ y^* > 1 \\ y^* & otherwise \end{cases}$$

It then uses an indicator function to determine the probability that an observation has been censored and uses maximum likelihood to estimate health status conditional upon not being censored. Both parts are estimated using the same covariates: age, sex and socioeconomic status. By comparison, the two stages of the 2PM can be specified separately. The first part is a binomial logit regression that regresses a dummy variable indicating whether EQ-5D equals one on a set of covariates including smoking status, employment status and whether a long-standing illness is reported. The second part is the OLS regression model used in the base case analysis, but now performed on the subset of observations where utility is less than one. Thus both the tobit and the 2PM estimate the probability of being in perfect health and the conditional expectation of HRQL given imperfect health. These are then combined in the same way to get the expected value of utility:

$$E(y|x) = [P(y = 1).1] + [P(y < 1).E(y|y < 1)]$$

Judgement on the suitability of the estimators is based on several diagnostic measures. Mean-squared error (MSE) and mean absolute error (MAE) are computed

for each estimator, with MAE also reported for the mean group scores and for the 360 QALE predictions. Whilst these might not be conclusive as to which estimator is the most appropriate, they should give some indication as to how well they model EQ-5D and QALE for the purposes of this study. We performed these estimations on the complete case data. This enables us to compare their relative performance with OLS without the additional complications of using the imputed dataset, particularly for the 2PM. A complete list of variables used in each regression model is reported in Table 2.3.

As a further sensitivity analysis, we change the socioeconomic variable used in the regression analysis predicting EQ-5D to NS-SEC category, which classifies individuals into 8 groups based on their occupation (Class I being the highest socioeconomic group and VIII the lowest). This is to validate the legitimacy of using IMD as our socioeconomic variable. Since the age-specific mortality data is by IMD quintile groups and not NS-SEC group, it is necessary to create a mapping between the two so that the mean EQ-5D scores by the latter (8 categories) could be applied to the life tables partitioned by the former (5 and 10 categories). This mapping, along with descriptions of the NS-SEC occupation categories, is reported in Table A2.9.

Table 2.3: Variables used within each regression model

Model	**Variables**
OLS & Tobit	Age (C), Gender (B), IMD Quintile (M5)
OLS (NS-SEC)	Age (C), Gender (B), NS-SEC (M8)
2PM – Part 1 (logit)	Age (C), Gender (B), IMD Quintile (M5), Recorded Heart Attack (B), Diabetic (B), Long-standing Illness (B), Smoker (B), Obese (B), Unemployed (B)
2PM – Part 2 (OLS)	Age (C), Gender (B), IMD Quintile (M5)

Note:

1. Variable type in parentheses: C = Continuous, B = Binary, M*n* = Categorical with *n* categories
2. IMD = Index of Multiple Deprivation; NS-SEC = National Statistics Socioeconomic Classification; OLS = Ordinary Least Sqaures; 2PM = Two-part model

2.4 Results

2.4.1 Descriptive statistics

Descriptive statistics for a number of relevant subgroups are shown in Table 2.4. HRQL is correlated with age, gender and socioeconomic status, supporting our variable selection. EQ-5D scores declined with age; are higher for men; and decrease in a linear fashion as deprivation increased. The mean scores in the least and most deprived quintile groups are 0.88 in the least deprived quintile group to 0.79 in the most deprived.

2.4.2 EQ-5D prediction

Regression results using the imputed datasets are reported in Table 2.5. Compared with the complete case regressions, there is a smaller constant and smaller effect size for all covariates. Prediction performance also improves following imputation, with lower MAE for individual and group mean utility scores. Interactions between age, gender and socioeconomic status are not statistically significant and are excluded from the model. All other covariates are statistically significant ($P<0.01$). The signs on all coefficients are consistent with the descriptive relationships: lower deprivation is associated with higher HRQL, whilst ageing and being female are associated with lower HRQL. The quadratic age term is also negative, indicating larger HRQL decrements for each additional year. Mean group scores ranged from 0.98 (16-19 year-old males in IMD quintile group 5) to 0.68 (85+ females in quintile group 1). The full distribution of EQ-5D scores is provided in Table A2.10.

2.4.3 QALE

Adjusting life expectancy for HRQL has substantial impacts on inequality, shown in Table 2.6 and Table 2.7. The direction of this impact depends on which dimension of inequality we focus on. Adjusting for HRQL increases the absolute difference between the least and most deprived quintile groups from 6.5 years to 11.9 QALYs. Conversely, gender inequality is reduced, as the lower mortality for females is

partially offset by higher morbidity. Inequality at birth between genders drops from 3.6 years to 1 QALY. The standard errors generated from the simulation are also reported in Table 2.6 and Table 2.7. These average 0.06 life years and 0.29 QALYs across IMD and gender groups. The full range of QALE predictions are provided in Table 2.8.

Table 2.4: Sample statistics with average EQ-5D scores

Variable	**N**	**% Sample**	**Utility**	**Variable**	**N**	**% Sample**	**Utility**
Total	35,062	-	0.842	**IMD Quintile**			
Age				1st (most)	6,665	19%	0.793
0-15	9,742	28%	-	2nd	6,763	19%	0.826
16-24	2,555	7%	0.928	3rd	6,935	20%	0.839
25-34	3,642	10%	0.915	4th	7,060	20%	0.860
35-44	4,340	12%	0.877	5th (least)	7,639	22%	0.880
45-54	4,423	13%	0.844	**NS-SEC**			
55-64	4,077	12%	0.799	I	2,886	8%	0.905
65-74	3,434	10%	0.795	II	5,556	16%	0.871
75+	2,849	8%	0.723	III	3,580	10%	0.847
Gender				IV	2,160	6%	0.839
Male	16,204	46%	0.856	V	1,791	5%	0.814
Female	18,858	54%	0.832	VI	4,440	13%	0.810
Race				VII	3,270	9%	0.784
White	30,617	87%	0.870	VIII	484	1%	0.792
Non-white	4,334	12%	0.840	Not available	10,895	31%	-
Not available	111	0%	-				

Note:

1. IMD = Index of Multiple Deprivation; NS-SEC = National Statistics Socioeconomic Classification
2. Percentages are rounded and may not exactly sum to 100.
3. Those aged 0-15 were assumed to have utility equal to those aged 16-19.

Table 2.5: Results of regressions predicting EQ-5D scores

Variables	OLS (Imputed)	OLS (Complete)	Tobit	2PM	OLS (NS-SEC)
IMD Quintile Group					
1 (most deprived)	Ref	Ref	Ref	Ref	
2	0.0341***	0.0372***	0.0652***	0.0425***	
	(0.00650)	(0.00598)	(0.0117)	(0.00874)	
3	0.0475***	0.0512***	0.0881***	0.0629***	
	(0.00631)	(0.00585)	(0.0116)	(0.00855)	
4	0.0747***	0.0802***	0.148***	0.0954***	
	(0.00579)	(0.00549)	(0.0112)	(0.00806)	
5 (least deprived)	0.0857***	0.0937***	0.175***	0.108***	
	(0.00563)	(0.00539)	(0.0114)	(0.00782)	
Age	-0.00219***	-0.00200***	-0.00751***	-0.00477***	-0.00333***
	(0.000472)	(0.000416)	(0.000962)	(0.000654)	(0.000494)
Age-squared	-7.98e-06*	-1.55e-05***	2.87e-07	2.28e-05***	-9.13e-07
	(4.79e-06)	(4.29e-06)	(8.91e-06)	(6.24e-06)	(4.94e-06)
NS-SEC					
I					Ref
II					-0.0285***
					(0.00414)
III					-0.0457***
					(0.00519)
IV					-0.0550***
					(0.00645)
V					-0.0801***
					(0.00728)
VI					-0.0868***
					(0.00532)
VII					-0.106***
					(0.00628)
VIII					-0.146***
					(0.0176)
Gender					
Male	Ref	Ref	Ref	Ref	Ref
Female	-0.0210***	-0.0253***	-0.0590***	-0.0148***	-0.0199***
	(0.00324)	(0.00286)	(0.00629)	(0.00466)	(0.00316)
Constant	0.935***	0.943***	1.327***	0.798***	1.079***
	(0.0115)	(0.00947)	(0.0252)	(0.0164)	(0.0119)
Observations	25,320	22,143	22,143	10,469	21,252
MAE	0.156	0.155	0.155	0.135	0.155
Group MAE	0.0273	0.0245	0.0212	0.0213	0.025
QALE MAE	0.541	0.399	0.351	0.434	0.660

1. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are in parentheses
2. IMD = Index of Multiple Deprivation; NS-SEC = National Statistics Socioeconomic Classification; OLS = Ordinary Least Sqaures; 2PM = Two-part model; MAE = Mean absolute error
3. 25,320 are 16 and over. 22,143 had complete EQ-5D responses. 21,252 had complete NS-SEC data. 10,469 had EQ-5D less than one
4. MAE is the mean absolute distance across observations between the predicted and observed values

Table 2.6: Comparisons of absolute and relative inequality in life expectancy by Index of Multiple Deprivation (IMD) quintile group and gender

Life expectancy						
IMD quintile group	**Male**		**Female**		**Combined**	
1 (most deprived)	75.2	(0.061)	79.9	(0.060)	77.5	(0.044)
2	78.0	(0.061)	81.9	(0.056)	80.0	(0.041)
3	79.8	(0.058)	83.3	(0.054)	81.6	(0.040)
4	81.3	(0.056)	84.3	(0.052)	82.8	(0.038)
5 (least deprived)	82.6	(0.058)	85.4	(0.055)	84.0	(0.040)
Mean	**79.4**		**83.0**		**81.2**	
Absolute IMD gap	*7.4*	*(0.084)*	*5.5*	*(0.082)*	*6.5*	*(0.060)*
Relative IMD gap	*0.10*	*(0.001)*	*0.07*	*(0.001)*	*0.08*	*(0.001)*
Absolute gender gap	*3.60*					
Relative gender gap	*0.05*					

Note:

1. Standard errors are given in parentheses
2. Absolute gap is Q5-Q1. Relative gap is (Q5/Q1)-1

Table 2.7: Comparisons of absolute and relative inequality in quality-adjusted life expectancy by Index of Multiple Deprivation (IMD) quintile group and gender

Quality-adjusted life expectancy						
IMD quintile group	**Male**		**Female**		**Combined**	
1 (least deprived)	62.3	(0.348)	64.1	(0.375)	63.2	(0.343)
2	67.0	(0.327)	68.2	(0.330)	67.7	(0.306)
3	69.5	(0.309)	70.4	(0.317)	70.0	(0.289)
4	72.8	(0.264)	73.4	(0.267)	73.2	(0.236)
5 (most deprived)	74.8	(0.183)	75.2	(0.181)	75.1	(0.134)
Mean	**69.5**		**70.3**		**70.3**	
Absolute IMD gap	*12.5*	*(0.323)*	*11.2*	*(0.339)*	*11.9*	*(0.328)*
Relative IMD gap	*0.20*	*(0.006)*	*0.17*	*(0.006)*	*0.19*	*(0.006)*
Absolute gender gap	*0.74*					
Relative gender gap	*0.01*					

Note:

1. Standard errors are given in parentheses.
2. Absolute gap is Q5-Q1. Relative gap is (Q5/Q1)-1

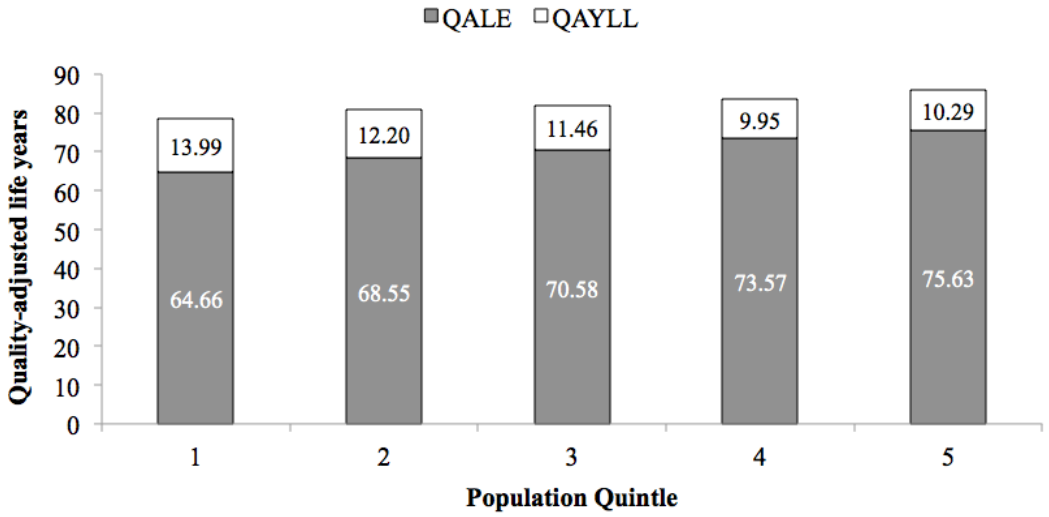
Table 2.8: Predicted quality-adjusted life expectancy by 5-year age group, gender and Index of Multiple Deprivation quintile group

	IMD Quintile Group (Males)						*IMD Quintile Group (Females)*					
Age	**1**	**2**	**3**	**4**	**5**	**Mean**	**1**	**2**	**3**	**4**	**5**	**Mean**
0-4	62.3	67.0	69.5	72.8	74.8	**69.3**	64.1	68.2	70.4	73.4	75.2	**70.5**
5-9	58.3	62.7	65.1	68.2	70.2	**64.9**	60.1	64.0	66.1	68.9	70.7	**66.2**
10-14	53.8	58.1	60.4	63.4	65.3	**60.3**	55.7	59.5	61.5	64.2	65.9	**61.6**
15-19	49.4	53.5	55.7	58.6	60.4	**55.6**	51.4	55.0	56.9	59.5	61.1	**57.0**
20-24	45.0	49.0	51.1	53.9	55.6	**50.9**	47.1	50.5	52.3	54.8	56.4	**52.5**
25-29	40.7	44.5	46.6	49.2	50.8	**46.4**	42.8	46.1	47.9	50.2	51.7	**48.0**
30-34	36.5	40.1	42.1	44.6	46.1	**42.0**	38.6	41.7	43.4	45.6	47.1	**43.6**
35-39	32.4	35.8	37.7	40.0	41.5	**37.6**	34.5	37.4	39.1	41.1	42.5	**39.3**
40-44	28.5	31.6	33.4	35.6	37.0	**33.4**	30.6	33.3	34.8	36.7	38.1	**35.1**
45-49	24.7	27.6	29.3	31.2	32.6	**29.3**	26.8	29.2	30.7	32.5	33.7	**31.0**
50-54	21.2	23.7	25.2	27.0	28.3	**25.3**	23.1	25.4	26.7	28.3	29.5	**27.0**
55-59	17.8	20.0	21.4	23.0	24.1	**21.5**	19.6	21.6	22.9	24.3	25.4	**23.1**
60-64	14.6	16.5	17.7	19.2	20.2	**17.9**	16.4	18.1	19.2	20.5	21.5	**19.5**
65-69	11.8	13.4	14.4	15.6	16.5	**14.5**	13.3	14.8	15.8	16.8	17.7	**16.0**
70-74	9.3	10.5	11.3	12.2	13.0	**11.5**	10.6	11.7	12.5	13.4	14.2	**12.8**
75-79	7.1	8.0	8.6	9.3	9.9	**8.8**	8.2	9.0	9.6	10.2	10.9	**9.9**
80-84	5.3	5.9	6.2	6.8	7.3	**6.5**	6.1	6.6	7.0	7.4	8.0	**7.3**
85+	3.9	4.3	4.4	4.8	5.2	**4.7**	4.4	4.7	4.9	5.2	5.6	**5.3**

Note: The predictions apply to the first year of each group.

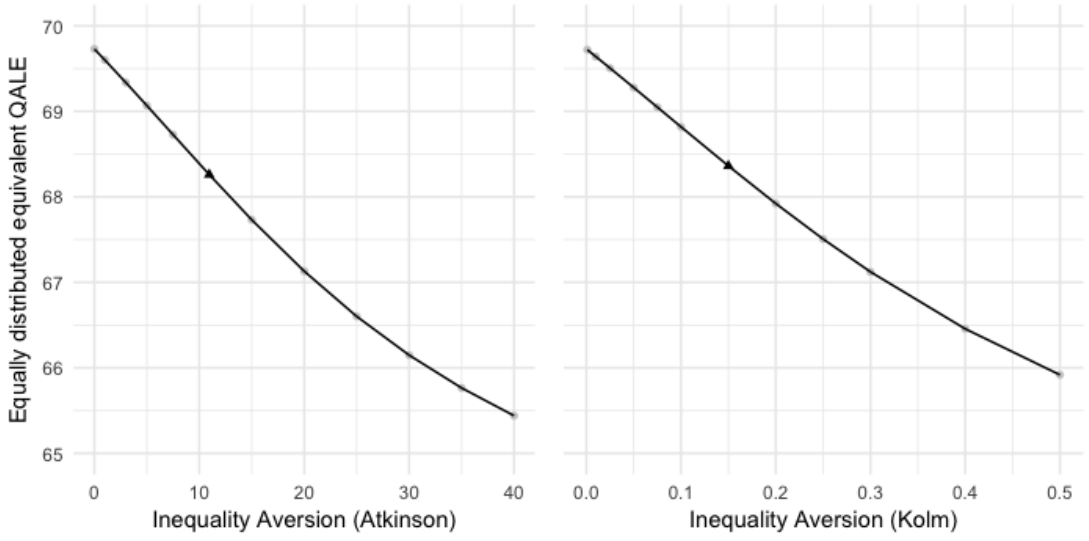
The population distribution of life expectancy and QALE at birth, which reflects both types of inequality and the relative group sizes, is shown in Figure 2.2 and Table A2.11. The quality-adjusted years of life lost from life expectancy due to morbidity increases from 10.29 for the most healthy to 13.99 for the least healthy. This means that the expected lifetime quality-adjusted health of the top quintile is equivalent to 88% of life expectancy in full health, compared to 82% for the bottom quintile. The gap between the top and bottom quintiles is 11 QALYs. The population distribution at 25, 40 and 60 (Table A2.12) shows that relative inequality between the least and most healthy increases with age, from 0.17 at birth to 0.32 at 65.

Figure 2.2: Social distribution of life expectancy and quality-adjusted life expectancy (QALE) in England



Note: Individuals are ranked from least to most healthy based on socioeconomic-gender sub-group, and divided into quintile groups. The total height of each bar is the LE estimate, which can be divided into QALE and the quality-adjusted years of life lost (QAYLL) due to morbidity over an individual's lifetime.

Figure 2.3: Mean equally distributed equivalent (EDE) health by degree of inequality aversion. EDE at the inequality aversion level elicited from the general population in Robson et al. (2016) is signified by a black triangle



Note: QALE = quality-adjusted life expectancy

The results from the social welfare analysis are shown in Figure 2.3. The mean QALE in the population (in which there is no inequality aversion) is 69.7 QALYs. At the general population values of inequality aversion for the Atkinson and Kolm indices, EDE QALE is 68.3 and 68.4 QALYs, respectively. As inequality aversion increases, a social decision-maker is willing to trade-off more average health to eradicate inequality. When using the Atkinson index, EDE decreases from 69.1 QALYs for ε=5 to 65.4 QALYs for ε=40.

The cumulative distribution of QAD is shown in Figure 2.4. The curve for males lies above that for females up to QADs of approximately 77 QALYs. The median QAD, for instance, is 68 for males and 72.8 for females. Males have the higher maximum quality-adjusted lifespan at 82.7, however. Approximately 9% experience more lifetime QALYs than the female maximum of 80.8.

Figure 2.4**:** Cumulative distribution of quality-adjusted age at death in England

2.4.4 Sensitivity analysis

Regression output when using the alternative estimators is reported in Table 2.5. In terms of predicting the mean subgroup EQ-5D scores, all estimators perform similarly. These differences are shown in Figure A2.6 in the appendix, where it is

seen that the observed scores for those in IMD 1 are more unstable than those in IMD 5, for which prediction is more accurate. In terms of QALE, OLS, tobit and 2PM again all yield comparable results, with none substantially outperforming the others. Mean absolute deviations in QALE prediction from those using observed scores, shown in Table 2.5, were all comfortably under 0.5 QALYs; the imputed scores fared best with an MAE of 0.327 QALYs, whilst the 2PM had the highest at 0.434 QALYs.

QALE predictions do not markedly change when NS-SEC replaces IMD quintile groups as the socioeconomic variable in the HRQL regression. All of the coefficients in the regression analysis were statistically significant ($P<0.01$). The average difference in predicted QALE across population quintile groups between mapped and non-mapped estimates is 0.7 QALYs. One statistic that was noticeably different was absolute inequality between the least and most healthy population quintile groups, which increased from 11.0 to 14.0 QALYs.

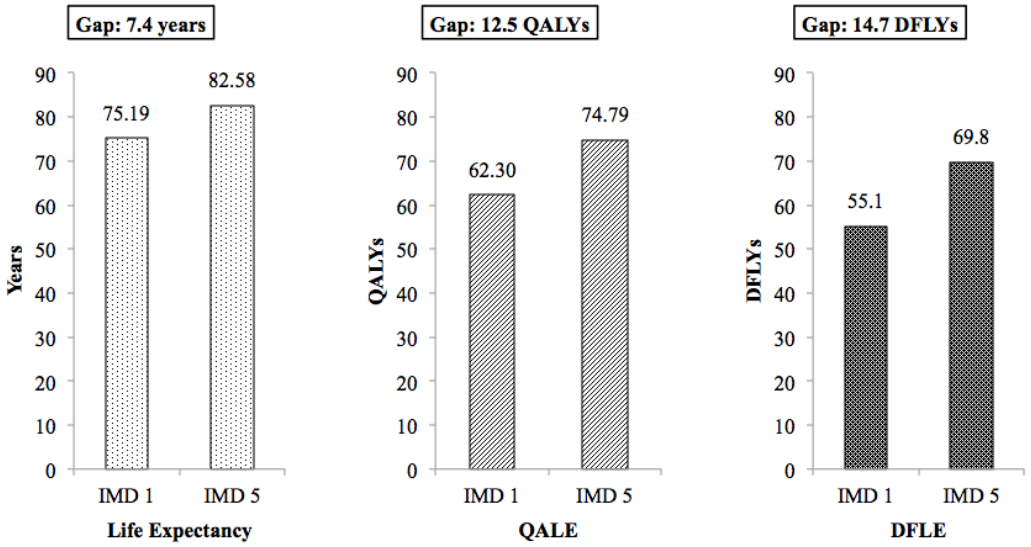
2.5 Discussion

2.5.1 Principal findings

Adjustment for morbidity using detailed patient reported data on HRQL substantially increased the size of socioeconomic health inequality when compared with life expectancy alone. Figure 2.5, using additional numbers from the ONS (2013), compares estimates of life expectancy, QALE and DFLE at birth for males in the least and most deprived quintile groups. A male in quintile group 1 is expected to experience 83 years, 75 QALYs or 70 disability-free life years. The discrepancies between these figures clearly demonstrate the impact of using the QALY, rather than a binary disability indicator, to measure morbidity. The increased sensitivity of the former to states of illness and disability creates a more realistic picture of health experience and a more accurate measure of health inequalities. Consequently, the inequality between those in the least and most deprived quintile groups for QALE of 12.5 QALYs sits between those predicted for life expectancy and DFLE, at 7.4 and 14.7, respectively. These estimates of socioeconomic inequality are consistent with

the figure of 12.7 QALYs that is estimated in an analysis of QALE disparities in one region of England (Collins 2013).

Figure 2.5: Inequalities in life expectancy, quality-adjusted life expectancy (QALE) and disability-free life expectancy (DFLE) for males at birth in the most and least deprived quintile groups



Note:

3. QALY = quality-adjusted life year; DFLY = disability-free life years
4. DFLE estimates are estimates for 2007-10 and taken from the ONS (2013). Life expectancy and QALE estimates are for 2010-12

The univariate distribution of health estimated in Figure 2.2, the principal use of which is in distributional analyses of new health interventions (see Section 2.5.3), estimates inequality between the most and least healthy at 11 QALYs. Although no direct comparisons could be found for this estimate, it is consistent with estimates from Marmot et al. (2010), who calculate DFLE for all small areas in England and find disparities of 13 DFLYs between the most and least healthy neighbourhoods.

Our social welfare analysis helps to quantify the social value lost through inequality. At the general population values of inequality aversion, we would be willing to sacrifice approximately 1.4 QALYs from the average lifespan in order to obtain a perfectly equal distribution. The distribution of quality-adjusted age at death also provides a different perspective on inequality. The influence of premature death on

QALE is clearly shown: nearly 25% of men experience in excess of 75 QALYs in their lifetime, despite the highest mean estimate of QALE being 74.8. We also see the lower mortality of women driving the differences in lifetime health and a small proportion of men who survive to old age experiencing greater life time QALYs due to lower average morbidity.

Causal inference is not required or attained in the regression framework used in this study, since we only aim to describe how expected lifetime health varies by age, sex, and socioeconomic status. Nevertheless, the selected covariates do exhibit strong associations with utility scores, justifying the measurement of inequality with respect to them.

Results are robust to a wide range of differing assumptions and estimation methods. The OLS predictions did not differ drastically from the 2PM or the tobit, performing similarly on diagnostic measures. Interestingly, despite the 2PM performing better and displaying lower MAE than both OLS and tobit for EQ-5D prediction, we find a greater MAE when estimating QALE. This is explained by the fact that the 2PM predicts the EQ-5D with larger error for the youngest age group, who have the largest impact on QALE estimates.

The anticipated difficulties in mapping IMD quintile groups to NS-SEC category are also surprisingly small. Despite the seemingly crude method of the mapping of mean IMD-specific EQ-5D scores (Table A2.10), the resulting QALE estimates are not substantially different than when mean EQ-5D scores are estimated for IMD quintile groups (Table 2.5). Multiple imputation of the missing data had a marginal influence, with a mean difference of 0.3 QALYs across the 360 predictions when using the imputed and non-imputed scores. Thus, despite tests indicating that the MCAR assumption was violated, assuming that missing values were Missing At Random (MAR) had little impact on the final results.

2.5.2 Strengths and limitations

This study utilised mortality data for the entire population of England and HRQL data for a large and representative sample of over 25,320 individuals. The validity of

the results is further substantiated by the robustness of the findings to alternative socioeconomic variables, regression estimators and ways of handling missing data.

The limitations of the analysis are largely over the estimates of HRQL. First, mean EQ-5D scores for the groups of interest are likely to be overestimates since the HSE is only representative of the non-institutionalised population. Those who reside in institutions such as nursing homes or prisons are on average likely to be unhealthier than the HSE sample. Since they are also likely to be in lower deprivation groups, this implies that our estimates of inequality are likely to be conservative underestimates.

Second, the subjective nature of EQ-5D reporting must be acknowledged. As noted elsewhere in the literature (Minet Kinge & Morris 2010, p.1869), there may be a systematic reporting bias associated with an individual's socioeconomic status. For example, those in higher socioeconomic groups who are in more sedentary work may have their usual activities inhibited less by illness or injury, resulting in a higher utility score. Conversely, people in low socioeconomic groups may have relatively low health expectations and so self-report feeling in relatively good health in particular dimensions, resulting in a higher utility score than people in high socioeconomic groups whilst experiencing the same level of health from an external clinical perspective. This again will have the effect of under-estimating the true degree of health inequality related to socioeconomic status.

Third, the assumption that those under 16 have HRQL equal to those in the 16-19 age group is unavoidable but improbable. This is verified by the fact that in the HSE sample, 95% of those under 16 reported 'Very Good' or 'Good' health, compared with only 90% in the 16-19 group, suggesting that our assumption may lead us to underestimate HRQL for the former, thereby again providing a conservative under-estimate of inequality.

2.5.3 Implications and conclusions

This study is the first to estimate the social distribution of QALE in England. Compared with previous estimates based on simple binary measures of morbidity,

our estimates, based on more detailed data on health-related quality of life, provide more accurate and credible estimates to inform policy-makers about the overall extent of health inequalities. The Marmot Review (Marmot et al. 2010) and the Department of Health's Public Health Outcomes Framework (DH 2013), which respectively use DFLE and HLE as indicators, are two examples of prominent public initiatives using crude binary indicators of morbidity.

Our results are also of interest in health technology assessment. First, the individual predictions of EQ-5D and QALE in Table A2.10 and Table 2.8 can be used as reference values within decision models. Second, as discussed in Chapter 1, the use of the QALE population distribution is an important empirical step in monitoring the health inequality impacts of health interventions in distributional cost-effectiveness analysis (Cookson et al. 2017). These types of evaluations would model the health benefits by IMD and gender, and use a distribution of health opportunity costs (estimated in Chapter 3) over the same groups to acquire an estimate of the net impact. These impacts can be modelled on the estimates presented here and can be examined, using a range of inequality measures, to see how health inequality has changed (Asaria, Griffin & Cookson 2015; Asaria et al. 2015). This process is exemplified in Chapters 4 and 5.

Future work could develop our model to introduce additional variables, such as race, into the analysis, paving the way for a more nuanced distribution that reflects stakeholder judgements on fair and unfair determinants of inequality. The techniques we describe can also be applied to other countries and settings where health inequalities are of concern. The accuracy of the inequality estimates could also be improved by incorporating more granular data on EQ-5D and mortality, for example by small area rather than IMD quintile. This would naturally place greater demands on subgroup evidence when evaluating new interventions. At present, however, this study brings us a step closer to explicit analysis of the equity-efficiency trade-offs involved in health care resource allocation.

Appendix 2

Table A2.9: Mapping process outlining how predicted utility scores for NS-SEC categories were combined with mortality rates for IMD quintile groups as part of a sensitivity analysis. The objective was to match group sizes as closely as possible

Social Class (NS-SEC)	Deprivation (IMD quintile group)
I (Higher managerial / Professional)	5 (Least deprived)
II (Lower managerial / Professional)	5
III (Intermediate occupations)	4
IV (Small employers / Own account holders)	3
V (Lower supervisory / Technical occupations)	3
VI (Semi-routine occupations)	2
VII (Routine occupations)	1
VIII (Never worked and long-term unemployed)	1 (Most deprived)

Table A2.10: Predicted EQ-5D scores by age, gender and IMD quintile group

	IMD Quintile Group (Males)						*IMD Quintile Group (Females)*					
Age	**1**	**2**	**3**	**4**	**5**	**Mean**	**1**	**2**	**3**	**4**	**5**	**Mean**
15-19	0.89	0.93	0.94	0.97	0.98	0.94	0.87	0.91	0.92	0.95	0.96	0.92
20-24	0.88	0.92	0.93	0.96	0.97	0.92	0.86	0.90	0.91	0.94	0.95	0.90
25-29	0.87	0.90	0.92	0.95	0.96	0.91	0.85	0.88	0.90	0.92	0.93	0.89
25-29	0.86	0.89	0.90	0.93	0.94	0.90	0.84	0.87	0.88	0.91	0.92	0.88
35-39	0.84	0.88	0.89	0.92	0.93	0.89	0.82	0.86	0.87	0.90	0.91	0.87
40-44	0.83	0.86	0.88	0.90	0.91	0.88	0.81	0.84	0.86	0.88	0.89	0.86
45-49	0.81	0.85	0.86	0.89	0.90	0.86	0.79	0.83	0.84	0.87	0.88	0.84
50-54	0.80	0.83	0.85	0.87	0.89	0.85	0.78	0.81	0.83	0.85	0.86	0.83
55-59	0.78	0.82	0.83	0.86	0.87	0.84	0.76	0.80	0.81	0.84	0.85	0.82
60-64	0.77	0.80	0.82	0.84	0.85	0.82	0.75	0.78	0.79	0.82	0.83	0.80
65-69	0.75	0.79	0.80	0.83	0.84	0.81	0.73	0.77	0.78	0.81	0.82	0.79
70-74	0.74	0.77	0.78	0.81	0.82	0.79	0.72	0.75	0.76	0.79	0.80	0.77
75-79	0.72	0.75	0.77	0.79	0.80	0.77	0.70	0.73	0.75	0.77	0.78	0.75
80-84	0.70	0.74	0.75	0.78	0.79	0.76	0.68	0.72	0.73	0.76	0.77	0.73
85+	0.68	0.72	0.73	0.76	0.77	0.74	0.66	0.69	0.70	0.73	0.74	0.71

Note:

1. Since EQ-5D is not estimable for individuals aged 0-14, their predicted scores are assumed to be equal to those aged 15-19.

Table A2.11: Life expectancy and quality-adjusted life expectancy estimates by population quintile group. The univariate distribution from which these are estimated ranks the entire population from least to most healthy, based on socioeconomic-gender sub-groups. Standard errors are in parentheses

Population quintile	Life Expectancy		Quality-adjusted Life Expectancy		Ratio of QALE to LE
1 (least healthy)	78.6	(0.084)	64.66	(0.060)	0.82
2	80.7	(0.078)	68.55	(0.059)	0.85
3	82.0	(0.072)	70.58	(0.055)	0.86
4	83.5	(0.074)	73.57	(0.058)	0.88
5 (most healthy)	85.9	(0.074)	75.63	(0.058)	0.88
Absolute gap (5-1)	7.3		11.0		
Relative gap (5/1)-1	0.093		0.170		

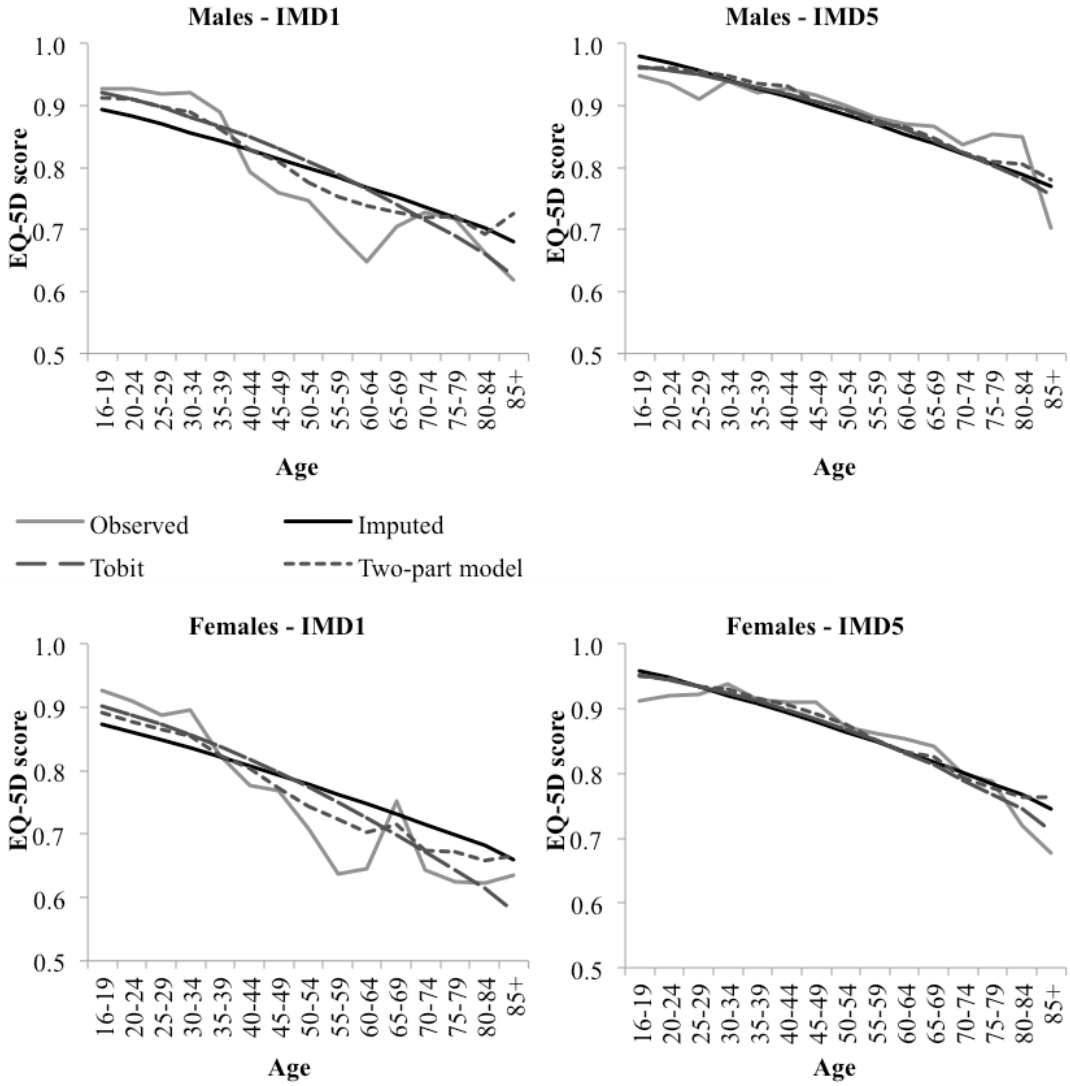
Note: QALE = quality-adjusted life expectancy; LE = life expectancy

Table A2.12: The social distribution of health as the baseline year is increased

	Population quintile group						
	1	**2**	**3**	**4**	**5**	**Absolute Gap**	**Relative Gap**
At birth	64.7	68.5	70.6	73.6	75.6	11.0	0.170
At 25	43.4	46.3	48.0	50.3	52.1	8.7	0.200
At 40	31.1	33.5	34.9	36.6	38.4	7.3	0.236
At 65	13.7	14.6	15.8	16.7	18.0	4.4	0.320

Note: Absolute gap is Q5-Q1. Relative gap is (Q5/Q1)-1.

Figure A2.6: Observed and predicted EQ-5D scores over age groups. Plots for the highest and lowest index of multiple deprivation (IMD) quintile groups are shown for each gender



Note: IMD 1=most deprived, IMD 5 = least deprived

Chapter 3: Estimating the social distribution of health impacts from changes in English NHS expenditure

3.1 Introduction

Two central objectives of public health care systems are to help improve population health and to help reduce health inequality. In England, for example, the leaders of the universal, publicly funded National Health Service (NHS) have a legal obligation to consider reducing inequalities between the people of England in the health benefits they obtain from health care services (NHS Health and Social Care Act 2012). Whilst many studies have examined the impact of public health care expenditure on population health (Cochrane et al. 1978; Crémieux et al. 1999; Nixon & Ulmann 2006; Martin et al. 2008; Marton et al. 2015; Singh 2014), much less is known about the impact on health inequality. If policy makers wish to reduce health inequality, they need to know whether investment in health care will deliver a larger health inequality reduction than other social programmes, such as education and social protection, and which kinds of health care expenditure deliver the largest health inequality reductions. In this chapter we provide an empirical estimate of the distribution of health impacts resulting from changes in NHS expenditure by socioeconomic status, age, gender and disease area.

To date only one study has examined the health inequality impacts of health care expenditure in the UK. Barr et al. (2014) used longitudinal data for local authorities in England to analyse associations between changes in NHS spending and changes in mortality considered amenable to influence by health care. They found greater reductions in amenable mortality in more deprived local authorities, which they attribute to i) larger increases in health expenditure per head in higher-deprivation areas and ii) higher productivity with respect to mortality reduction for every £1 spent in the higher deprivation areas.

In contrast, our study explores how health impacts from expenditure changes are distributed by age, gender and disease area as well as socioeconomic status, and utilises existing evidence that carefully attempts to identify causation using an

instrumental variables approach. We develop the results of a recent study (Claxton, Martin, et al. 2015) that explores the relationship between health care spending and mortality in 152 different sub-national administrative areas of the English NHS, adjusting for quality of life differentials. It estimates the cost of producing a quality-adjusted life year (QALY) through additional health care expenditure, accounting for differences between broad clinical areas of expenditure (e.g. respiratory disease, circulatory disease and so on), controlling for potential endogeneity bias in the estimation of the mortality effects of expenditure using an instrumental variables approach. This evidence provides a benchmark cost-per-QALY value for appraising new NHS expenditure programmes, because it can be interpreted as the health opportunity cost to the English NHS of displacing alternative health care services that could have been funded instead.

However, this benchmark is not sufficient where policy makers are also concerned about the health inequality impacts of public expenditure. By extending the results of Claxton and colleagues, we provide the first evidence on how the health effects of changes in NHS expenditure are distributed between social groups in terms of both quality and length of life. Our findings can therefore show how health opportunity costs at the margin are distributed socially when introducing a new, cost-increasing programme. When combined with an estimate of who gains most from the new health programme, for example in distributional cost-effectiveness analysis (Asaria, Griffin, Cookson, et al. 2015), this allows the net health inequality impact to be evaluated, as we demonstrate in chapters 4 and 5 of this thesis.

3.2 Background

3.2.1 Overview

We estimate the socioeconomic distribution of health effects arising from marginal changes in health care expenditure in the English NHS. We use underpinning evidence from previous work by Claxton and colleagues on the causal link between health outcomes and health care expenditure by disease area. By adding information about the socioeconomic distribution of health care utilisation, we provide (i) the link between socioeconomic characteristics and health outcomes; and (ii) measures for

summarising the impact on socioeconomic inequality in the distribution of health. Since this evidence is the foundation on which we build in the present study, it is important to understand the nature and limitations of that evidence. Sections 3.2.2 to 3.2.4 describe the complex methods and data used in this previous study, with section 3.3 describing the approach, methods and data undertaken in the present analysis.

3.2.2 Effect of expenditure on mortality

Since 2003, each regional spending body of the English NHS (formerly Primary Care Trusts, now Clinical Commissioning Groups) has been required to categorise all expenditure into one of 23 programme budgeting categories (PBC). Each PBC covers a broad clinical area such as cancer or infectious disease, and is defined by a subset of International Classification of Disease (ICD) Version 10 codes. Martin et al. showed how observations on expenditure could be linked to mortality using routine Primary Care Trust level data within each PBC (Martin et al. 2008; Martin et al. 2012). Claxton and colleagues built on this analysis using more recent data that included all 152 Primary Care Trusts and covered all programmes of care.

Analysis of cross-sectional data can potentially suffer from endogeneity bias; for example, reverse causality if poor health outcomes motivate decision-makers to increase health expenditure. These and other problems (Gravelle & Backhouse 1987) may account for the substantial variation in published estimates of the magnitude of the health effect of additional health care expenditure (Crémieux et al. 1999; Or 2001; Nixon & Ulmann 2006; Gallet & Doucouliagos 2015; Vallejo-Torres et al. 2016). Claxton and colleagues therefore use a two-stage least squares instrumental variables approach to account for endogeneity. Two equations are estimated for each PBC: an expenditure equation linking the NHS budget to PBC expenditure and an outcome equation linking PBC mortality to PBC expenditure in a particular year. The equations control for need by including the Department of Health's formula for regional need (Morris et al. 2007) and/or programme-specific variables (such as diabetes prevalence rates). Three-year averages of mortality are used to account for temporal fluctuations, the first year of which aligns with that of the expenditure data, thereby allowing for a lagged effect of the latter on the former to be captured. Since

mortality data are only available for eleven PBCs, the productivity of the remainder are assumed to be equal to the average of those where health effects could be estimated (with the exception of PBC 23, 'other', which is assumed to have zero health gain). The expenditure and outcome equations, respectively, are as follows:

$$x_i = \alpha + \beta n_i + \gamma m_i + \theta y_i + \varepsilon_i \text{ for } i = 1, \dots, 152$$

$$h_i = \rho + \delta n_i + \pi x_i + \varepsilon_i \text{ for } i = 1, \dots, 152$$

where x_i is expenditure; n_i is the own programme need for care; m_i is the need for care in other programmes; y_i is the total budget and h_i is the health gain in PCT i. The variables are log-transformed such that the coefficients of interest, θ and π, represent elasticities: θ is the percentage change of a PBC budget with respect to a percentage change in the overall NHS budget; π is the percentage change in health for a percentage change in a PBC budget.

To account for the endogeneity in both equations, a large number of instruments are acquired from census data and tailored to each PBC. A battery of tests, including the Hansen–Sargen test, the Kleibergen–Paap Lagrange multiplier test and the Kleibergen–Paap F-statistic, are performed; where instruments are weak or invalid, a combination of other census-derived variables are used instead. A full list of all the instruments considered by the authors is given in Table 92 of their report (Claxton, Martin, et al. 2015, p.347). This strategy ensured that, even though endogeneity is indeed found to be present in many of the expenditure and outcome equations, the instrument set used for each is valid and sufficiently strong, thereby giving consistently estimated coefficients.[5]

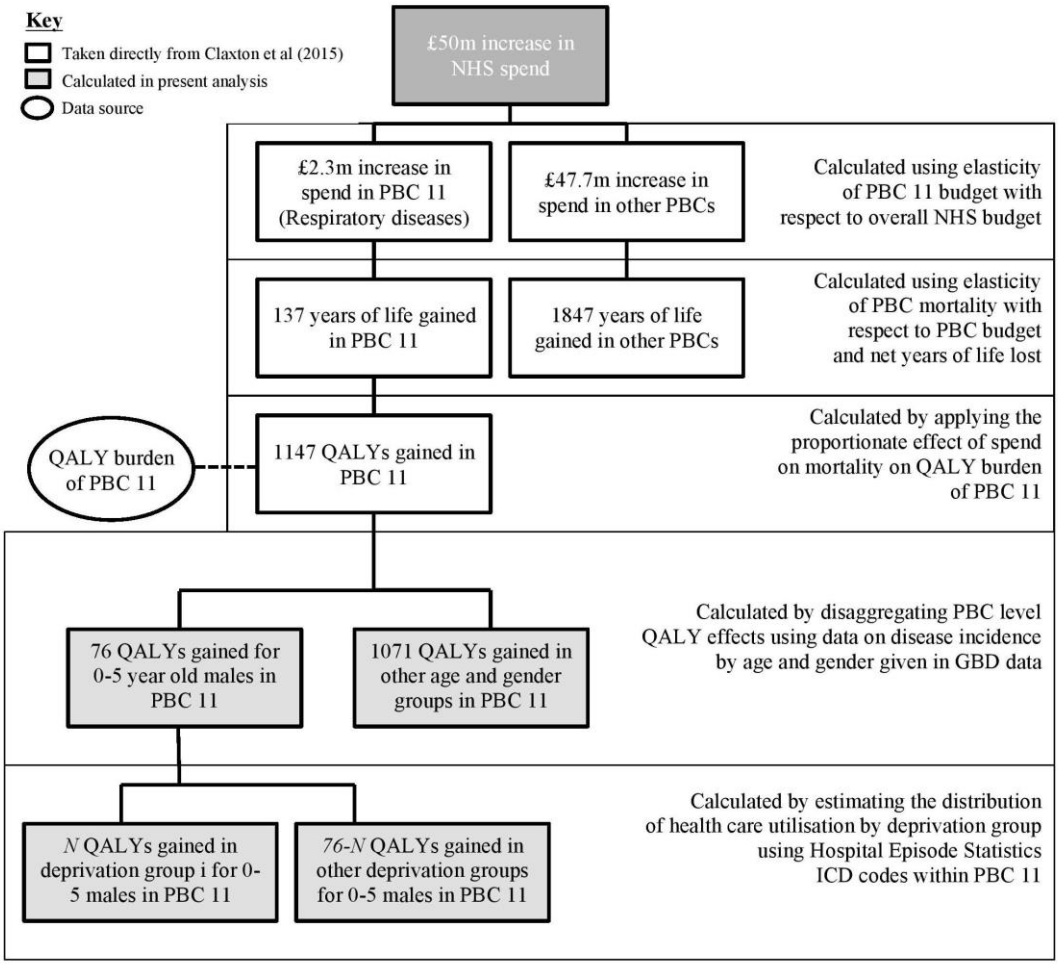
[5] The instruments included for each expenditure and outcome equation and their performance in the tests are detailed in the section *"Analysis of programme budgeting expenditure for 2008/9 and mortality data for 2008/9/10"* in Appendix 2 of Claxton et al's report (Claxton, Martin, et al. 2015, p.314).

3.2.3 Extending outcome to QALYs

To estimate the effect of spending on QALYs, mortality effects are first converted into 'net' years of life lost by disease area using data on age of death by PBC from the Office for National Statistics (ONS). This accounts for counterfactual deaths that would have occurred in the at-risk populations; Chapter 4 of their report provides full details of the calculations (Claxton, Martin, et al. 2015, p.45). When the elasticities from the mortality equations are applied to net years of life lost, they provide an estimate of life years gained from additional expenditure, as shown in the third tier of Figure 3.1.

Claxton and colleagues then translate these into QALYs by adding the health lost due to reduced quality of life whilst living with a disease to the health lost due to premature death, using data at the ICD code level. The process is described in Chapter 4 of their report (Claxton, Martin, et al. 2015, p.56) and involves weighting years of life lived with, and lost to, each disease by combining evidence on net years of life lost, incidence, duration of disease, age and gender with quality of life scores by disease. The QALY burdens for each ICD code within a PBC are then summed to generate a PBC level QALY burden. The change in QALYs for a change in expenditure is then yielded by applying the proportionate effect of spend on mortality (i.e. the spend elasticity multiplied by the outcome elasticity) to the QALY burden for a given PBC, as shown in the fourth tier of Figure 3.1. This implies that the ratio of PBC-level health effects of life extension to quality of life gains is identical to that in the respective QALY burden. We hereafter refer to the notional marginal QALY from a change in expenditure as an additional QALY brought about by an expenditure increase, although it is equally legitimate to conceive of it as a forgone QALY brought about by a budgetary reduction.

Figure 3.1: How a £50m National Health Service (NHS) budget increase is translated into quality-adjusted life year (QALY) effects for the respiratory disease programme budgeting category (PBC)



Notes:

1. ICD = International Classification of Disease; GBD = Global Burden of Disease.
2. Proportionate effect of spend on mortality for PBC 11 is calculated by multiplying the respective elasticities from the outcome and spend equations.
3. Net years of life lost for each PBC are calculated using a method described in Chapter 4: Translating mortality effects into life-years and quality-adjusted life-years in Claxton et al. (2015)

3.2.4 Disaggregating health effects by age, gender and disease

The main work by Claxton and colleagues only requires health effects at the PBC level, but a subsequent publication (Claxton, Sculpher, et al. 2015) provides more detailed ICD level breakdowns. These are published online in "Appendix A

(Displacement by ICD Code)"[6]. These data can be re-aggregated back to PBC level to get the proportion of an additional QALY attributable to that PBC (in line with the original analysis), denoted p_k. Since $\sum_k p_k = 1$, these proportions are analogous to the probability that a QALY falls in PBC k at the margin, denoted $P(Q_k)$. This is based on the assumption that QALY gains are distributed within a PBC population according to the age and gender distribution of the incident population, using disease-specific estimates from the World Health Organization's Global Burden of Disease Study (WHO, 2015). This provides the probability an additional QALY falls on an individual of gender g and in age group a and PBC k, denoted $P(Q_{a.g,k})$ and represented in the fifth tier of Figure 3.1. These data are the starting point for our analysis of the social distribution of health effects.

3.3 Methods

3.3.1 Analytical framework

Our aim is to disaggregate each additional QALY by social groups of interest. We focus on three characteristics: (i) socioeconomic status, due to its political importance and strong correlation with health outcomes, which is commonly viewed as inequitable, (ii) age and (iii) gender. We start with data on the probability the additional QALY affects an individual in socioeconomic group d given that it falls on a specific age-gender-PBC group, given by $P(Q_d|Q_{a,g,k})$. We then multiply this by the age-gender-PBC probabilities, $P(Q_{a.g,k})$ to obtain the probability that an additional QALY falls on a particular age, gender, PBC and socioeconomic group:

$$P\left(Q_d\middle|Q_{a,g,k}\right)P(Q_{a,g,k}) = P(Q_{a,g,k,d})$$

We then sum over PBCs to describe the social distribution of an additional QALY from NHS expenditure, represented by the sixth tier in Figure 3.1:

$$\sum_k P(Q_{a,g,k,d}) = P(Q_{a,g,d})$$

[6] Available at http://www.york.ac.uk/che/research/teehta/thresholds/

In the absence of direct observations of the socioeconomic characteristics of those whose health services are actually affected following budget changes, we require a data source that acts as an appropriate proxy.

3.3.2 Data and variables

Hospital Episode Statistics (HES) is a database containing information on all NHS funded activity in public and private hospitals. The primary unit of measurement is the 'consultant episode'; patients whose care is transferred between consultants during a single stay in hospital may have multiple episodes. The HES data include a wide range of personal and geographical variables for each patient, including a unique patient identifier code, age at the beginning of the episode, gender and postcode. We used HES data from 2013/14, the most recent data available at the time the analysis was conducted. Whilst these data do not temporally align with those used in Claxton et al (mortality data for 2008-10 and expenditure 2008/9), they provide a more up-to-date estimate of the socioeconomic distributions. The HES inpatient dataset includes both day cases and overnight stays, encompassing a total of 19,578,568 unique episodes.

The Index of Multiple Deprivation (IMD) is used as our measure of socioeconomic status. The IMD is a weighted index of 38 variables covering seven dimensions of deprivation (employment, income, education, health, crime, living environment, and housing/services) that is given to each of the 32,482 lower layer super output areas (LSOA) in England. Each postcode belongs to an LSOA, giving each patient a deprivation score. The 2004 version of the IMD is provided in HES for the financial year 2013/14 (Noble et al. 2004), which uses LSOA boundaries from the 2001 census. Newer versions of the IMD are available and can be attributed by postcode. Differences in the IMD score will mean that the areas of some individuals will have changed in deprivation level between 2004 (the time the score was attributed) and 2014 (when entered secondary care). However, the proportion of patients transitioning between quintiles is expected to be minimal: the number of areas remaining in the most deprived decile group between 2004 and 2010, for example, was 82.7% (McLennan et al. 2011). As less movement is expected between quintiles

groups due to their larger size, the pre-coded 2004 variable is used in our analysis. This gathers the LSOAs (and their populations) into equally sized quintile groups to obtain a five-level socioeconomic status variable.

ICD codes are included as diagnosis variables, of which up to twenty 4-digit codes can be recorded for each episode. These are converted to 3-digit codes, providing 1,562 diagnostic categories that are mapped to the 23 PBCs. The comprehensive and nationally representative nature of the data justifies the use of HES in our analysis. No other source containing the age, gender and socioeconomic information of patients for all disease areas could be identified.

We anticipate that the proportion of episodes with missing data would be small. It is also thought that the reasons would be largely administrative and due to data entry errors, and therefore we assume that data are missing completely at random and removed observations with missing age, gender IMD quintile values, or at least one diagnosis code from the sample. We further remove observations with gender unspecified. Age is grouped using the eight age bands used in the GBD study, containing a mixture of 5-, 10-, and 15-year bands.[7]

An episode is deemed related to an ICD code if the latter appeared in any of the 20 diagnosis codes. Consequently, episodes with multiple diagnosis codes will be 'counted' multiple times. We then construct a matrix containing the episode counts for each age, gender, ICD and IMD group.

3.3.3 Calculating QALY probabilities

Although eight age groups, both genders and the 23 PBCs would generate 368 subgroups, we analyse 320 since three PBCs are not allocated any health effects by Claxton and colleagues for reasons detailed in their report (Claxton, Martin, et al. 2015, p.103): Trauma and Injuries (PBC 16), Social Care (PBC 22) and General Medical Services (PBC 23). For each group we count the number of episodes in each

[7] These age groups are 0-4, 5-14, 15-29, 30-44, 45-59, 60-69, 70-79, 80 and over.

of the five IMD quintile groups. The count matrices are produced using Stata 12, with all subsequent analyses performed in R.

Figure 3.2: Calculation of the proportion of the overall health effect from a change in overall health care expenditure attributed to 30-44 year old females in the highest deprivation group

Step 1: Effect of health care expenditure change on PBC expenditure

- Calculate elasticity of respiratory disease (PBC 11) expenditure (£3,994m) with respect to percentage change in total health care expenditure (£78,398m)
  - Instrumental variable regression model estimated across Primary Care Trusts in England
  - 1% change in total expenditure (£784m) yields 0.9% change in PBC 11 expenditure (£36m)

Step 2: Effect of PBC expenditure on health

- Calculate elasticity of years of life lost in PBC 11 (73,922 life years) with respect to percentage change in PBC 11 budget
  - Instrumental variable regression model estimated across Primary Care Trusts in England
  - 1% change in PBC 11 expenditure yields 1.808% change in years of life lost in PBC 11
- Assume that the propotionate effect of expenditure on quality-adjusted life years is the same as for unadjusted life years
- The effect of a 1% change in overall expenditure on health in PBC 11 is therefore 0.9% x 1.808% x 100 = 1.627%
  - This is applied to the annual QALY burden for PBC 11 of 1,105,027 QALYs, yielding a change in QALYs of 17,981

Step 3: Effect of health care expenditure change on health

- Steps 1 and 2 are repeated for all PBCs to obtain the total health effects of a change of a 1% change in overall expenditure
  - 12 PBCs did not have appropriate mortality data to perform the second regression analysis in step 2.
  - For these PBCs, a weighted average effect is calculated from the 11 PBCs where analyses were conducted and applied to PBC-specific QALY burden
- Health opportunity costs calculated as the change in overall expenditure / change in overall QALYs
  - Estimated cost of a QALY at the margin of £12,937
- Respiratory illness accounts for 29.67% of the overall health change in QALYs

Step 4: Disaggregating health effects to social groups

- Global burden of disease data are used to calculate the size of the patient population for PBC 11, by mapping from u-codes to ICD code to PBC
  - Patient population for PBC 11 is 8,588,210. 527,447 (6.1%) of these are females aged 30-44
  - The proportion of the total health effect accruing to 30-44 year old females in PBC 11 is therefore 29.67% x 6.1% = 1.81%
- Hospital episode statistics for 2013 are used to calculate the socioeconomic distribution within each age and gender subgroup in PBC 11
  - 342,759 hospital episodes relating to PBC 11 are for 30-44 year old females. 90,864 (26.5%) of these are for individuals in the most deprived quintile group.
  - The proportion of the total health effect accruing to 30-44 year old females in the most deprived quintile group in PBC 11 is therefore 1.81% x 26.5% = 0.48%

Note: PBC = programme budgeting category; QALY = quality-adjusted life year; ICD = International Classification of Disease.

The counts are converted into proportions for each row to obtain the socioeconomic distribution of each age, gender and PBC group. These proportions are considered proxies for the probability that an additional QALY falls on deprivation group d given that it falls on a particular age, gender and PBC group, $P(Q_d|Q_{a,g,k})$. Multiplying these by the probability of a QALY change falling on an age-gender-PBC group, $P(Q_{a,g,k})$ yields the probabilities that the new NHS QALY will fall on a particular age-gender-PBC-socioeconomic group, or $P(Q_{a,g,k,d})$. Only one group, 0-5 males in PBC 18, had no episodes associated with it. We assumed a flat socioeconomic distribution for this group, which accounts for less than 0.001 of the total QALY probability.

3.3.4 Analysing social inequalities

To analyse the social distribution of an additional QALY, we firstly sum over PBCs to obtain the distribution of probabilities of receiving the marginal QALY by age, gender and socioeconomic status. From here we investigate socioeconomic inequality by age, gender and socioeconomic status: (i) summing over age groups and genders yields an aggregate socioeconomic distribution; (ii) summing over genders yields the socioeconomic distribution of probabilities by age group; (iii) summing over age groups yields the socioeconomic distribution of probabilities by gender.

In order to present the results and distributions by 5-year age group we split the probabilities of 10- and 15-year age groups from the GBD study into 5-year bands, using the respective population proportions of the 5-year bands from which they are composed, obtained from the ONS (2014). For example, the probability associated with 70-79 year-old men was disaggregated into the 70-74 and 75-79 bands according to their proportions within the 70-79 band, which are 0.56 and 0.44, respectively. These are general population estimates and are thus not specific to each disease area.

3.3.5 Inequality measures

To compare socioeconomic distributions, we compute two summary measures of inequality: the slope index of inequality (SII) and the relative index of inequality (RII) for IMD quintile groups.

The SII measures absolute inequality using Ordinary Least Squares regression to estimate the effect of IMD on the probability of gaining the marginal QALY. Rather than using an ordinal socioeconomic variable in the regression equations, we convert each IMD level into a ridit score, a continuous variable on a 0-1 scale. Although IMD quintile groups are equally sized in the whole population, we also analyse inequality by age group, where the distribution over IMD groups differs. Ridit scores resolve this issue by accounting for the relative size of each socioeconomic group: the ridit score of each group represents the mid-point of the cumulative distribution of a subgroup population by IMD. For example, as IMD quintile 1 represents 20% of the population, the ridit score would be 0.2/2=0.1 (for further details see Beder and Heim (1990)). SII is then estimated using the following model:

$$q_i = \alpha + \beta_{SII} r_i + \varepsilon_i$$

Where β_{SII} is the SII value, q_i is the probability of the QALY accruing to deprivation quintile i, r_i is deprivation quintile ridit score, ε_i is the idiosyncratic error, α is the constant term. β_{SII} is the slope parameter of r_i on q_i and represents the SII. It is interpreted as the absolute change in probability of moving from the highest to the lowest deprivation groups.

A greater negative SII value indicates a steeper 'pro-poor' gradient and means that low socioeconomic status has a higher absolute probability of receiving the QALY. For example, an SII of -0.02 in PBC 2 (Cancer) would mean that cancer patients in the poorest fifth of neighbourhoods would be 2 percentage points more likely to receive an additional NHS QALY than the richest fifth. RII uses the same information to calculate the relative difference in moving from the most to least deprived quintile group. An RII of -0.5, for instance, would imply that cancer

patients in the most deprived fifth of neighbourhoods are twice as likely to receive an additional NHS QALY relative to the least deprived. To obtain RII, we simply divide the SII by the mean probability across socioeconomic groups, $\overline{q_i}$, such that $\beta_{RII} = \beta_{SII}/\overline{q_i}$. The SII is analogous to risk difference and describes the inequality on the same scale as the outcome being analysed. The RII may be more useful when comparing inequality across multiple distributions where the mean of the outcome is different (such as comparing inequality over PBCs), and is akin to a risk ratio.

3.3.6 Social welfare analysis

The gender and socioeconomic distribution of health effects can be used to calculate an 'equity-adjusted' threshold, which in turn can be used to evaluate the social welfare impacts of new interventions. We first assume that a £12,937 expenditure reduction will impose a loss of one QALY on the health system, as estimated by Claxton and colleagues. This QALY loss for each gender and socioeconomic group is then modelled on to the baseline distribution of health estimated in chapter 2 (a process described in more detail in section 4.2.3.2 of chapter 4. We can then calculate the change in social welfare by evaluating both health distributions using the functions described in section 2.3.5 of chapter 2. The difference in the equally distributed equivalent (EDE) mean health at baseline and after the QALY loss is multiplied by the population size to obtain the change in population EDE QALYs:

$$\Delta PopEDE_{A,\varepsilon} = (PostEDE_{A,\varepsilon} - BaseEDE_{A,\varepsilon})Pop$$

Where $EDE_{A,\varepsilon}$ is the EDE mean health for the distribution when using the Atkinson inequality index (Atkinson 1970) and an inequality aversion parameter ε. Thus, if the more deprived groups bear a greater share of the health opportunity costs, the social value of displacing one QALY will be greater than one and the cost-per-QALY threshold should be reduced. The size of this adjustment will depend upon the inequality aversion parameter: we therefore calculate the adjusted thresholds for a range of parameters.

The adjusted threshold is calculated by finding the reduction in expenditure that yields a loss of one EDE QALY. We do this by multiplying the distribution of one QALY by an adjustment factor and calculating $\Delta PopEDE_{A,\varepsilon}$. An optimisation algorithm is run to obtain the factor that yields a value of minus one, which can then be applied to the cost-per-QALY threshold of £12,937 to obtain the cost-per-EDE QALY threshold.

3.3.7 Sensitivity analysis

3.3.7.1 Unique patient counts

Using episode counts to infer the socioeconomic distribution of the QALY assumes that every episode within each age, gender, and ICD group is associated with an equal probability of generating a QALY regardless of socioeconomic group. To demonstrate, consider two diabetic patients, A and B, who have four and two recorded episodes, respectively. An episode tally assumes that A has twice the probability of generating a QALY, whilst a patient tally supposes that A requires twice the number of episodes as B in order to achieve the same QALY-generating probability. To determine whether this assumption is influential, we estimate our results using an alternative set of socioeconomic distributions calculated using the distributions of unique patients.

To account for the fact that some patients' age and IMD quintile group changed across episodes, these values are fixed to their values in the first episode (when listed chronologically). We wanted to capture all ICD codes associated with each patient across their episodes. However, with some patients associated with in excess of 200 episodes this becomes computationally unfeasible. We therefore count a patient in the matrix if an ICD code appears in the diagnosis codes of any of their ten most coded episodes (where most coded means the highest number of diagnosis codes entered).

3.3.7.2 Previous years of data

A second sensitivity analysis is conducted by repeating our analysis using HES data from the previous two years to test whether there are any noticeable differences in inequality over time.

3.3.7.3 Primary care data

Whilst Hospital Episode Statistics (HES) provides comprehensive coverage of inpatient secondary care utilisation by age, gender, socioeconomic status and disease (measured by International Classification of Disease (ICD) code), it may not be the most appropriate data source from which to estimate socioeconomic distributions for some clinical areas. The socioeconomic patterns observed in inpatient secondary care might not be reliable proxies for how the health benefits accruing to each disease area are distributed, especially for disease areas where the proportion of total health care activity taking place in inpatient secondary care is small.

We therefore sought data that provides information on diseases and conditions typically treated in primary care by socioeconomic status (preferably Index of Multiple Deprivation (IMD)). One such source is the Quality and Outcomes Framework (QOF) dataset, which includes information on the socioeconomic (but not age or gender) distribution of diseases at the level of general practitioner (GP) clinic. QOF is an incentive scheme for NHS GPs in the UK that provides financial rewards to each practice for achieving specific clinical goals within their patient population, known as 'indicators'. Many of the indicators involve ensuring that a sufficient proportion of the practice population with a certain condition receive a test or treatment (for example, the proportion of patients with coronary heart disease who have received an influenza immunisation). This necessitates having an estimate of the at-risk population for each condition of interest for each practice so their achievement can be measured, from which a practice-level prevalence rate can be calculated. These are provided for all conditions relating to the indicators and are published annually (Health and Social Care Information Centre 2015).

The Quality and Outcomes Framework (QOF) primary care data provides prevalence rates by disease for each GP practice in England. However, the practices do not align with LSOAs that are used to calculate local area deprivation scores using the Index of Multiple Deprivation, as some practices straddle multiple LSOAs. Therefore we use the Attribution Dataset on GP Registered Populations, which disaggregates each practice population by LSOA.

With these data, we can calculate prevalence by IMD using the following process for each condition:

i. Apply the practice-level prevalence rate to each postcode portion of the practice population, giving the expected number of cases of a condition by practice, disaggregated by postcode;
ii. Add the number of expected cases for each postcode (which is split between multiple practices). Each postcode is assigned its IMD score and IMD quintile group;
iii. Add up the number of cases for each IMD quintile group and divide by the total number of cases to obtain the relative proportions over IMD (i.e. the socioeconomic distribution);

The prevalence rates from QOF reflect differences in the utilisation of services, and therefore do not capture the differences in the likelihood of certain groups receiving the health benefits from treatment. Nor are the rates by QOF condition direct substitutes for the HES dataset, as they are not defined by ICD code or IMD quintile. We therefore mapped the conditions to their ICD codes (or subset of codes), shown in Table A3.6. The QOF distributions are then used to replace the episode distributions extracted from HES, and are applied to all the age-gender-ICD groups that constitute each condition. To investigate the impact of using this alternative source of data, we compared the socioeconomic distribution of the QALY proportions using only the ICD codes covered by the QOF dataset. These 54 codes relate to 37.4% of the health effects that result from a change in expenditure; 0.37 of each additional QALY therefore goes to patients that are covered by QOF.

3.4 Results

3.4.1 Descriptive statistics

Descriptive statistics for HES are reported in Table 3.1. In total, 119,569 (0.006%) observations are excluded from the sample. Another 51,344 are deleted as suspected duplicates, leaving a remaining sample size of 19,407,655 episodes covering a total patient population of 8,882,110.

Females accounted for a larger proportion of patients (56.1%) than males. Whilst this is also true for episodes, the proportion was slightly lower (54.5%), implying fewer episodes per female patient. A near-identical socioeconomic gradient in both episode and patient counts is found. For patients, 23.5% are in the most deprived IMD quintile group and 17.4% in the least deprived. The respective statistics using episodes are 23.3% and 16.9%. The number of episodes attributed to each PBC and IMD quintile group are provided in Table A3.5 in the appendix. The ratio of counts in the most deprived to least deprived groups ranges from 0.95 for PBC 2 (Cancers and Tumours) to 2.87 for PBC 19 (Neonates).

Table 3.1: Descriptive statistics for Hospital Episode Statistics 2012/13.

Variable	**Patients**	**% Sample**	**Episodes**	**% Sample**
Total	8,882,110	100%	19,407,655	100%
Age				
0-4	999,334	11.3%	1,463,253	7.5%
5-14	363,592	4.1%	564,144	2.9%
15-29	1,183,033	13.3%	2,141,345	11.0%
30-44	1,427,015	16.1%	2,642,378	13.6%
45-59	1,520,374	17.1%	3,297,482	17.0%
60-69	1,204,898	13.6%	2,983,189	15.4%
70-79	1,143,281	12.9%	3,201,919	16.5%
80+	1,040,583	11.7%	3,113,945	16.0%
Gender				
Male	3,896,899	43.9%	8,826,364	45.5%
Female	4,985,211	56.1%	10,581,291	54.5%
IMD				
1 (most deprived)	2,090,295	23.5%	4,530,436	23.3%
2	1,799,620	20.3%	3,998,631	20.6%
3	1,804,243	20.3%	4,018,339	20.7%
4	1,641,355	18.5%	3,571,730	18.4%
5 (least deprived)	1,546,597	17.4%	3,288,519	16.9%

Note: IMD = index of multiple deprivation

Table 3.2: Distribution of quality-adjusted life years by age and index of multiple deprivation (IMD) quintile group for a £50m change in the English National Health Service budget

	IMD Quintile Group					
Age band	**1**	**2**	**3**	**4**	**5**	**Total**
0-4	135	103	101	65	59	463
5-9	47	39	39	27	27	180
10-14	46	38	38	27	26	176
15-19	78	62	61	40	33	274
20-24	86	68	68	44	37	303
25-29	87	69	69	44	37	307
30-34	49	39	39	24	20	172
35-39	47	37	37	23	19	163
40-44	52	42	41	26	21	183
45-49	69	57	57	40	34	257
50-54	62	51	51	36	30	231
55-59	54	45	44	31	26	200
60-64	53	46	47	40	34	221
65-69	49	43	43	38	32	204
70-74	41	38	38	36	31	183
75-79	34	31	31	30	26	152
80-84	20	20	20	21	19	100
85+	20	20	19	20	18	96
Total	**1029**	**850**	**844**	**612**	**530**	**3865**

Note:

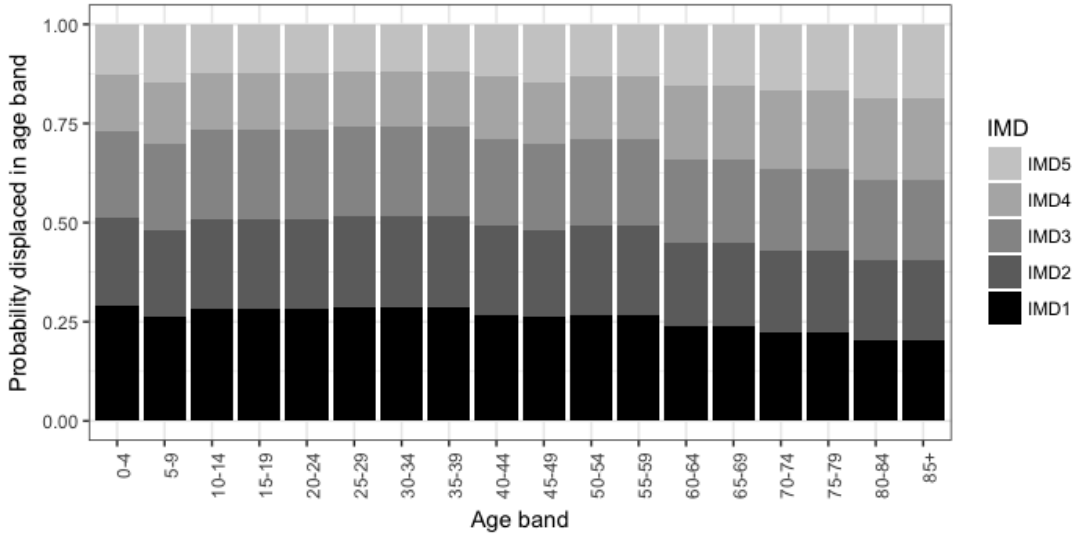
1. IMD1=most deprived, IMD5=least deprived
2. A cost-per-QALY estimate of £12,937 from Claxton et al. (2015) is used to predict the expected number of QALYs

3.4.2 Main findings

Table 3.2 demonstrates how the health benefits of a £50 million budget increase would be distributed between age, gender and socioeconomic subgroups using Claxton and colleagues' estimate that the cost of each additional QALY is £12,937. Of the 3,865 QALYs generated from the increase, nearly twice as many accrue to the most deprived fifth (1,029) as to the least deprived (530), whilst 45% of the gains go to those under 30.

The likelihood of each deprivation quintile group receiving the QALY is given in Figure 3.3. The most deprived fifth are the most likely to be affected with a probability of 0.27, a value that decreases with deprivation up to the least deprived fifth who have an associated probability of 0.14. This disparity is summarised with a negative SII of -0.08. For each IMD quintile group, females had a higher probability of being affected than males. However, the relative differences between deprivation groups were greater for men, with a RII of -0.85, compared to -0.82 for women.

Figure 3.3: How the socioeconomic distribution of quality-adjusted life year gains varies by age group

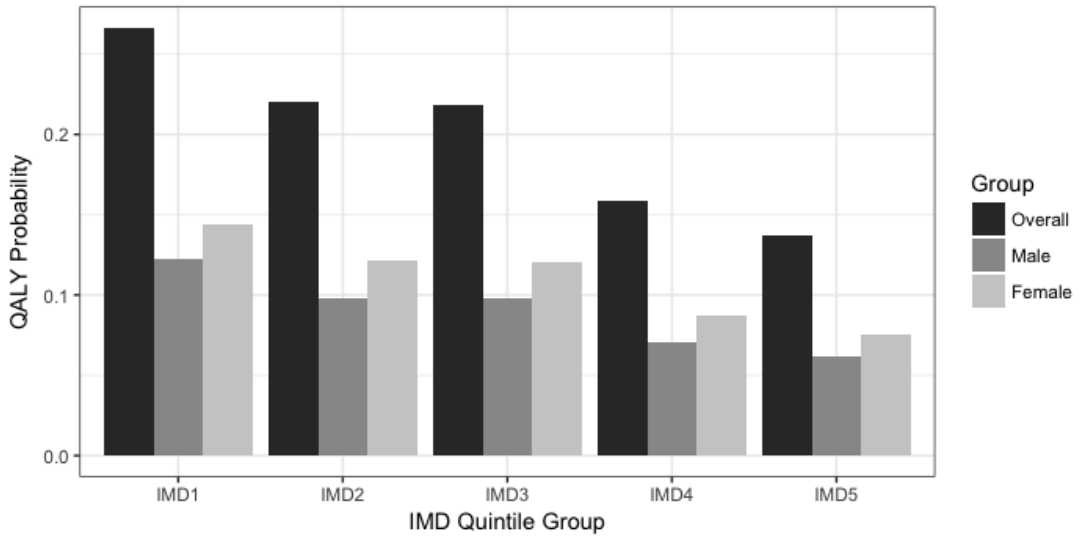


Note: Socioeconomic status is measured by Index of Multiple Deprivation (IMD) quintile (1=most deprived, 5=least deprived)

3.4.2.1 Inequality by age band

Figure 3.4 shows how the likelihood of receiving the additional QALY is distributed between socioeconomic quintiles for each age group. Inequality is most pronounced in young age bands, with a large social gradient clear from birth until the 40-44 band. RII values are consistently around -1.0 up to this group, indicating that the QALY probability for the most deprived group is twice that of the least deprived group. Thereafter disparities reduce to a minimal level, with an RII of -0.08 for the 85+ band.

Figure 3.4: Probability distribution of an additional quality-adjusted life year (QALY) over age and Index of Multiple Deprivation (IMD) quintile group



Notes:

1. IMD1 = most deprived group, IMD5 = least deprived
2. The differences in QALY effects between genders should be treated with caution. This is because the larger effects for women reflect their levels of health care utilisation rather than any systematic differences in the health care services being affected by expenditure changes

1.1.1 Inequality by Programme Budgeting Category

Table 3.3 shows how the social gradient in the probability of receiving an additional QALY differs by PBC. The respiratory programme, within which nearly 30% of effects accrue, exhibits average levels of inequality, with an RII of -0.89, with the gastro-intestinal programme being one of the most unequal with a score of -1.43. The cancer PBC distributes its gains most equally, yielding an RII of -0.05.

1.1.2 Social welfare analysis

The cost-per-EDE thresholds, which adjust the cost-per-QALY threshold based on (i) the gender and socioeconomic distribution of a displaced QALY and (ii) the level of inequality aversion, are shown in Figure 3.5. At the base case level of inequality aversion of 10.95 (estimated by Robson et al. (2016)), the adjustment factor is 0.867, yielding an adjusted threshold of £11,220. As inequality aversion increases, the added societal cost of displacing more health in poorer groups increases the

opportunity cost and reduces the threshold: when ε=30, the threshold drops to £10,097.

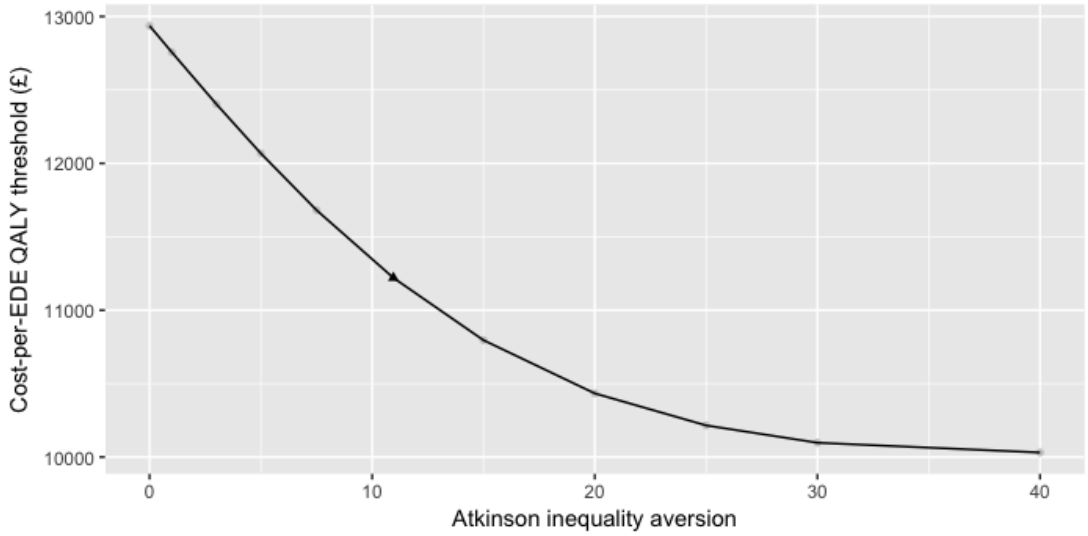
Table 3.3: Inequality in quality-adjusted life year (QALY) gains by Programme Budgeting Category (PBC)

PBC	QALY probability	QALYs from £50m spend increase	SII	RII
Total	**1**	**3865**	**-0.0848**	**-0.85**
Respiratory	0.297	1146	-0.0263	-0.89
Neurological	0.141	545	-0.0127	-0.90
Circulatory	0.139	539	-0.0096	-0.68
Mental health	0.123	476	-0.0159	-1.29
Endocrine	0.078	303	-0.0074	-0.95
Gastro-intestinal	0.057	219	-0.0081	-1.43
Cancer	0.034	132	-0.0002	-0.05
Muscuoskeletal	0.030	116	-0.0008	-0.27
Blood disorders	0.028	109	-0.0030	-1.06
Infectious diseases	0.020	78	-0.0017	-0.85
Problems of hearing	0.018	70	-0.0008	-0.42
Genito-urinary	0.014	53	-0.0005	-0.39
Dental problems	0.009	34	-0.0007	-0.80
Problems of eye and vision	0.005	21	-0.0002	-0.30
Skin	0.003	10	-0.0001	-0.35
Poisoning and A&E	0.001	4	-0.0001	-0.91
Learning	0.001	3	-0.0001	-1.30
Healthy Individuals	0.001	3	-0.0002	-1.93
Maternity + Neonate	0.001	2	0.0000	-0.89

Note:

1. SII = slope index of inequality; RII = relative index of inequality; A&E = accident & emergency
2. PBCs 16, 22 and 23 were not associated with any health effects of additional expenditure

Figure 3.5: Cost-per-EDE QALY threshold as inequality aversion increases. The base case inequality aversion parameter of 10.95 from Robson et al. (2016) is shown by the black triangle



Note: EDE = equally distributed equivalent; QALY = quality-adjusted life year

3.4.3 Sensitivity Analysis

The impact of using unique patient counts was negligible. Compared to the episode count distribution, RII slightly fell from -0.848 to -0.787, with a mean absolute difference over IMD quintile groups of 0.004. Similarly, no differences were found when using the HES datasets from 2011 or 2012, with the probability of QALY gain for each IMD quintile group staying consistent over time. These analyses are all summarised in Table 3.4.

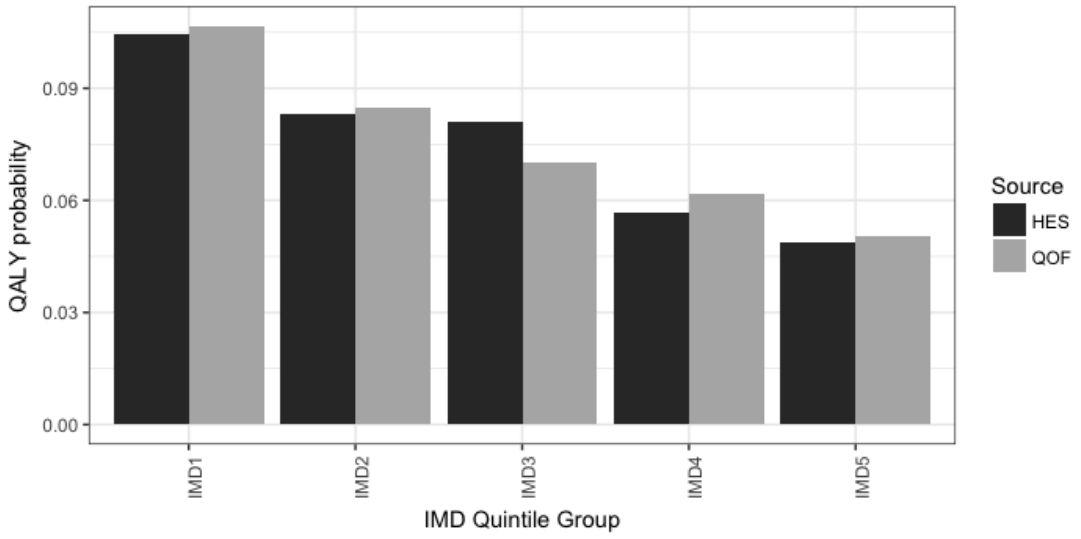
The comparison of the QALY distributions derived from the prevalence rates from QOF with those derived from the utilisation statistics from HES is shown in Figure 3.6. The gradient in the health effects, which for this analysis covers 0.37 of each additional QALY, is marginally more even when using QOF than when using HES; RII and SII decreased from -0.968 and -0.036 for HES to -0.913 and -0.034 for QOF, respectively. Stated alternatively, when analysing only the diseases included in QOF, the health changes accruing to the most deprived are 91% higher than those for the least deprived when using QOF, compared with 97% when using HES.

Table 3.4: Probability of receiving the additional quality-adjusted life year by Index of Multiple Deprivation (IMD) group using alternative data sources to estimate socioeconomic distributions, along with absolute (SII) and relative (RII) inequality measures

	IMD Quintile Group						
Data Source	**1**	**2**	**3**	**4**	**5**	**SII**	**RII**
Episodes							
2013	0.2663	0.2199	0.2184	0.1583	0.1370	-0.085	-0.848
2012	0.2677	0.2197	0.2180	0.1583	0.1363	-0.086	-0.857
2011	0.2686	0.2192	0.2186	0.1576	0.1360	-0.086	-0.865
Patients							
2013	0.2642	0.2167	0.2140	0.1619	0.1432	-0.079	-0.787

Note: IMD 1 = most deprived, IMD 5 = least deprived.

Figure 3.6: Socioeconomic distribution of the proportion of an additional quality adjusted life year (QALY) attributable to disease covered by the Quality and Outcomes Framework (QOF) dataset when using disease-specific socioeconomic patterns from Hospital Episode Statistics (HES) and QOF



Note: Includes only health effects attributable to diseases covered by the QOF dataset. These cover approximately 37% of the total.

3.5 Discussion

3.5.1 Main findings

Our analysis provides the first empirical estimate of the social distribution of health effects from marginal NHS budget changes. We find that such changes disproportionately affect the most socioeconomically deprived and are concentrated in younger age groups, where socioeconomic gradients are most pronounced.

Our results suggest that the real increases in NHS funding in recent years (Crawford & Emmerson 2012) have likely contributed to a reduction in socioeconomic health inequalities. This supports the conclusions of both Asaria et al. in their analysis of primary care access and outcomes and Barr et al., who found that proportionately larger budgetary increases for higher deprivation local authorities reduced the disparity in mortality rates to less deprived ones (Barr et al. 2014; Asaria et al. 2016).

We also provide breakdowns by broad disease area, known as "programme budgeting category" (PBC). The overall gradient is primarily driven by a small number of "influential" disease areas, where expenditure is particularly elastic with respect to overall budget changes. A key example of this is respiratory disease (PBC 3), which accounts for 30% of the change in health for a change in overall spend. Our analysis of inequality by PBC demonstrates how the aggregate socioeconomic gradient is driven by high levels of inequality in respiratory conditions such as asthma and chronic obstructive pulmonary disease.

Through sensitivity analyses, we find that the social distribution of an additional QALY is almost identical when using episode distributions from 2011 and 2012, and very similar when using primary care data on disease prevalence rather than secondary care data on health care utilisation. We can thus be satisfied that our results are not sensitive to quirks in health care utilisation specific to 2013, or to the use of secondary care utilisation data rather than prevalence data, and reflect consistent socioeconomic patterns by disease.

The social welfare analysis we conduct demonstrates an innovative application of the results by calculating equity-adjusted cost-effectiveness thresholds that reflect the size and distribution of health opportunity costs. Since opportunity costs fall more prominently on lower socioeconomic groups, we calculate adjustment factors for the cost-effectiveness threshold that are flexible to the level of inequality aversion of the social decision-maker. These thresholds can be utilised in economic evaluation as equity-informative benchmarks by which to judge the population health and health inequality impacts. For example, an intervention that costs £5m would need to generate 445 EDE QALYs instead of 386 QALYs in order to be cost-effective, as it must also account for the health inequality increases that come with displacing services that benefit those at the bottom of the health distribution.

3.5.2 Limitations and assumptions

This study is underpinned by evidence produced by Claxton and colleagues on the relationship between local expenditure and mortality, which they combined with other data to estimate the marginal effects of NHS expenditure on population QALYs. The complexity of the problem Claxton and colleagues faced, and the limited availability of data, especially in extending analyses from life years to QALYs, meant that several assumptions were made. A full list of these assumptions is given in Table 32 of their report (Claxton, Martin, et al. 2015, p.83). Other critiques and responses have subsequently been published (Barnsley et al. 2013; Claxton & Sculpher 2015; Raftery 2014). However, our analysis depends not on the absolute health effects of expenditure changes but on the relative contribution of each PBC to the overall health effect, as our focus is on distribution. Importantly, none of the critiques published to date question the plausibility of the results in this specific respect.

Additionally, our analyses assume that the effectiveness of health spending in terms of health production is the same for each socioeconomic, age and gender group conditional on PBC. Whilst it would have been desirable to estimate outcome-expenditure elasticities by socioeconomic group, this would have required expenditure and health outcome data by socioeconomic status as well as by region

and clinical area that are not currently available. Although empirical work suggests that the health outcomes from health care are generally better for less deprived groups (Cookson, Propper, et al. 2016), little work has been conducted on the direct link between health care inputs and health outputs by socioeconomic group. The impact on our results is thus unknown: the affluent may be more effective at producing health, but those in more deprived areas may only seek care when sicker, thereby obtaining more benefit.

We did not fully characterise the uncertainty around the probabilities of receiving the additional QALY. In their analysis, Claxton and colleagues propagate parameter uncertainty through their model using Monte Carlo simulation (as detailed in Chapter 5 of their report). Whilst we investigated uncertainty over episode counts from HES, we did not combine this with the uncertainty from the original analysis, thereby generating unrealistically small standard errors.

The socioeconomic episode distribution of particular diseases estimated from HES may also not be a reliable proxy for how health gains might actually be distributed following an expenditure increase. This is because the patterns observed in hospital activity might be different to those observed in other forms of health care. This is true for conditions principally treated in primary care, such as asthma, or in specialist facilities, such as schizophrenia or other mental health conditions. We are not able to obtain primary care utilisation data that could test this hypothesis. The sensitivity analysis we perform with primary care data from QOF indicates that results were marginally more evenly distributed and largely comparable to secondary care utilisation. However, as these are prevalence data, they do not account for patterns of utilisation and would not capture the additional health benefits that sicker patients in more deprived groups obtain from multiple visits to primary care, for example. Without evidence to inform how the patterns in these data sources might differ to those we estimate from HES, we are unable to speculate as to what direction of bias this might have on our results.

Last, the methods used to expand the eight age bands into 18 by using the population densities of 5-year bands within the 10- and 15-year ones likely oversimplify the SES

distribution by age. This can be seen in Figure 3.3, where small step changes can be seen every two or three bands.

3.5.3 Implications and further research

An important application of our results is their use in health technology assessment. For a universal health system like the NHS, the decision to fund a new technology will entail marginal reductions in existing services. The numbers in Table 3.2 can thus be considered as the distribution of health losses resulting from a decision to approve a new intervention costing £50 million. Our results therefore provide the first quantitative assessment of how health opportunity cost is distributed socioeconomically, and can help to inform decision-makers on what impact future interventions have on health inequality. This could be through informal consideration or a distributional cost-effectiveness model (Asaria, Griffin, Cookson, et al. 2015; Cookson et al. 2017), in which our estimates can be combined with an intervention's expected health benefits by age, gender and socioeconomic status to generate a distribution of net health effects. When these are modelled on to expected lifetime health (such as the estimates from chapter 2), decision makers can be offered evidence on how a decision might affect health inequalities. Our methods and estimates may be also be a useful building block in future work on developing resource allocation formula based on equity of outcomes rather than utilisation. However, several further theoretical and empirical challenges would need to be overcome before a robust formula of this kind could be operationalized.

There is scope to improve our estimates by using better data with which to estimate the socioeconomic distributions of age-gender-ICD subgroups. For example, linking in other datasets such as the Clinical Practice Research Datalink and the Mental Health Minimum Dataset could provide socioeconomic distributions of relevant conditions by age and gender in primary care and specialist mental health centres, respectively. Future research should also investigate the differences in health benefit achieved from receiving health care, which our analysis has assumed is the same for all socioeconomic groups. Lastly, similar analyses to this should be conducted for social care expenditure, as a comparison between the marginal effects of

expenditures of health and social care can help inform resource allocation priorities with respect to health inequalities.

However, this study demonstrates how insights into the distribution of health gains from expenditure can be made, whilst also providing an important contribution toward the role of public health care expenditure in reducing health inequalities.

Appendix 3

Table A3.5: Number of episodes from Hospital Episode Statistics by Index of Multiple Deprivation (IMD) quintile group and Programme Budgeting Category (PBC)

		IMD Quintile Group					
	PBC	**1**	**2**	**3**	**4**	**5**	**Ratio 1/5**
1	Infectious disease	73,206	65,773	65,792	55,640	51,106	1.43
2	Cancers and tumours	798,965	769,636	790,868	875,026	838,692	0.95
3	Blood disorders	288,574	260,946	266,963	218,611	193,114	1.49
4	Endocrine, nutritional	977,489	888,327	880,756	738,236	623,516	1.57
5	Mental health	1,182,758	935,190	917,735	603,987	466,905	2.53
6	Learning disability	55,110	44,280	43,390	28,797	21,897	2.52
7	Neurological	389,763	343,101	339,122	293,529	262,108	1.49
8	Vision problems	259,087	246,621	248,643	255,468	233,922	1.11
9	Hearing problems	75,419	69,193	69,268	63,735	57,986	1.30
10	Circulatory disease	2,625,616	2,415,341	2,422,808	2,299,099	2,023,139	1.30
11	Respiratory disease	1,215,577	1,050,923	1,046,605	840,736	710,081	1.71
12	Dental problems	79,505	74,205	72,900	50,626	43,713	1.82
13	Gastrointestinal system	1,288,223	1,160,694	1,170,376	1,051,384	954,478	1.35
14	Skin problems	287,882	255,144	251,871	213,796	186,235	1.55
15	Musculoskeletal system	894,024	827,430	829,136	784,789	712,146	1.26
17	Genitourinary system	1,086,429	1,001,856	1,001,596	860,605	740,267	1.47
18	Maternity	534,888	474,522	461,860	309,722	279,065	1.92
19	Neonate	227,315	129,349	124,651	87,885	79,231	2.87
20	Poisoning	21,984	18,753	19,307	15,167	13,946	1.58
21	Healthy individuals	340,182	313,719	305,312	242,280	209,867	1.62
	Total	**12,701,996**	**11,345,003**	**11,328,959**	**9,889,118**	**8,701,414**	**1.46**

Notes:

1. IMD 1 = most deprived; IMD 5 = least deprived
2. Episodes are attributed to a PBC/IMD group if they appear in any of its respective diagnosis codes. Counts therefore exceed the episode total of 19,407,655

Table A3.6: ICD codes covered by the disease included in the Quality and Outcomes Framework dataset

Disease	**ICD Codes**
Asthma	J45
Atrial Fibrillation	I48
CHD	I20-25
CKD	N18
COPD	J40-44, J47
Dementia	F00-F07
Depression	F32
Diabetes	E10-14
Epilepsy	G40-41
Heart Failure	I50
Hypertension	I10-15
Learning disability	F82
Obesity	E66
Osteoperosis	M81-82
PAD	I73
Rheumatoid Arthritis	M05-06
Stroke	I61-64, G45
Hypothyroidism	E00-07

Chapter 4: Simplified distributional cost-effectiveness analysis: an application to 27 health technologies approved in England

4.1 Introduction

Health systems around the world have increasingly turned to health technology assessment (HTA) as a process for systematically evaluating new health interventions and prioritising existing ones. The principal quantitative component of HTA has been cost-effectiveness analysis, which assesses technologies in terms of their ability to maximise health benefits relative to cost for the average patient (Mathes et al. 2013).

Health care decision-makers must also consider other aspects of benefit when evaluating treatments. Chief amongst these are how health gains are distributed and how health inequality will be affected (Marmot et al. 2010; Turrell et al. 2006; Canadian Institute for Health Information 2015; CDC 2013). However, the health inequality impacts of new interventions are not quantitatively analysed by HTA agencies. In the UK, for example, the National Institute for Health and Clinical Excellence's (NICE) methods guidance describes commitments to considering health inequalities, but does not routinely consider whether they will be improved, or by how much (NICE 2014).

Distributional cost-effectiveness analysis (DCEA) is a framework that has sought to address this shortcoming by using and extending the results of a traditional cost-effectiveness analysis (Asaria, Griffin, Cookson, et al. 2015). Health benefits and losses are disaggregated by social groups of interest (i.e. by socioeconomic status or gender), which are then combined to calculate a gradient of net effects. These can then be modelled on to a distribution of expected lifetime health to understand how health inequality might change as a result of a recommendation.

DCEAs require the estimation of the distribution of health benefits from an extended or adapted decision analytic model, such as using socioeconomic subgroup-specific

epidemiology, quality of life scores or treatment effects. This approach is exemplified in chapter 5 with respect to smoking cessation interventions. This paper outlines a simplified version that does not adapt or extend the decision model itself, but takes the average cost-effectiveness results and scales them up using patient population numbers. The population-level benefits are then disaggregated according to aggregate data on gender and socioeconomic patterns observed for the relevant disease, such as health care utilisation data. The method we propose for conducting quantitative inequality impact assessment, therefore, requires only the standard outputs of a cost-effectiveness analysis.

We apply this simplified approach to a sample of 27 NICE single technology appraisals (TAs), where one new treatment is compared against existing standard care, conducted between 2012 and 2014. We show how the net health benefits are distributed between gender and socioeconomic groups for each intervention, as well as how their implementation affects lifetime health inequality in the English population.

This new and simple way of calculating the health inequality impacts of health technologies can thus provide NHS decision-makers and stakeholders with an evidence-based technique for evaluating whether new interventions can help address the important social objective of reducing health inequalities.

4.2 Methods

4.2.1 Overview

The method for estimating the net health impacts by gender and socioeconomic group is shown in Figure 4.1. Incremental costs and benefits are extracted from the manufacturer's submission to NICE, along with patient population estimates to calculate population-level effects. Benefits are distributed according to healthcare utilisation patterns observed in Hospital Episode Statistics (HES) for the relevant disease, identified by a 3-digit International Classification of Disease (ICD) code. HES is selected as it is a convenient, consistent data source that provides information on all ICD codes. Costs are converted into health losses using a recent estimate of the

cost-effectiveness threshold by Claxton et al. (2015), which are disaggregated by gender and socioeconomic groups using the results from chapter 3. The difference between the health benefits and health opportunity costs provides net health effect for each group. These can then be modelled, both by individual TA and collectively, onto a baseline distribution of lifetime health, such as the one estimated in chapter 2, in order to assess the impact upon health inequality.

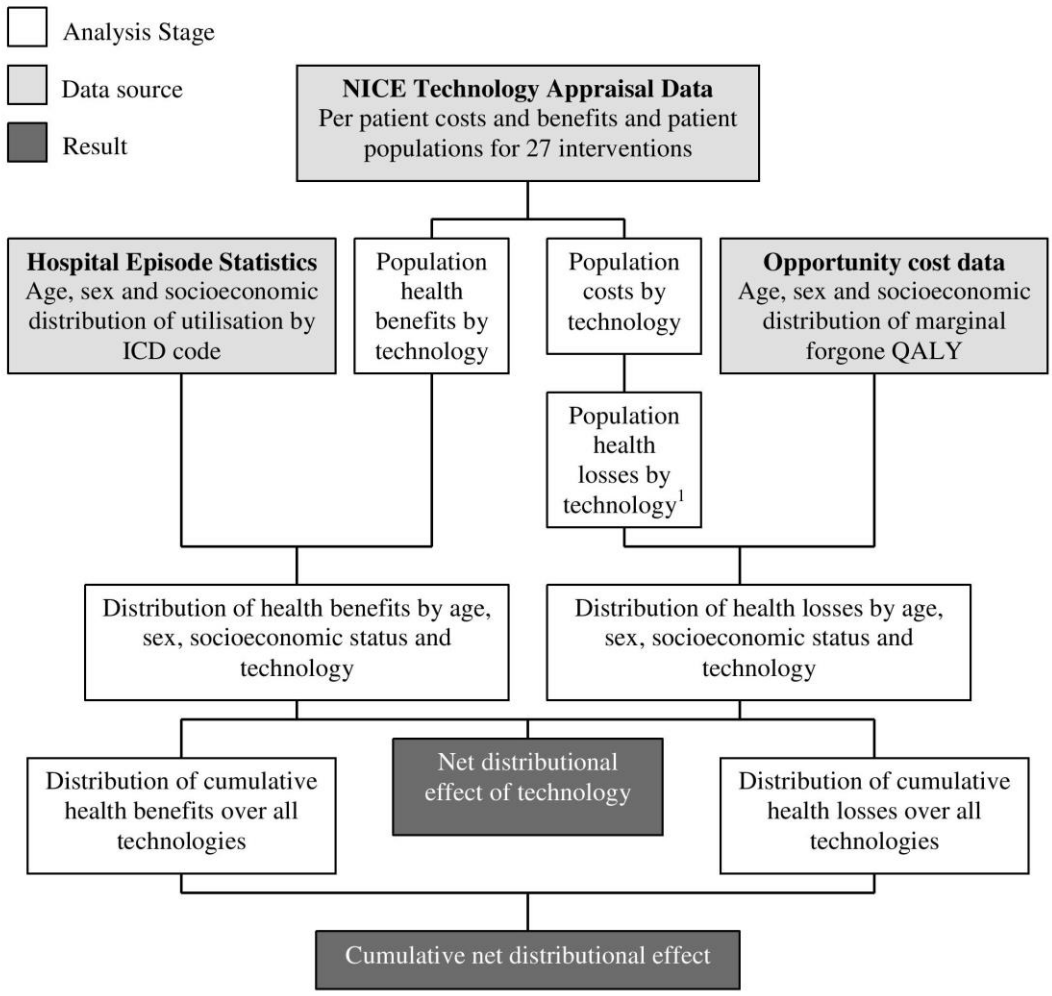
4.2.2 Data and variables

4.2.2.1 NICE Technology Appraisal Data

Information on population health benefits and costs is taken from cost-effectiveness evidence and associated reimbursement decisions for NICE single TAs issued between January 2012 and November 2014. Information is extracted from guidance documents, manufacturers' submissions and costing templates (CTs), all obtained from the NICE website. Where multiple treatments for the same condition are appraised separately, these are treated as independent and no attempt is made to combine the results. This may occur, for example, where one treatment has been recommended and subsequently another treatment has been appraised and found to be superior.

For each TA, we extract information on the expected lifetime costs and health benefits of both the new treatment and the comparator treatment (or treatments). Health benefits are expressed in quality-adjusted life years (QALYs). All costs and health benefits had been discounted at a rate of 3.5% in line with NICE's methods guidance. The task of identifying the most plausible set of estimates is often complicated by the reluctance of the appraisal committee to specify the exact estimates used to inform their decision. For the purposes of this study, we use the expected costs and benefits put forward by the manufacturers in its base case scenario, since these were available for all TAs.

Figure 4.1: Influence diagram demonstrating how our data sources are combined to estimate the net distributional effect of interventions



Note: NICE = National Institute of Health and Care Excellence; ICD = International Classification of Disease; QALY = Quality-adjusted life year

[1] Costs are converted into health losses using the cost-effectiveness threshold of £12,937 estimated by Claxton et al. (2015)

Information on the number of patients in England who would be eligible for treatment is extracted from the costing templates provided by NICE, covering incident cases and, where appropriate, prevalent cases. Thus, the proceeding calculation of population net health benefits assumes that the intervention will be provided to all eligible patients. For appraisals involving multiple comparator treatments, we calculate a 'blended' estimate of incremental costs and QALYs. First,

the incremental data are extracted for each comparator from the manufacturer's submission; each was then assigned a weight according to the proportion of its current market share, the data on which is provided in the TA costing template. The blended estimates of incremental costs and QALYs are then calculated as the weighted average over the comparators.

We exclude TAs from our analysis if (i) the appraisal committee did not recommend the treatment for adoption into the NHS, (ii) it was an update of a previous appraisal that did not change the adoption decision, or (iii) relevant information was withheld on the grounds of it being *commercial in confidence*. The latter is typically the case when manufacturers negotiated a patient access scheme with the Department of Health that allowed patients access to the new treatment at a reduced price. We also exclude multiple technology appraisals that compared a number of new technologies on the ground that the detailed manufacturer's submissions were not made publicly available on the NICE website.

4.2.2.2 Hospital Episode Statistics

HES is a database containing extensive information on all NHS funded activity in public and private hospitals in England. The primary unit of measurement is the 'finished consultant episode'; patients whose care is transferred between consultants during a single stay in hospital may have multiple episodes. The HES data include a wide range of socio-demographic and geographical variables for each patient, including a unique patient identifier code, age, gender and postcode. The latter is used to assign the patient a deprivation score using the 2004 Index of Multiple Deprivation (IMD). The IMD is a weighted index of 38 variables covering seven dimensions of deprivation and is described in full in section 3.3.2 of chapter 3. We again use the 2004 version of IMD because it is the one provided in the years of HES we analyse.

HES are used here to calculate the socioeconomic and gender distribution of health benefits. We take two years of HES data (financial years 2011 and 2012) and count the number of episodes associated with each of the 1562 3-digit ICD codes that make up NHS spending, then disaggregate them by gender and IMD group.

4.2.2.3 Health opportunity costs

Costs are first converted into health losses using a value representing the cost-per-QALY of likely displaced NHS services at the margin. We use the best available estimate of £12,937 derived by Claxton et al (2015). The details of this approach are detailed in section 3.2 of chapter 3. Each QALY that comprises this health opportunity cost is distributed over gender and socioeconomic groups using the results from chapter 3. These are shown in Table 4.1. A pro-poor distribution is observed, with 27% of health losses incurred by the IMD1 compared with 13% for IMD5. Health losses also fall more heavily on females (55%) than males (45%).

Table 4.1: Estimate of the gender and socioeconomic distribution of health opportunity costs in the English NHS

Gender	**IMD1**	**IMD2**	**IMD3**	**IMD4**	**IMD5**
Males	0.12	0.10	0.10	0.07	0.06
Females	0.14	0.12	0.12	0.09	0.08

Note: Socioeconomic status measured by Index of Multiple Deprivation quintile group (IMD1=most deprived)

4.2.3 Analysis

4.2.3.1 Modelling net health changes

The NICE TA data are used to calculate the population benefits and costs for each TA. The incremental costs and QALYs are multiplied by the population using each comparator included in the TA. By summing over each of the J comparators in TA t, we can calculate NPB_t, the net population benefit of that TA:

$$NPB_t = \sum_{j=1}^{J} h_{tj}p_{tj} - \sum_{j=1}^{J} \frac{1}{k}(c_{tj}p_{tj}) = PB_t - PC_t$$

Where h_{tj} is the incremental QALYs, p_{tj} the patient population, c_{tj} the incremental costs, k the cost-effectiveness threshold and PB_t and PC_t the population benefits and costs, respectively. By dividing through PB_t and PC_t by $\sum_j p_{tj}$, we obtain the 'blended' incremental health and costs for each technology.

Two additional TAs are excluded at this stage: TA308 because costs and QALYs could not be identified for the recommended subgroups, and TA311 due to a mismatch between the comparators in the submission and the costing template. Mismatches also occur for comparators in other TAs. Where the comparator appears in the submission but not the costing template. This is deemed acceptable since the latter represents current practice and so no net health change is expected. Where comparators are included in the costing template but not the submission, we are unable to model the net health impacts of switching those patients onto the new treatment, thereby reducing the patient population.

The next step involves estimating the health benefits likely to be accrued for different genders and socioeconomic groups in each appraisal. In order to do this, we seek to extract the socioeconomic and gender distribution of health care utilisation from the HES data for each of the relevant disease areas. Each TA is first allocated to a 3-digit ICD code (or group of codes) via its respective disease area. These mappings are given in Table 4.2. The gender and socioeconomic distributions relating to the disease for each TA are then extracted from HES, and are used as a proxy for the distribution of benefit.

We then obtain the net benefits from implementing each technology by subgroup:

$$NPB_{tds} = PB_t z_{tds} - PC_t u_{ds}$$

Where NPB_{tds} is the net population benefit accruing to deprivation group d and gender s from TA t, z_{tds} are the proportions estimated from HES described above and u_{ds} are the proportions health opportunity costs accruing to each subgroup.

4.2.3.2 Inequality impacts

To model changes in lifetime health inequality, we first extract estimates of quality-adjusted life expectancy (QALE) at birth by gender and IMD quintile group provided in chapter 2. Post-intervention QALE is estimated by adding the net QALY benefits for each TA to the lifetime QALYs for each respective subgroup:

$$Q`_{tds} = \frac{Q_{ds}n_{ds} + NPB_{tds}}{n_{ds}}$$

Where Q_{ds} and $Q`_{tds}$ are baseline and post-intervention QALE, respectively, and n_{ds} is the population of the deprivation-gender subgroup. Combining each subgroup's QALE estimate with its respective population figure and ordering the whole population from least to most healthy yields univariate distributions of pre- and post-intervention health.

4.2.3.3 Inequality measures

Applying inequality measures to these respective distributions informs us how health inequality may change as a result of the intervention. We choose two inequality measures for this analysis: (1) the slope index of inequality (SII), which is an absolute measure of inequality, changes in which are only sensitive to absolute changes in health and not baseline levels of health, and (2) the relative inequality index (RII), which is a relative measure of inequality, changes in which are sensitive to baseline levels of health as well as changes. The SII measures absolute inequality using Ordinary Least Squares regression to estimate the difference in QALE between the least and most healthy population quintile groups. SII is estimated using the following regression model:

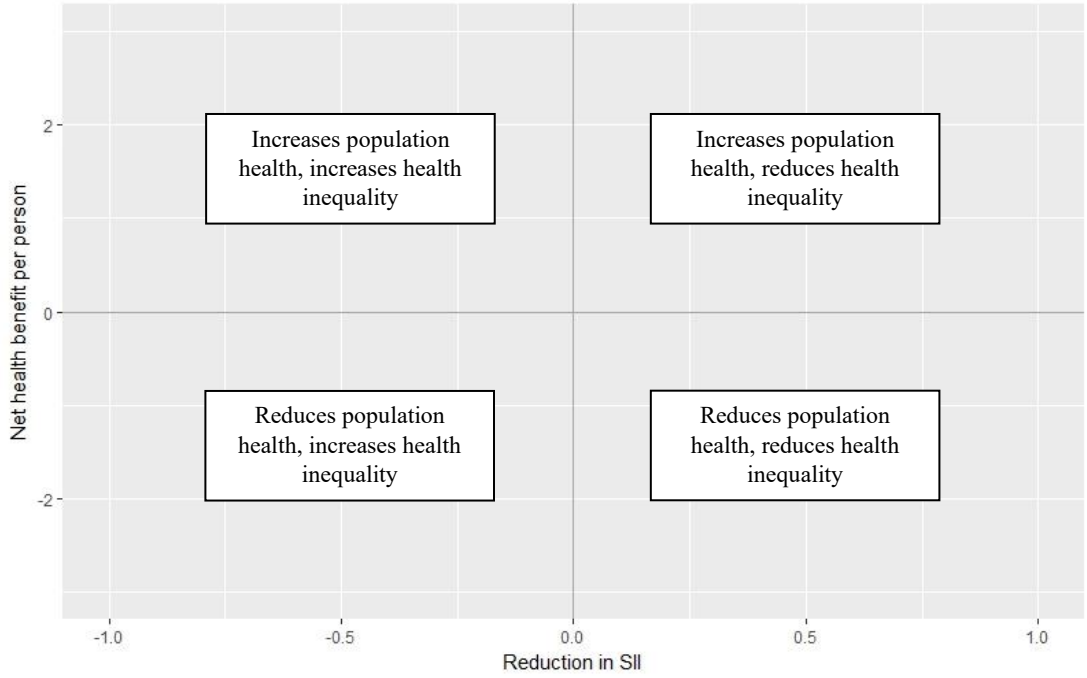
$$Q_r = \alpha + \beta_{SII}r + \varepsilon_r$$

Where Q_r is the QALE estimate of population quintile group r, β_{SII} is the SII value, ε_r is the idiosyncratic error and α is the constant term. β_{SII} is interpreted as the fitted difference in QALE when moving from the least to most healthy population quintile; an SII of 10 would mean the healthiest fifth of the population experience 10 more lifetime QALYs than the least healthy fifth. RII, our relative measure, can be interpreted as the relative change in QALE when moving from the least to most healthy population quintile. For example, a RII value of 0.1 would mean that the healthiest experience 10% more lifetime QALYs than the poorest. RII is obtained by dividing the SII by the mean QALE in the population, $\bar{Q}$, such that $\beta_{RII} = \beta_{SII}/\bar{Q}$. The inequality impact is the difference between SII values pre- and post-intervention:

we report the reduction in SII/RII so that a positive value means health inequality has reduced.

To represent the impact of interventions on health inequalities and total population health, we plot SII reductions against the per person net health benefits on the health equity impact plane. Interventions that have a positive incremental net health benefit increase total health and fall in the north of the plane. Interventions that reduce inequality as measured by the slope index of inequality fall in the east of the plane. Interventions that fall in the northwest and southeast quadrants are those for which some trade-off between the objectives of inequality reduction and total health improvement exists. This orientation of the plane is used in line with guidance provided by Cookson et al. (2017).

Figure 4.2: Equity impact plane



Note: SII = slope index of inequality

The impacts of each intervention on health-related social welfare is analysed using social welfare functions, as described in chapter 2 (section 2.3.5). We use the Atkinson and Kolm indices to measure social welfare changes solely as a function of

changes in the population distribution of health. The Atkinson index, A_ε, measures inequality relatively and is given by:

$$A_\varepsilon = 1 - \left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{Q_i}{\bar{Q}}\right)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}$$

Where N is the total population, Q_i is the QALE estimate of the ith individual, $\bar{Q}$ is the mean QALE and ε the inequality aversion parameter that quantifies the concern for relative inequality. Alternatively, the Kolm index, K_α, incorporates inequality on an absolute scale, where absolute inequality aversion is represented by the parameter α:

$$K_\alpha = \left(\frac{1}{\alpha}\right)\log\left(\frac{1}{N}\sum_{i=1}^{N} e^{\alpha[\bar{Q}-Q_i]}\right)$$

For both indices, we again use base case inequality aversion estimates of 10.95 for ε and 0.15 for α, which are estimated from a survey of the general public in England by Robson et al (2016). These values implicitly weight health gains for the least healthy fifth of the population six to seven times more highly than gains for the healthiest. Social welfare is calculated by combining each index with the mean level of health in the distribution to obtain the 'equally distributed equivalent' (EDE) level of health:

$$EDE_{A,\varepsilon} = (1 - A_\varepsilon)\bar{Q}$$
$$EDE_{K,\alpha} = (\bar{Q} - K_\alpha)$$

Where $EDE_{A,\varepsilon}$ and $EDE_{K,\alpha}$ are the Atkinson and Kolm welfare scores, respectively. The population equally distributed equivalent multiplies these values by the population size, and is the level of population health (expressed in QALYs) in a completely equal distribution that yields an equivalent amount of social welfare to the distribution being evaluated. We calculate the population EDEs pre- and post-intervention; with the difference indicating the change in health-related social welfare. Positive values indicate that health-related social welfare has increased: for

example, if an intervention is estimated to increase net population health by 100,000 QALYs and to reduce health inequality such that the Atkinson index fell by 0.01, the population equally distributed equivalent change in QALE would be (1-(-0.01))*100,000 = 101,000. The increase in social welfare associated with the inequality reduction would therefore be equivalent to 1,000 QALYs.

4.2.3.4 Sensitivity analysis

We investigate the possibility that the incremental costs and QALYs cited in the manufacturer submissions may be biased in favour of the new treatment. To do this, we use a recent study by Versoza et al. (2015) that found that manufacturer incremental cost-effectiveness ratios (ICER) were £6,200 lower on average than those found by the Evidence Review Groups (ERGs) employed to independently evaluate the manufacturer's analyses. Assuming that the latter estimates are closer to the 'true' ICER, we use this number to adjust our data. However, this ICER increase does not inform us as to how much of the change is due to higher incremental costs or lower incremental health. We therefore calculate two new datasets: (i) one where the incremental costs are adjusted such that the ICER change is solely due to underestimated costs (health remains constant) and (ii) one where the incremental benefits are adjusted such that the ICER change is solely due to overestimated benefits (costs remain constant). The cost adjustment is given by $\Delta c = (c/h + 6200)h$ and the health adjustment is given by $\Delta h = c/(c/h + 6200)$, Where Δc and Δh are the respective incremental cost and QALY adjustments required in order for the ICER to increase by £6,200 and c and h are the respective baseline incremental cost and health.

Some interventions are health improving and cost saving (known as dominant); these result in negative ICERs that are difficult to interpret and cannot be adjusted using the figure from Versoza et al. (2015). We predict the adjustments for both of the analyses described above from a linear regression analysis. We regress the cost/QALY change (due to the £6,200 adjustment) on baseline costs/QALYs for the sample of non-dominant interventions. The following equations are therefore estimated:

$$\Delta c_t = \beta_0 + \beta_1 c_t + \beta_2 c_t^2 + \varepsilon_t$$
$$\Delta h_t = \alpha_0 + \alpha_1 h_t + \mu_t$$

Where ε_i and μ_t are the idiosyncratic error terms. The equations are used to predict the cost/QALY changes for the dominant interventions so that the full sample of technologies can be included in the sensitivity analysis.[8]

Two further sensitivity analyses are conducted. First, in order to explore the impact of the distribution of health opportunity costs on our results, we assume they are distributed evenly over gender and socioeconomic groups, rather than falling more heavily on females and those with low socioeconomic status. We also explore the effects of the inequality aversion parameter and the value of the cost-effectiveness threshold on results.

4.3 Results

4.3.1 Descriptive statistics

Details on the incremental health, costs and cost-effectiveness for each technology, blended over their respective comparators, are reported in Table 4.2. Incremental health ranges from 0.01 to 1.57 QALYs per person and incremental costs range from savings of £6,200 to costs of £46,935. In terms of cost-effectiveness, these blended estimates yield ten dominant interventions, whilst the highest ICER of £133,523 is reported for dimethyl fumarate for multiple sclerosis.

As several TAs covered the same disease area, 18 health benefit distributions were extracted from HES. 14 of these distributions are pro-poor, with the steepest gradients seen for hepatitis C and alcohol dependence. The remaining four had roughly uniform distributions and included multiple sclerosis and atrial fibrillation. These are shown in full in Table A4.5 in the appendix. The net population benefits

[8] Where baseline incremental costs were negative, the absolute value is used so that a positive adjustment is made (i.e. incremental costs become less negative or positive).

are shown in Table 4.3. 19 interventions had a positive net health impact, the highest of which was apixaban for atrial fibrillation, with 62,745 population QALYs.

Table 4.2: Sample of technology appraisals (TAs) used in the analysis

TA	Technology	Disease Area (ICD code)	Inc. Health	Inc. Costs	ICER	Patients
245	Apixaban	Thromboembolism (I82)	0.04	-£244	Dominant	91,100
248	Exenatide	Type 2 Diabetes (E11)	0.08	-£282	Dominant	39,765
249	Dabigatran etexilate	Atrial Fibrillation (I48)	0.19	£1,410	£7,501	137,124
252	Telaprevir	Hepatitis C (B18)	0.97	£10,930	£11,226	17,456
253	Boceprevir	Hepatitis C (B18)	1.35	£8,508	£6,296	17,456
254	Fingolimod	Multiple Sclerosis (G35)	0.69	£19,012	£27,429	2,449
256	Rivaroxaban	Atrial Fibrillation (I48)	0.04	£740	£18,974	137,124
260	Botulinum	Migraine (G43)	0.09	£543	£6,033	35,180
261	Rivaroxoban	Deep vein thrombosis/ Pulmonary embolism (I26)	0.02	-£258	Dominant	39,828
265	Denosumab	Bone Cancer (C40, C41)	0.01	-£1,351	Dominant	86,656
266	Mannitol	Cystic Fibrosis (E84)	1.57	£46,935	£29,895	200
267	Ivabradine	Coronary Heart Disease (I50)	0.28	£2,376	£8,486	10,466
275	Apixaban	Atrial Fibrillation (I48)	0.24	£1,326	£5,498	452,463
283	Ranibizumab	Macular Oedema (H35)	0.24	£1,581	£6,457	10,663
287	Rivaroxoban	Thromboembolism (I82)	0.06	£591	£9,821	18,497
288	Dapagliflozin	Type 2 Diabetes (E11)	0.25	-£99	Dominant	155,086
292	Aripriprazole	Bipolar I Disorder (F31)	0.01	-£686	Dominant	20
297	Ocriplasmin	Vitreomacular Traction (H43)	0.09	£1,781	£20,777	954
303	Teriflunomide	Multiple Sclerosis (G35)	0.30	-£6,200	Dominant	9,780
306	Pixantrone	B-cell Lymphoma (C85)	0.20	£4,759	£23,796	1,650
312	Alemtuzumab	Multiple Sclerosis (G35)	1.10	-£3,424	Dominant	6,906
315	Canagliflozin	Type 2 Diabetes (E11)	0.11	£547	£4,939	711,444
318	Lubiprostone	Chronic idiopathic constipation (K59)	0.00	-£20	Dominant	25,500
320	Dimethyl fumarate	Multiple Sclerosis (G35)	0.24	£31,979	£133,523	4,891
322	Lenalidomide	Myelodysplastic syndrome (D46)	0.72	£17,677	£24,551	200
325	Nalmefene	Alcohol dependence (F10)	0.07	-£397	Dominant	57,820
326	Imatinib	Gastrointestinal Stromal Tumours (D37)	1.43	£22,931	£16,036	170

Notes:

1. Inc. = incremental; ICER = incremental cost-effectiveness ratio; ICD = International Classification of Disease
2. The incremental costs, benefits are 'blended' estimates, calculated by combining the estimates for each technology over their relevant comparators and combining them into one figure, weighted by their respective patient populations.

4.3.2 Health inequality impacts

14 technologies had a lower post-intervention SII compared with pre-intervention, indicating that health inequality has been reduced. The changes in SII are shown in Table 4.3. The largest inequality reduction of 0.00048 is found for canagliflozin for type 2 diabetes. Of the 13 technologies that increased inequality, the largest increase in SII of 0.00053 is found for apixaban for atrial fibrillation.

When the change in SII is plotted against per person net benefit on the health equity impact plane in Figure 4.3, we see that all 14 inequality reducing interventions are also health improving. However, of the 13 inequality increasing interventions, eight also reduce population health. This leaves five interventions which involve a trade-off between health gain and equity, as they increase health but also inequality.

The analysis of social welfare is also shown in Table 4.3. Taking into account both health improvement and the change in inequality, 19 have a positive impact on social welfare – the same 19 that provide net health improvements. The additional social value of reducing inequality, as defined by the difference between population NHB and EDE, is largest for canagliflozin for type 2 diabetes. The inequality reductions for this intervention are, when using the Atkinson index (and an inequality aversion parameter of 10.95), equivalent to 5,981 additional QALYs.

The impact of inequality aversion on the ranking of interventions is shown in Figure 4.4. When we place a greater concern on inequality, apixaban is replaced as the most valuable intervention by canagliflozin. Pro-poor interventions, such as those for hepatitis C (252) and coronary heart disease (267) also rise in rank, whilst others for atrial fibrillation (275, 249) are demoted due to their negative health inequality impacts.

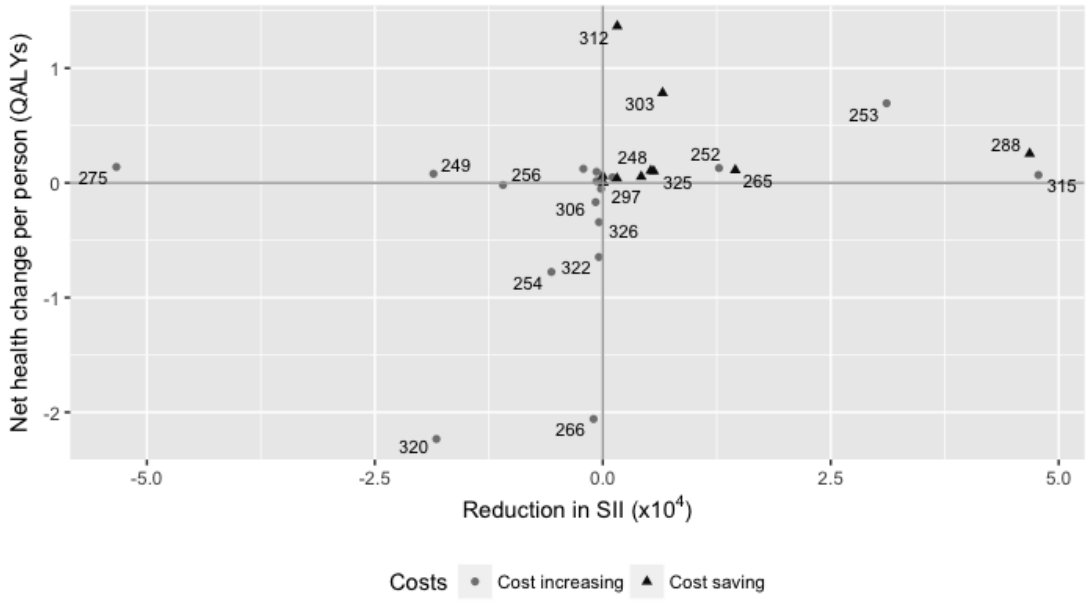
Table 4.3: Health, inequality and social welfare impact of each technology

			Inequality measures		*Social welfare measures*	
TA	**Technology**	**Population NHB**	**ΔSII**	**ΔRII**	**$\Delta EDE_{K,\alpha}$**	**$\Delta EDE_{A,\varepsilon}$**
245	Apixaban	4,917	0.00004	0.0000008	5,658	5,845
248	Exenatide	4,230	0.00005	0.0000010	4,684	4,815
249	Dabigatran etexilate	10,834	-0.00019	-0.0000022	9,169	9,286
252	Telaprevir	2,248	0.00013	0.0000019	4,722	5,021
253	Boceprevir	12,109	0.00031	0.0000050	16,635	17,334
254	Fingolimod	-1,902	-0.00006	-0.0000009	-2,347	-2,426
256	Rivaroxaban	-2,496	-0.00011	-0.0000017	-3,417	-3,542
260	Botulinum	1,690	0.00001	0.0000003	1,672	1,701
261	Rivaroxoban	1,541	0.00002	0.0000003	1,670	1,714
265	Denosumab	9,661	0.00015	0.0000026	10,821	11,124
266	Mannitol	-412	-0.00001	-0.0000002	-475	-488
267	Ivabradine	1,008	-0.00001	-0.0000001	971	991
275	Apixaban	62,745	-0.00053	-0.0000050	57,747	58,778
283	Ranibizumab	1,308	-0.00002	-0.0000002	1,085	1,091
287	Rivaroxoban	268	-0.00001	-0.0000001	351	371
288	Dapagliflozin	39,436	0.00047	0.0000085	43,548	44,757
292	Aripriprazole	1	0.00000	0.0000000	1	1
297	Ocriplasmin	-50	0.00000	0.0000000	-57	-58
303	Teriflunomide	7,667	0.00007	0.0000014	8,222	8,430
306	Pixantrone	-277	-0.00001	-0.0000001	-301	-308
312	Alemtuzumab	9,435	0.00002	0.0000008	9,619	9,830
315	Canagliflozin	48,668	0.00048	0.0000087	53,185	54,649
318	Lubiprostone	60	0.00000	0.0000000	66	67
320	Dimethyl fumarate	-10,919	-0.00018	-0.0000032	-12,393	-12,750
322	Lenalidomide	-129	0.00000	-0.0000001	-167	-173
325	Nalmefene	5,880	0.00006	0.0000010	7,086	7,349
326	Imatinib	-58	0.00000	-0.0000001	-82	-85

Note:

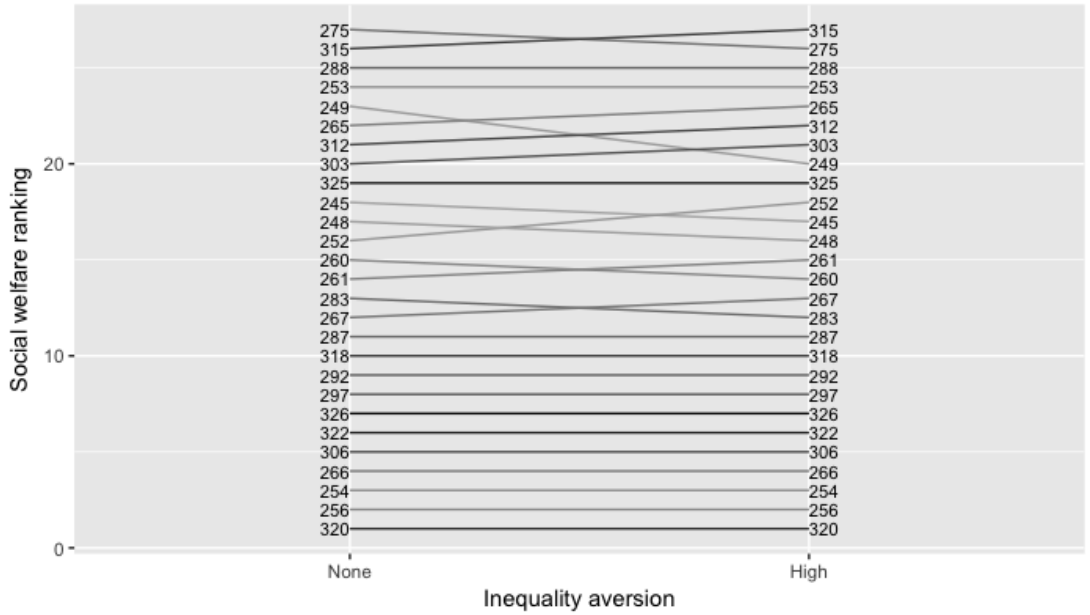
1. SII = slope index of inequality; RII = relative index of inequality; EDE = equally distributed equivalent
2. Inequality aversion parameters of 10.95 and 0.15 are used to calculate the Atkinson and Kolm EDEs.

Figure 4.3: Equity impact plane showing the change in net QALYs generated by a treatment and the impact on lifetime health inequality



Note: SII = slope index of inequality

Figure 4.4: Change in Atkinson social welfare ranking of sample technologies when moving from no inequality aversion (ε=0) to high inequality aversion (ε=20)



4.3.3 Sensitivity analysis

The results from the linear regressions that are used to predict the adjusted incremental costs and QALYs for the interventions with negative ICERs are given in Table A4.6. The effects of changing base case assumptions on the location of interventions on the equity impact plane are shown in Figure A4.7 to Figure A4.9. The changes in the numbers of interventions in each quadrant are summarised in Table 4.4. When the ICER increase of £6,300 is attributed to higher costs, 11 interventions moved to a new quadrant in the equity impact plane, with a majority moving from being health increasing and inequality reducing ('win-win') to vice versa ('lose-lose'). When the ICER increase is attributed to lower QALYs, the corresponding number is seven. For both scenarios, one intervention (TA252) became inequality moved from the north to south east quadrant. When equal opportunity costs over gender and socioeconomic status are assumed, most interventions shift horizontally to the right: the number increasing health and reducing inequality rose from 14 to 16 and the number reducing health and inequality rose from zero to eight.

Table 4.4: Total health and health inequality impacts of the 27 technologies for each scenario analysis

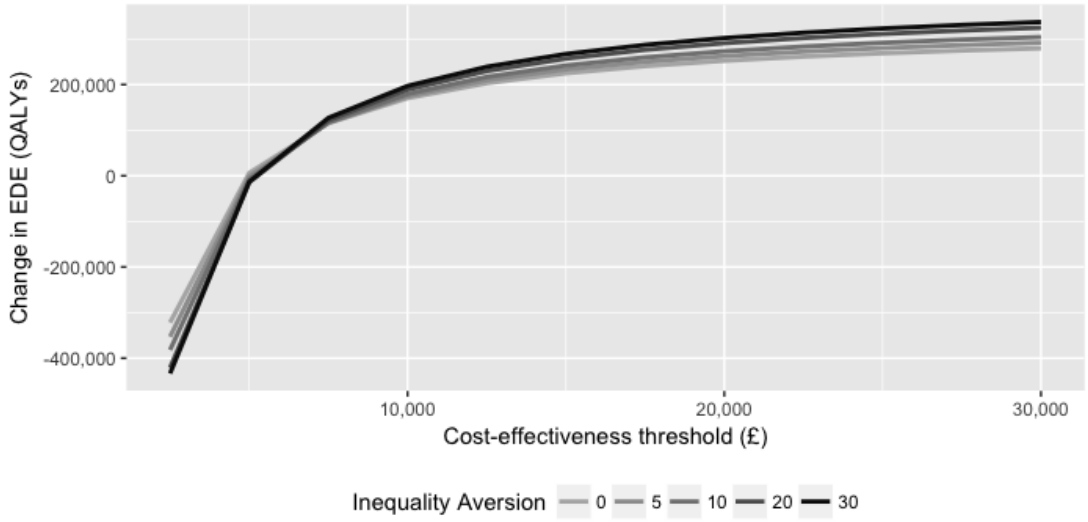
Health / inequality impact	**Manufacturer estimates**	**ERG-adjusted costs**	**ERG-adjusted QALYs**	**Uniform HOC**
Increase health, reduce inequality	14	6	10	16
Increase health, increase inequality	5	5	4	3
Reduce health, reduce inequality	0	1	1	8
Reduce health, increase inequality	8	15	12	0

Note: ERG = evidence review group; QALYs = quality-adjusted life years; HOC = health opportunity costs

The joint impact of the cost-effectiveness threshold and inequality aversion parameter are shown in Figure 4.5. The cumulative social welfare impact of the 27 interventions becomes positive between threshold values of £5,000 and £6,000. At a

threshold of £30,000, the change in population EDE ranges from 280,000 QALYs for ε=0 and 340,000 for ε=30.

Figure 4.5: Effect of the cost-effectiveness threshold and inequality aversion parameter on the cumulative net health impact of 27 health technologies



4.4 Discussion

4.4.1 Principal findings

This study proposes a simple method to quantitatively analyse the expected health inequality impacts of new health technologies, and applies this simplified form of DCEA to retrospectively estimate distributional impacts for a wide range of recommended interventions. Five interventions in our sample involve a trade-off between health inequality and health improvement. However, when these effects are combined using the Atkinson and Kolm social welfare indices incorporating general population estimates of inequality aversion, a positive change in EDE is still observed, indicating that the increases in health inequality are compensated for by the total health improvements.

Although the data on NICE TAs is systematically extracted from published documentation, our results do not constitute a health inequality impact analysis of NICE decisions over the time period. The assumption that treatments will acquire

100% market share is highly optimistic in calculating population net benefit, which will thus be overestimated. Using the incremental costs and QALYs cited by the manufacturer is similarly optimistic as we expect them to be biased towards the new treatment. When the manufacturer estimates are adjusted to reflect the average difference in cost-effectiveness with the ERG analyses, the number of interventions involving trade-offs remained relatively constant, as interventions moved from the 'win-win' quadrant to the 'lose-lose'. Our results are more sensitive when costs were underestimated, as opposed to when health gains were overestimated. This is explained by the fact that higher costs impose additional health opportunity costs, for which there is a strong socioeconomic gradient. When we assume that health opportunity costs are uniformly distributed, cost-increasing interventions shift toward the east of the plane whilst cost saving interventions move west. The overall impact is to move interventions into the trade-off quadrants, and illustrates how the distribution of opportunity costs plays an important role in the distributional impact and has contrary effects depending on the cost impact on the health sector.

4.4.2 Limitations and strengths

A number of assumptions that we adopt, along with some limitations in the available data, point to aspects that future work should focus upon to improve the robustness of this approach.

The main limitation of this type of simplified analysis is that the health benefits are distributed based only on the patterns observed in secondary care utilisation. Although the technologies included in our sample are largely administered in secondary care, our results do not reflect other factors influencing the heterogeneity in health gains, such as treatment effectiveness and uptake. As evidence suggests that treatment uptake and adherence is improved in higher socioeconomic groups (Cookson et al. 2016), we would expect that including these factors would reduce the pro-poor socioeconomic gradient in health benefits and shift interventions to the left on the health equity impact plane. This greater productivity from health care inputs in more educated groups is also supported by the theory of health production (Grossman 1972), although these effects may be smaller with the interventions in our sample when compared with behaviour change interventions, for example. Although we do not model uptake in this analysis, it can be easily incorporated in this

simplified approach. A comparison between the simplified approach (both with uniform and differential uptake included) and a full DCEA is made in chapter 5.

It is also the case that HES may not be an appropriate proxy for distributing the expected benefits of a new technology for some diseases. For many chronic conditions such as asthma or diabetes the majority of activity will take place in primary care, whilst mental health treatment primarily takes place in specialist centres not included in HES. If the socioeconomic distribution of activity recorded in these settings were systematically different to that seen in hospitals then net distributional effect estimated using our current approach would be incorrect. The sensitivity analysis conducted in chapter 3 (section 3.4.3) suggests that the socioeconomic distributions observed in the Quality and Outcomes Framework primary care dataset, which relate to prevalence rather than utilisation, do not markedly differ, however.

There are also a number of intervention comparators not included in the manufacturer's submissions, even though the costing template indicated their usage in clinical practice. For example, the costing templates for TAs 249 and 256 suggest that nearly 1.4 million patients receive either no treatment or aspirin. Since the incremental benefits of switching patients on these regimens to the new treatment are not captured, they cannot be factored in to any inequality impact analysis presented to a decision-maker. How these exclusions affect our results depends on the socioeconomic gradient of the disease area, whilst the magnitude of the health inequality impact, whether the intervention increases or reduces disparities, will be greater if more patients are included in the population health change.

Another feature of our data was the lack of recommended oncology-related interventions, with only two of the 27 in our sample intended for cancer patients. This may be due to (i) the confidential discounts agreed for patient access schemes and (ii) the operation of the Cancer Drugs Fund (CDF) over this period. How this impacts upon our results is uncertain, since the gradient of any net health benefits depends upon the type of cancer. Incidence ranges from highly pro-poor for laryngeal and lung cancer, to pro-rich for the likes of breast cancer and malignant melanoma (Public Health England 2014).

The mapping of ICD codes to some of the disease areas is also inexact, with the 3-digit codes occasionally too broad for the indication for which the intervention is intended. An example of this is for TAs 252 and 253 for Hepatitis C patients. The most appropriate 3-digit code for this disease is B18; however this counts all chronic hepatitis patients including Hepatitis B, potentially distorting the socioeconomic pattern we extract.

Last, we do not account for parameter uncertainty in our analysis. Ideally we would like to jointly simulate from the probability distributions of benefit and cost proportions as well as the incremental costs and QALYs themselves. However, informative standard errors for any of these quantities were not available to us and this could not therefore be reflected in our results.

This study does, however, provide decision-makers with unique information on a decision criterion important to the NHS and society, and requires only limited information on socioeconomic variation. A large sample of interventions using systematically extracted data is utilised to analyse past NICE recommendations and the best available data on the distribution of health opportunity costs are used to calculate the net impact of decisions.

4.4.3 Conclusion

Our analysis presents a novel way of quantitatively examining health inequality impacts and demonstrates the potential utility of the DCEA framework in aiding decisions to allocate funding to new treatments in the English NHS. The approach we propose is highly flexible and can be easily applied to any disease type.

However, additional research will greatly improve the validity of the approach. Future work should focus on refining the data sources used to make these simple DCEAs. Appropriate sources could be obtained that more realistically reflect the expected distribution of benefits, such as the Clinical Practice Research Datalink for diseases treated mostly in primary care, or the Mental Health Minimum Dataset for mental health conditions. Empirical work can help inform adjustments that could be made to benefit distributions to account for factors outside utilisation, such as uptake

and adherence. Work can also be done to allow for parameter uncertainty to be included in analyses since many inputs will be highly uncertain. Of particular interest would be work that could validate the seemingly high inequality aversion estimates from Robson et al. (2016), which implicitly weight gains for least healthy between six and seven times more highly the most healthy.

Quantifying the distributional impact of new technologies, despite the importance of health inequality to policy-makers and the general public, has not been undertaken in health technology assessment. This study and the proposed method can help to rectify this omission from the decision-making process.

Appendix 4

Table A4.5: Socioeconomic distribution of an additional quality-adjusted life year (QALY) for each of the disease areas in our sample

	% of each incremental QALY by IMD quintile				
Disease Area (ICD Code)	**IMD 1**	**IMD 2**	**IMD 3**	**IMD 4**	**IMD 5**
Thromboembolism (I82)	25%	22%	21%	17%	16%
Type 2 Diabetes (E11)	24%	22%	22%	18%	14%
Atrial Fibrillation (I48)	21%	20%	20%	21%	19%
Hepatitis C (B18)	32%	24%	23%	12%	10%
Multiple Sclerosis (G35)	20%	21%	20%	20%	19%
Migraine (G43)	24%	22%	21%	18%	16%
Deep vein thrombosis/Pulmonary embolism (I26)	22%	21%	20%	20%	17%
Bone Cancer (C40, C41)	26%	22%	22%	16%	14%
Cystic Fibrosis (E84)	22%	20%	20%	20%	18%
Coronary Heart Disease (I50)	23%	21%	21%	19%	16%
Macular Oedema (H35)	20%	19%	20%	22%	19%
Bipolar I Disorder (F31)	27%	22%	23%	16%	12%
Vitreomacular Traction (H43)	23%	20%	20%	19%	18%
B-cell Lymphoma (C85)	25%	21%	21%	17%	15%
Chronic idiopathic constipation (K59)	23%	21%	21%	18%	16%
Myelodysplastic syndrome (D46)	19%	19%	21%	21%	20%
Alcohol dependence (F10)	29%	22%	22%	15%	13%
Gastrointestinal Stromal Tumours (D37)	23%	20%	20%	19%	18%

Note:

1. Treatments are linked to a disease area by three-digit International Classification of Disease (ICD) code. This is used to extract the distribution of health utilisation by Index of Multiple Deprivation (IMD) quintile group.

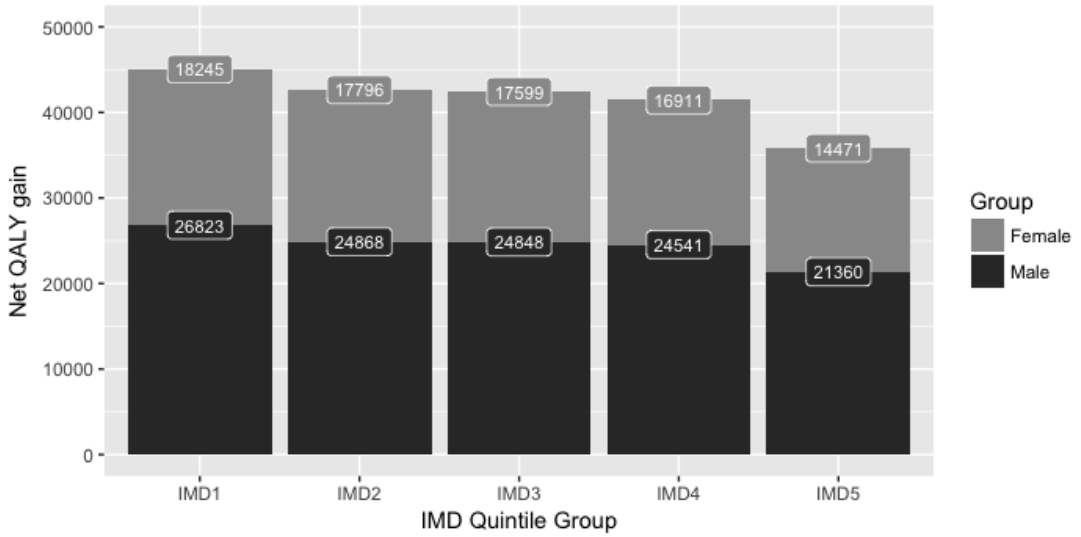
Table A4.6: Output for linear regressions of per person health (and cost) change on baseline health (and cost)

Variable	(1) Δcost	(2) Δhealth
Cost	0.3*	
	(0.136)	
$Cost^2$	-3.31×10^{-6}	
	(3.22×10^{-6})	
QALYs		-0.286***
		(0.0458)
Constant	948.8	-0.011
	(856.7)	(0.0326)
Observations	17	17
Prob > F	0.006	< 0.001
Adj. R-squared	0.451	0.704

Notes:

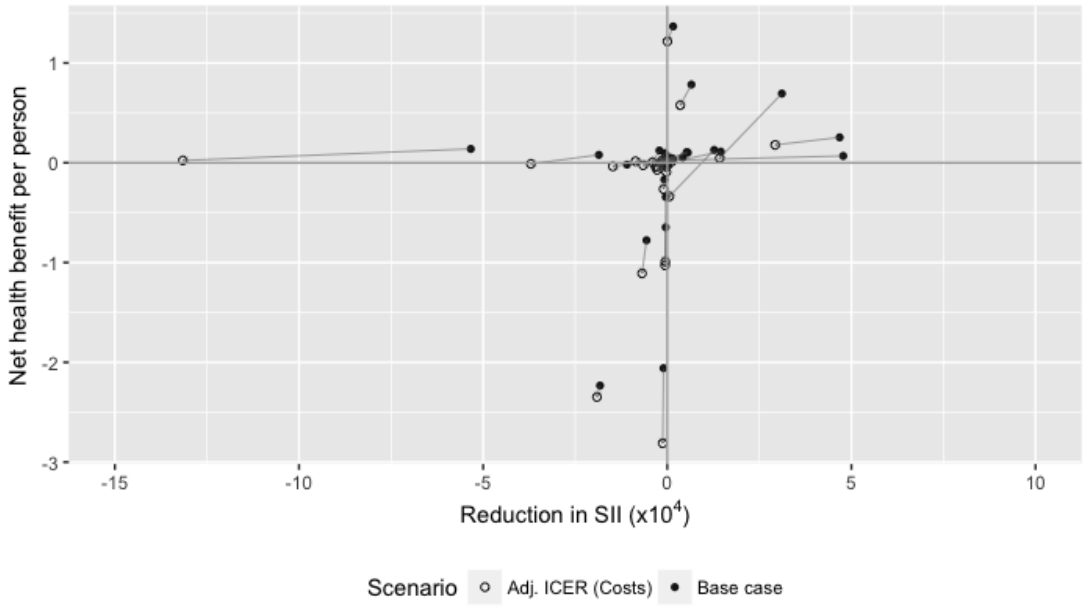
1. Standard errors in parentheses
2. *** p<0.01, ** p<0.05, * p<0.1

Figure A4.6: Cumulative distribution of net quality-adjusted life year gains by Index of Multiple Deprivation quintile groups, broken down by gender



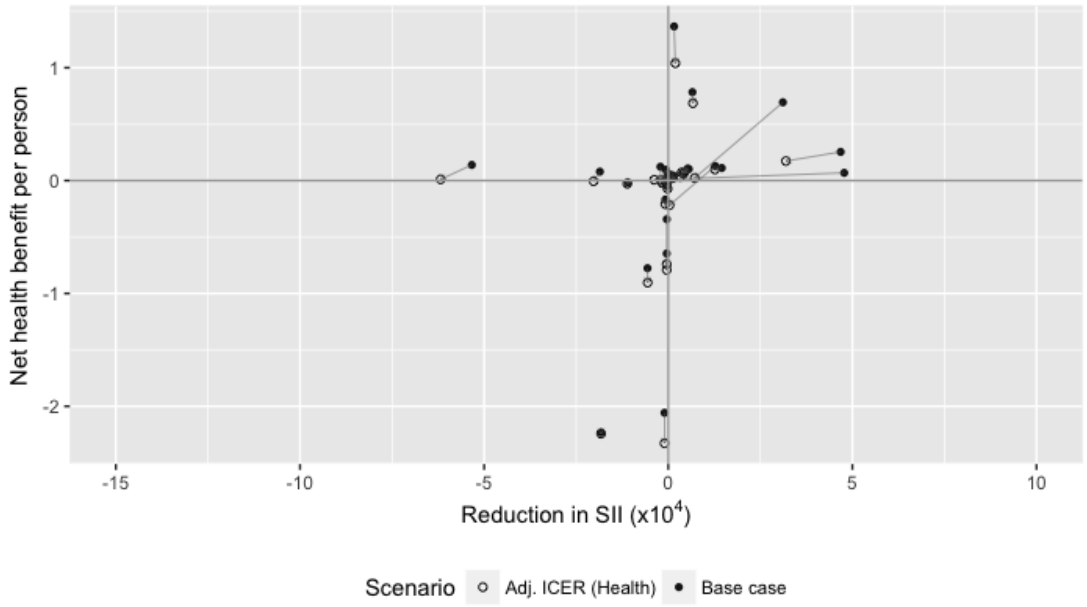
Note: IMD1 = most deprived, IMD5 = least deprived

Figure A4.7: Change in equity impact plane location of each technology when a £6,200 increase in the ICER is attributed to an increase in incremental cost



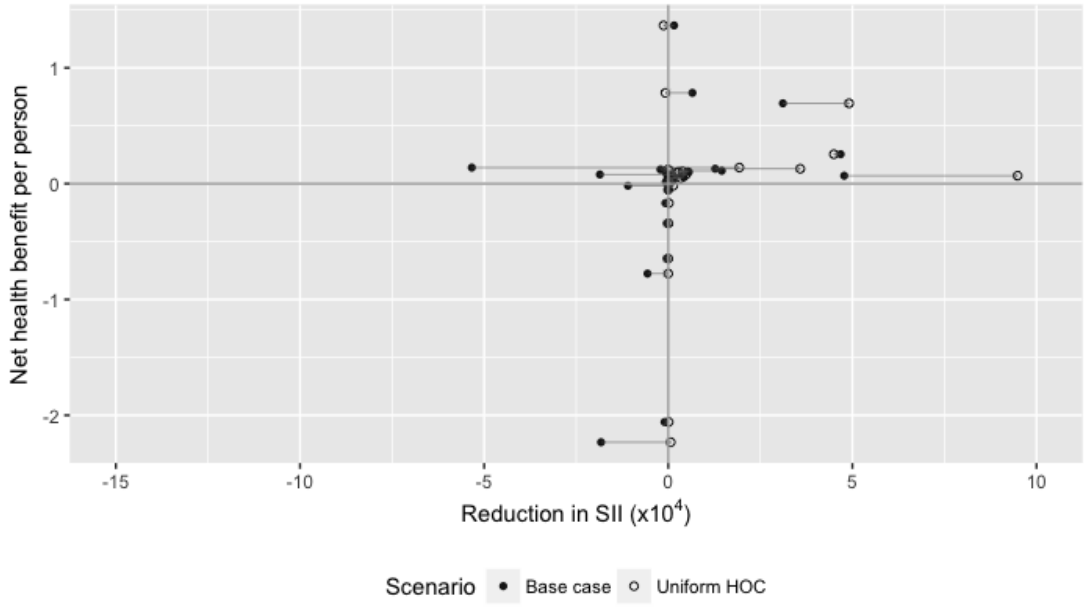
Note: SII = slope index of inequality

Figure A4.8: Change in equity impact plane location of each technology when a £6,200 increase in the ICER is attributed to a reduction in incremental health



Note: SII = slope index of inequality

Figure A4.9: Change in equity impact plane location of each technology when health opportunity costs are distributed uniformly over gender and socioeconomic groups



Note: SII = slope index of inequality; HOC = health opportunity costs

Chapter 5: Distributional cost-effectiveness analysis of smoking cessation interventions: a decision model adaptation

5.1 Introduction

Addressing inequalities in the social determinants of health, such as lifestyle and education, is considered a key tool in the quest to reduce health inequalities (Graham & Kelly 2004; Marmot et al. 2008). Tackling health inequalities has become one of the primary goals of public health campaigns, where interventions, such as those promoting healthy diet and exercise, focus on prevention.

The health inequality impacts of new public health interventions are not routinely estimated. In England, for example, The National Institute for Health and Care Excellence (NICE) produces evidence-based guidance for public health commissioners, with the aim of improving population health and reducing unfair health inequalities (NICE 2014). As in the case of health care interventions, the quantitative component of the guidance centres principally on cost-effectiveness analysis, leaving the consideration of health inequalities as a qualitative exercise for decision-making committees. Yet, health inequality impact analysis is arguably more important in public health than for health care, because interventions can be prone to generate inequalities, such as through differences in uptake (Lorenc et al. 2013).

The objective of this work is to demonstrate how distributional cost-effectiveness (DCEA) can be used to directly model the health inequality impacts of public health interventions by extending decision models used to estimate cost-effectiveness, and can be achieved with limited additional resources by using evidence readily available in the literature. Conducting an in-depth analysis that incorporates numerous sources of socioeconomic variation also provides an opportunity to compare the full DCEA approach (Asaria, Griffin, Cookson, et al. 2015) with the simplified version proposed in chapter 4. This allows us examine the strengths and limitations of that approach when evaluating health inequality impacts.

Although DCEA models can be developed alongside a traditional decision analytic model, in this pilot study we retrospectively adapt an existing model developed for NICE to evaluate the cost-effectiveness of behavioural and pharmacological interventions to improve smoking cessation. Smoking remains a significant cause of ill health and death in England. It was estimated in 2014/15 that 19 per cent of adults in England were smokers, despite substantial public health efforts and increases in taxation. As a result, approximately 4 per cent of hospital admissions and 17 per cent of deaths in England were related to smoking (HSCIC 2016). Smoking prevalence is also strongly correlated with a number of important social variables. Those with low incomes, less qualifications and living in poor neighbourhoods are more likely to smoke, as are men. Smoking, therefore, represents a key determinant of both health inequality and population health, providing an ideal case study with which to pilot the DCEA framework.

5.2 Methods

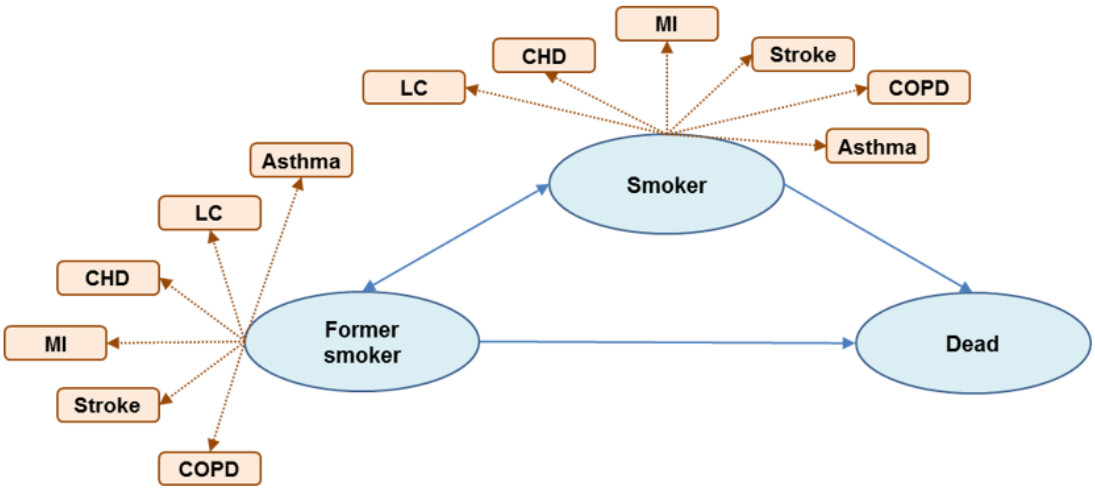
5.2.1 Overview

The DCEA analysis we present consists of four stages. First, we retrospectively adapt the decision model for smoking cessation interventions. The adaptation incorporates the socioeconomic variation in key parameters and does not alter the model structure. This is done through literature searches that identify data on these variations for six sets of model inputs (see Table 5.1). Next, the model is rerun to estimate the incremental costs and health benefits for recipients of each intervention in each of the five socioeconomic groups. After adjusting for the relative uptake across groups, we then calculate costs and health benefits at the population level. This is done because (i) health opportunity costs fall across the whole population and (ii) health inequality is estimated at population level. Using the same process described in chapter 4, we then model these impacts on to the baseline distribution of health and evaluate the change in health inequality using inequality and social welfare measures.

5.2.2 Decision model adaptation

The existing economic model adapted in this study was created to help inform recommendations for an ongoing update of NICE public health guidance on smoking cessation, and is an update of previous NICE guidelines on brief advice, referral and services for smoking cessation (NICE 2006; NICE 2008b). Our analysis adopts the same scope as the original model, and evaluates a wide range of interventions (listed in Table 5.4), including brief advice, behavioural support, pharmacotherapies and e-cigarettes from the perspective of the NHS and personal social services over a lifetime horizon. The principal health outcome is quality-adjusted life years (QALYs), with cost-effectiveness summarised using the incremental cost-effectiveness ratio: the cost-per-QALY gained over the comparator intervention. We determined that this was a suitable case study because smoking prevalence and the success of smoking cessation interventions are known to vary systematically according to socioeconomic status (Hiscock et al. 2015).

Figure 5.1: Model structure for smoking cessation interventions



Note: LC = lung cancer; CHD = coronary heart disease; MI = myocardial infarction; COPD = chronic obstructive pulmonary disease; asthma = asthma exacerbation.

The model adopts a Markov structure, shown in Figure 5.1. A cohort of smokers enters the model, with a proportion transitioning to the 'former smoker' state based

on the effectiveness of an intervention. The effectiveness outcome used in the model is the probability of quit success after one year. The cohort faces mortality and disease risks specific to age and smoking status, with former smokers facing uniformly lower risks and higher health-related quality of life. The model includes six smoking-related comorbidities: lung cancer, coronary heart disease, chronic obstructive pulmonary disease, myocardial infarction, stroke and asthma. Those who receive no intervention have a 'background' quit rate of 2%, which represents the proportion of current smokers who naturally quit each year (West 2006). To account for age-based heterogeneity, the model is run for each year of age from 16 to 100. Results are calculated using a weighted average, in which each weight is the relative density of smokers for the respective year of age.

The model is run separately for five equity-relevant subgroups defined by their socioeconomic status. The socioeconomic variable used is the Index of Multiple Deprivation (IMD), a weighted index measure based on seven dimensions of deprivation: employment, income, education, crime, living environment and housing/services. A unique score is assigned to each of the 32,482 small areas in England, which are then grouped into fifths. Each individual is therefore associated with an IMD quintile group according to the score given to their place of residence.

Subgroup analysis is conducted by varying a number of key model inputs by socioeconomic status that relate to (i) baseline levels of health and (ii) intervention impacts. A pragmatic literature review identified six inputs, shown in Table 5.1.

5.2.2.1 Baseline levels of health

Four aspects of baseline health disaggregated by socioeconomic status are identified in the literature. First, a bespoke data extraction from the ONS from 2013 provided mortality rates specific to each IMD quintile group, shown in Figure A5.7. As these data are only available by 5-year age bands, the rates are converted into one-year probabilities for use in the adapted model using standard techniques (Fleurence & Hollenbeak 2007).

Table 5.1: Summary of model inputs and data sources disaggregated by index of multiple deprivation (IMD) quintile group

Model input	Disaggregation source
Mortality rates by age and gender	Deaths and population estimates by IMD quintile group 2012.
Health-related quality of life by smoking status	Health Survey for England (pooled data from 2012 and 2014)
Proportion of smokers, former smokers and non-smokers within each age and gender group	Health Survey for England (pooled data from 2012 and 2014)
Relative risk of developing smoking-related diseases	Eberth et al. (2013)
Intervention quit rate	Dobbie et al. (2015)
Intervention uptake rate	NHS Stop Smoking Services Statistics (2014-15)

Secondly, EQ-5D estimates disaggregated by smoking status and IMD quintile group are obtained from the most recent Health Survey for England datasets that included the EQ-5D questionnaire (2012 and 2014). These included a collective sample size of 20,413, of which 6,076 were excluded due to missing values for EQ-5D, age, IMD or smoking status, and 6,298 excluded as they were neither current nor former smokers. To avoid double counting the impact of comorbidities included in the model, we excluded a further 2,592 individuals with self-reported respiratory or circulatory conditions, leaving a sample size of 5,447. Linear regression analysis is used to predict EQ-5D scores for smokers and former smokers by IMD quintile, regressing EQ-5D score on age, smoking status and IMD quintile. Interaction terms are initially included, but removed as none are statistically significant. The regression output used to predict the EQ-5D scores used in the economic model is reported in Table A5.8. The resultant health-related quality of life values are shown in Table 5.2.

A third set of inputs expected to vary by socioeconomic status are the baseline prevalence of comorbidities. Although some of the observed variation will be explained by differences in smoking prevalence, evidence from a Scottish study is also found on the independent impact of deprivation level on the probability of smoking-related disease (Eberth et al. 2014). The regression coefficients and baseline characteristics of the study sample are used to estimate the relative risk of a smoking-related event for each Scottish IMD quintile relative to quintile 3, shown in Table 5.2. Assuming that the population prevalence of each smoking related disease in the existing model is equivalent to Scottish IMD quintile 3, our calculated relative risks are used to estimate prevalence for the remaining quintiles.

Table 5.2: Model inputs disaggregated by socioeconomic status

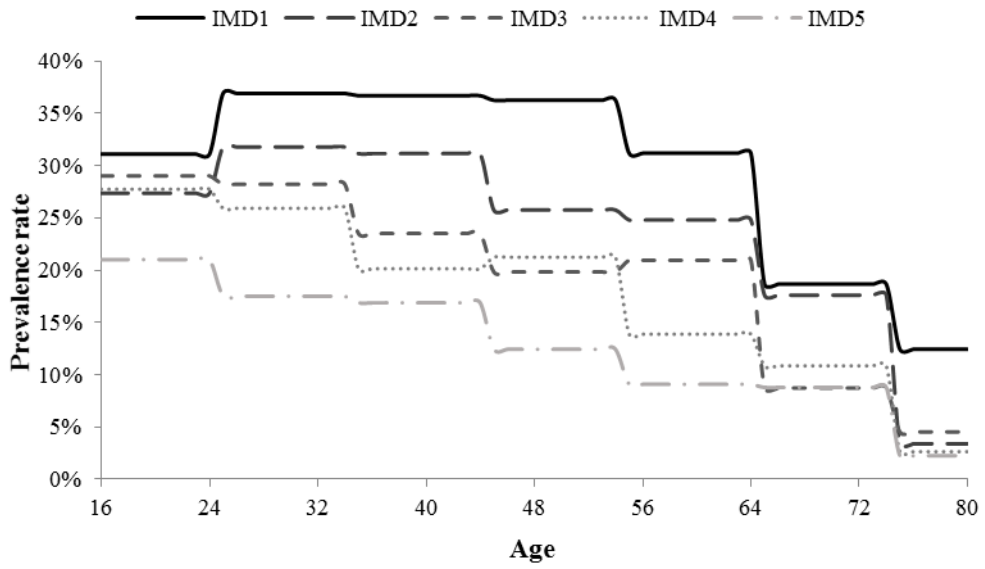
	IMD quintile group				
Model input	**1**	**2**	**3**	**4**	**5**
Smoking population					
Number of smokers	2,531,937	2,021,674	1,679,300	1,346,821	1,062,170
Proportion of male smokers	55%	56%	56%	55%	57%
EQ-5D scores					
Smokers	0.786	0.835	0.830	0.858	0.877
Former smokers	0.803	0.845	0.839	0.863	0.887
Comorbidity risk					
Relative risk of smoking-related illness	1.03	0.99	1.00	0.95	0.84
Odds of quit success					
All intervention types	1.00	1.58	1.34	1.43	1.61
One-to-one	1.00	1.04	1.06	1.08	1.07
Drop-in clinic	1.00	1.05	1.07	1.08	1.07
Open group	1.00	1.27	1.33	1.37	1.35
Closed group	1.00	1.15	1.24	1.44	1.49
Utilisation of NHS stop smoking services					
Any service	4.03%	6.48%	6.62%	10.14%	9.92%
One-to-one	3.39%	5.36%	5.34%	7.24%	8.01%
Drop-in clinic	0.32%	0.43%	0.65%	1.35%	1.14%
Open group	0.03%	0.12%	0.14%	0.54%	0.08%
Closed group	0.14%	0.10%	0.07%	0.13%	0.13%

Notes:

1. IMD = index of multiple deprivation. 1 = most deprived and 5 = least deprived
2. EQ-5D score and smoking prevalence by deprivation level and gender are estimated from the Health Survey for England. Comorbidity relative risk is estimated for Scottish IMD quintiles and applied to the English population. Quit success and utilisation are estimated from NHS stop smoking services data.
3. The list of stop smoking services categories is not exhaustive, thus the sum of the four service types does not equal the 'Any service' percentage.
4. 'All intervention' quit odds are taken from Dobbie et al. (2015), whilst the intervention-specific odds are from Hiscock et al. (2013)

Lastly, we estimate smoking prevalence by age and socioeconomic status using the same pooled Health Survey for England data used in the EQ-5D regression. The resultant rates are shown in Figure 5.2. A number of other parameters are assumed constant over socioeconomic groups as no data are identified to estimate any variation. Notable amongst these are the relative risks of a successful quit attempt on all-cause mortality and smoking-related disease; the disutility associated with each disease; and annual disease costs.

Figure 5.2: Prevalence of smoking by age and index of multiple deprivation (IMD)



5.2.2.2 Socioeconomic variation in intervention impacts

A pragmatic review was undertaken to identify evidence on how quit attempts from specific smoking cessation interventions vary by socioeconomic status. Several publications from the ELONS study investigated whether the quit rates for smoking cessation interventions differ according to socioeconomic status (Hiscock et al. 2013; Hiscock et al. 2015; Dobbie et al. 2015). Of these, one by Dobbie et al. (2015) provided odds ratios for cessation for those using NHS stop smoking services by IMD quintile. These data do not allow us to differentiate between the various types of interventions, and are only available for the probability of quitting at four weeks, rather than for the 52-week quit rates used in the existing cost-effectiveness model. While odd ratios of cessation at 52 weeks are estimated from a separate prospective study, the binary measure of socioeconomic status used in this analysis could not be mapped to IMD quintiles and are not used. We therefore assume that the socioeconomic pattern observed for four week quit rate is reflective of the pattern in 52-week quit rate.

We sought additional evidence that could inform how quit success varies by intervention type as well as socioeconomic status. However, we are only able to identify one study that provided data on these patterns (Hiscock et al. 2013).

Although this study, which used the same English NHS stop smoking service data as Dobbie et al., explores interactions between cessation and socioeconomic status, it does so only for selected behavioural intervention types. Their results in principle allow us to characterise up to four different socioeconomic patterns in quit rates for behavioural interventions. However, only 11 of the 21 interventions included in the model mapped to these intervention types (nine 'one-to-one' and two 'closed group' interventions). Furthermore, the measure of socioeconomic status is an NS-SEC occupation-based measure, thereby requiring the odds ratios to be mapped to IMD using a cross-tabulation of NS-SEC and IMD quintile group estimated from the Health Survey for England. Our base case analysis therefore applies the same socioeconomic pattern in treatment efficacy across all intervention types estimated by Dobbie *et al*., with the results of Hiscock *et al*. adopted in a sensitivity analysis. Both sets of odds ratios are shown in Table 5.2.

The socioeconomic distribution of uptake was another expected critical determinant of net benefit (Brown et al. 2014; Bell et al. 2007). To estimate this, we used NHS stop smoking services statistics from 2014/2015, which supply information on the number of interventions provided within each local authority (HSCIC 2014; HSCIC 2016). By mapping local authorities to an IMD quintile group (using the average score of the small areas that comprise them) and summing the number of provided interventions, we calculated the rate by socioeconomic group. However, rather than uptake, these data are more akin to utilisation because the availability of services varies between local authorities. These utilisation rates are summarised in Table 5.2 and show that a greater proportion of smokers in the least deprived quintile groups are utilising services. However, in terms of the absolute numbers of smokers utilising services, a slight pro-poor gradient is observed due to smoking prevalence being much greater in more deprived groups. As a sensitivity analysis, we re-estimate our results using a uniform uptake rate of 6.8% for all quintiles, this being the average value for the smoking population. These results indicate the potential inequality impact of equalising the utilisation of services across socioeconomic groups.

5.2.3 Modelling health inequality impact

5.2.3.1 Health benefits

The adapted cost-effectiveness model is run for all 21 interventions and each of the five IMD groups. The model estimates the incremental costs and QALYs for an individual receiving each intervention, weighted by gender. We translate these per recipient estimates into population level costs and QALYs by multiplying the estimates for each IMD quintile by its respective uptake rate and smoking population. The QALY benefits of smoking cessation interventions in the existing model are specific to smokers, and are a weighted gender average. We disaggregate these QALY benefits by gender using the ratio of male to female smokers for each IMD quintile group, shown in Table 5.2.

5.2.3.2 Opportunity cost of displaced activities

The additional costs of implementing an intervention represent health opportunity costs; which from an NHS perspective equates to the health losses that are associated with withdrawing funding from existing services to pay for the new one. The population costs we estimate for each intervention are therefore converted into health opportunity costs at a rate of £20,000 per QALY. This is the lower bound of the cost-effectiveness threshold range that NICE uses to decide whether an intervention is cost-effective. As well as health opportunity costs, NICE states that the threshold factors in criteria such as severity of illness and innovative value (Dillon 2015). The health opportunity costs can fall to any individual in the population regardless of smoking status, as changes in NHS spend affect the level of general NHS activities, which are not solely for smokers. We therefore distribute the health opportunity costs to gender and socioeconomic groups using the results from Chapter 3, shown in Table 5.3.

The results indicate that new interventions that impose additional costs on the NHS will increase inequality and interventions that save costs will reduce inequality. This distribution of opportunity cost is directly relevant to any changes in health care resource use as a consequence of smoking cessation interventions (e.g. cost savings from avoided smoking-related disease) but may not match the opportunity cost of

other public sector funding sources used to provide smoking cessation services (e.g. intervention costs funded by local authorities). However, changes in NHS and health sector resource use form the biggest component of cost changes attributed to smoking cessation, with intervention implementation costs smaller by comparison. In the absence of an estimate of the distribution of the health impact of displaced wider public sector spending, we therefore assume the same gradient as estimated for NHS services.

Table 5.3: Estimate of the gender and socioeconomic distribution of health opportunity costs in the English NHS.

Gender	IMD1	IMD2	IMD3	IMD4	IMD5
Males	0.14	0.12	0.12	0.09	0.08
Females	0.12	0.10	0.10	0.07	0.06

Note: IMD = Index of multiple deprivation, an area-level measure of socioeconomic status. IMD1 = most deprived, IMD5 = least deprived

5.2.3.3 Net health inequality impact

Taking the difference between the QALY gains and the health opportunity cost provides the net health impact by gender and IMD for each intervention. We add these net QALY impacts on to a baseline distribution of health estimated in chapter 2. The subsequent distribution of health provides a picture of health inequality following the implementation of the intervention or guideline.

For example, if an intervention cost £100 million and generated 5000 QALYs for males in the least deprived quintile group, their health opportunity costs would be:

$$(1{,}000{,}000 \times 0.14)/20{,}000 = 700$$

The net health benefit for this group is therefore 4300 QALYs. With a baseline quality adjusted life expectancy of 62.3 per person and a population size of 5,393,565, we can therefore work out the post-intervention quality-adjusted life expectancy for this subgroup:

$$((62.3 \times 5{,}393{,}565) + 4300)/5{,}393{,}565 = 62.3008$$

Giving an increase of 0.008 lifetime QALYs per person in this subgroup. Repeating this for each sex and IMD group will yield a new quality adjusted life expectancy for every subgroup defined by gender and IMD quintile in the population. Ordering the whole population from least to most healthy then provides a univariate distribution of health that can be compared against the baseline distribution.

5.2.3.4 Inequality impact measures

The analysis of the pre- and post-intervention health distributions follows the same procedure as described in section 4.2.3.3 chapter 4. The slope index of inequality (SII) and the relative index of inequality (RII) are calculated to measure absolute and relative inequality, respectively, whilst the Kolm and Atkinson indices are estimated to measure social welfare.

The slope index of inequality is calculated as the slope coefficient estimated from a linear regression of QALE on population health quintile that uses the following specification:

$$Q_r = \delta + \beta_{SII} r + \gamma_r$$

Where Q_r is the QALE estimate of population quintile r, β_{SII} is the slope index of inequality value, γ_r is the idiosyncratic error and δ is the constant term. β_{SII} is interpreted as the fitted difference in QALE between the least to most healthy population quintile.

The relative index of inequality is interpreted as the relative change in QALE when moving from the least to most healthy population quintile. RII is obtained by dividing the slope index of inequality by the mean QALE in the population, $\bar{Q}$, such that $\beta_{RII} = \beta_{SII}/\bar{Q}$. The inequality impact of interventions is calculated by taking the difference between SII and RII values pre- and post-intervention. Reduction in SII and RII is reported; a positive value means that they are lower post-intervention and health inequality has reduced.

The net health benefit per person (for the average smoker) and change in SII are plotted simultaneously on the health equity impact plane, as shown in Figure 4.2. Interventions that have a positive incremental net health benefit increase total health and fall in the north of the plane. Interventions that reduce inequality fall in the east of the plane.

Health-related social welfare is summarised by absolute and relative social welfare indices. These include an inequality aversion parameter, which indicates the amount by which the least healthy should be prioritised for health improvement. We use the Atkinson and Kolm indices to measure social welfare changes solely as a function of changes in the population distribution of health. The Atkinson index, A_ε, measures inequality relatively and is given by:

$$A_\varepsilon = 1 - \left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{Q_i}{\bar{Q}}\right)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}$$

Where N is the total population, Q_i is the QALE estimate of the ith individual, $\bar{Q}$ is the mean QALE and ε the inequality aversion parameter that quantifies the concern for relative inequality. Alternatively, the Kolm index, K_α, incorporates inequality on an absolute scale, where absolute inequality aversion is represented by the parameter α:

$$K_\alpha = \left(\frac{1}{\alpha}\right)\log\left(\frac{1}{N}\sum_{i=1}^{N} e^{\alpha[\bar{Q}-Q_i]}\right)$$

Increases in the inequality aversion parameter reflect a greater societal concern for the less healthy. A base case value of 10.95 for ε and 0.15 for α are again used, based on the results from Robson et al. (2016). Social welfare is calculated by combining each index with the mean level of health in the distribution to obtain the 'equally distributed equivalent' (EDE) level of health:

$$EDE_{A,\varepsilon} = (1 - A_\varepsilon)\bar{Q}$$
$$EDE_{K,\alpha} = (\bar{Q} - K_\alpha)$$

Where $EDE_{A,\varepsilon}$ and $EDE_{K,\alpha}$ are the Atkinson and Kolm welfare scores, respectively. The equally distributed equivalent is the level of average health (expressed in QALYs) in a completely equal distribution that yields an equivalent amount of social welfare to the distribution being evaluated.

We calculate the population-level $EDE_{A,\varepsilon}$ and $EDE_{K,\alpha}$ pre- and post-intervention by multiplying the change by the population size. A positive difference indicates that health-related social welfare has increased. The difference between the equally distributed equivalent and the net population health impact directly describes the extent of the change in health inequality and the strength of inequality aversion in terms of the trade-off between total population health and health inequality reduction. Conversely, interventions that increase health inequality would have an equally distributed equivalent lower than their net population health impact, with the difference showing the loss of social welfare in terms of QALYs.

5.2.3.5 Sensitivity analysis

Four key inputs are varied in sensitivity analysis. The first are the relative odds of successful cessation by IMD group, for which we use the set of intervention-specific quit odds provided in Hiscock et al. (2013), described in Section 5.2.2.2. To do this, we map interventions to the behavioural support categories (one-to-one support, closed group, open group and drop-in clinic) where possible and rerun the analysis using the odds ratios by IMD given in Table 5.2.

Second, we assume that, instead of the utilisation rate being substantially higher amongst the least deprived groups, it is uniform across gender and socioeconomic groups at the population average of 6.8%. Third, we look at the effect of the distribution of health opportunity costs on our results by reanalysing inequality and social welfare impacts. Again, this is done by assuming that the distribution is uniform over gender and socioeconomic status. Last, the value of the cost-

effectiveness threshold is varied to show how inequality impact is affected by the marginal productivity of competing health (and other local authority) services.

Our results can also be used to compare different DCEA modelling approaches. To do this, we produce two further sets of results. We first conduct the same type of simplified DCEA that was outlined in chapter 4, which calculates population health benefits and costs using (i) the whole population of smokers and (ii) the incremental costs and QALYs of the average participant and apportions benefits using the socioeconomic and gender distribution of smokers. As the patient populations used in chapter 4 are for those receiving treatment (extracted from the costing templates), we adjust the population health benefits using the population utilisation rate of 6.8% to account for intervention uptake. Comparing these results with our base case analysis measures the extent by which health inequality reductions are overestimated when additional socioeconomic variation is not accounted for. We next look at how incorporating differential uptake by socioeconomic analysis can be combined and incorporated into the simplified approach, adjusting the health benefits using the IMD-specific utilisation rates reported in Table 5.2 rather than a uniform rate.

To demonstrate the impact of not accounting for uptake at all, we conduct an additional simplified analysis that makes the highly optimistic assumption that all smokers receive the intervention.

5.3 Results

5.3.1 Descriptive statistics

The characteristics of the 21 interventions are reported in Table 5.4 (J. Brown et al. 2014; Chengappa et al. 2014; Rigotti et al. 2010; Jorenby et al. 2006; Heydari et al. 2012; Williams et al. 2006; Issa et al. 2013; Tranvåg et al. 2013; Wittchen et al. 2011; Blondal et al. 1999; Smith et al. 2009; Caponnetto et al. 2013). The net health benefits in Table 5.4 are for each recipient of the intervention, weighted across IMD groups according to the number of smokers. The quit success rate at 12-months is estimated to range from 7% for counselling to 40% for a combination of varenicline, bupropion and selective serotonin reuptake inhibitors (SSRI). Excluding over the

counter nicotine replacement therapy (for which there are only private costs), the intervention costs range from £19 for brief advice to £764 for nicotine patches and nasal spray. All interventions are estimated to have positive net health benefits, indicating that they would be associated with incremental cost-effectiveness ratios of less than £20,000 per QALY gained. Twenty out of the 21 save costs when we account for the fact that interventions avert lifetime health care costs by reducing the risks of smoking-related diseases, but patch and nasal spray is estimated to increase costs. The most and least effective interventions are also the most and least cost-effective; the combination of varenicline, bupropion and SSRI has an incremental net health benefit of 0.446 per recipient, the respective figure for counselling is 0.045.

Table 5.4: Intervention characteristics for the general smoking population

Intervention	Study	12-month quit rate	Cost	NHB	Abbreviation
NRT OTC	Brown et al. (2014)	10%	£0	0.080	NTC OTC
Placebo + counselling	Chengappa et al. (2014)	7%	£29	0.045	Co (1)
Varenicline	Chengappa et al. (2014)	19%	£220	0.164	Var
Placebo + counselling	Rigotti et al. (2010)	18%	£343	0.150	Co (2)
Varenicline + counselling	Rigotti et al. (2010)	32%	£507	0.286	Var, Co (1)
Placebo + counselling	Jorenby et al. (2006)	20%	£189	0.174	Co (3)
Varenicline + counselling	Jorenby et al. (2006)	34%	£353	0.325	Var, Co (2)
Brief advice	Heydari et al. (2011)	8%	£19	0.054	Brief Advice
Varenicline + brief advice	Heydari et al. (2011)	29%	£194	0.266	Var, Brief Advice
Self-determination intervention	Williams et al. (2006)	12%	£199	0.087	SDI
Sequence (var, bup, SSRI)	Issa et al. (2013)	45%	£269	0.446	Var, Bup
Minimal intervention	Wittchen et al. (2010)	34%	£43	0.329	MI
CBT + MI	Wittchen et al. (2010)	33%	£268	0.311	CBT, MI
Bupropion + CBT + MI	Wittchen et al. (2010)	24%	£352	0.209	Bup, CBT, MI
NRT + CBT + MI	Wittchen et al. (2010)	34%	£116	0.326	NRT, CBT, MI
Patch and nasal spray	Blondal et al. (1999)	31%	£764	0.262	Patch, nasal spray
Patch	Blondal et al. (1999)	13%	£120	0.102	Patch
Bupropion and lozenge	Smith et al. (2009)	29%	£79	0.279	Bup, Loz
Lozenge	Smith et al. (2009)	17%	£78	0.145	Loz
7.2mg e-cigarette	Caponetto et al. (2013)	15%	£42	0.130	E-cig
7.2mg then 5.4mg e-cig	Caponetto et al. (2013)	11%	£42	0.082	E-cig (2)

Note: NRT = nicotine replacement therapy; OTC = over the counter; Var = varenicline; Bup = bupropion; SSRI = selective serotonin reuptake inhibitors; MI = minimal intervention; CBT = cognitive behavioural therapy

Table 5.5: Population quality-adjusted life year (row 1) and cost (row 2) impacts by intervention and index of multiple deprivation quintile (IMD) group

	IMD Quintile Group				
Intervention	**1**	**2**	**3**	**4**	**5**
NTC OTC	4,813	9,330	6,567	8,526	7,820
	-£12,035,533	*-£26,989,667*	*-£23,203,684*	*-£30,891,515*	*-£25,330,582*
Co (1)	2,809	5,511	3,860	5,020	4,621
	-£4,070,424	*-£12,147,848*	*-£10,421,129*	*-£14,239,515*	*-£11,918,903*
Var	10,798	20,218	14,428	18,634	16,916
	-£4,539,964	*-£29,637,393*	*-£26,529,953*	*-£37,466,621*	*-£31,601,871*
Co (2)	10,256	19,261	13,729	17,739	16,117
	£9,414,201	*-£10,694,829*	*-£10,344,796*	*-£17,368,459*	*-£16,010,019*
Var, Co (1)	19,738	35,161	25,585	32,797	29,350
	£2,425,484	*-£35,216,898*	*-£34,035,168*	*-£49,558,358*	*-£41,608,735*
Co (3)	11,329	21,147	15,110	19,505	17,690
	-£9,002,030	*-£36,349,813*	*-£32,350,339*	*-£44,814,483*	*-£37,348,760*
Var, Co (2)	21,872	38,516	28,151	36,024	32,134
	-£18,643,989	*-£65,122,677*	*-£60,228,971*	*-£82,297,207*	*-£66,873,062*
Brief Advice	3,314	6,481	4,545	5,909	5,434
	-£6,335,375	*-£16,243,693*	*-£13,936,342*	*-£18,802,375*	*-£15,589,117*
Var, Brief Advice	17,386	31,372	22,715	29,174	26,203
	-£23,693,403	*-£65,341,032*	*-£58,730,047*	*-£79,239,421*	*-£64,454,312*
SDI	5,887	11,341	8,002	10,379	9,502
	£5,577,546	*-£6,737,468*	*-£6,175,773*	*-£10,446,062*	*-£9,818,743*
Var, Bup	30,214	50,928	37,841	48,115	42,408
	-£48,079,046	*-£112,027,260*	*-£103,804,361*	*-£137,572,991*	*-£109,003,021*
MI	21,132	37,362	27,265	34,912	31,177
	-£48,417,236	*-£102,382,623*	*-£91,528,010*	*-£120,570,495*	*-£96,420,748*
CBT, MI	20,639	36,587	26,673	34,166	30,534
	-£24,210,428	*-£70,647,681*	*-£64,423,567*	*-£87,137,476*	*-£70,618,340*
Bup, CBT, MI	14,129	25,960	18,664	24,036	21,700
	£597,162	*-£28,956,482*	*-£26,838,462*	*-£39,022,080*	*-£33,199,645*
NRT, CBT, MI	21,132	37,362	27,265	34,912	31,177
	-£41,032,487	*-£92,901,245*	*-£83,488,530*	*-£110,691,157*	*-£88,795,937*
Patch, nasal spray	19,005	33,991	24,695	31,675	28,378
	£30,448,875	*£1,794,680*	*-£2,378,809*	*-£10,453,775*	*-£11,419,619*
Patch	6,554	12,575	8,887	11,520	10,534
	-£4,174,306	*-£20,691,179*	*-£18,105,141*	*-£25,400,766*	*-£21,511,453*
Bup, Loz	17,871	32,161	23,311	29,927	26,859
	-£36,640,833	*-£82,692,093*	*-£73,610,252*	*-£97,664,368*	*-£78,694,552*
Loz	9,098	17,201	12,228	15,816	14,398
	-£14,813,840	*-£39,561,941*	*-£34,568,457*	*-£46,687,779*	*-£38,447,668*
E-cig	8,051	15,313	10,860	14,059	12,822
	-£15,796,282	*-£38,726,846*	*-£33,654,720*	*-£45,138,536*	*-£37,059,126*
E-cig (2)	5,071	9,814	6,911	8,971	8,225
	-£8,342,852	*-£22,820,270*	*-£19,702,784*	*-£26,702,723*	*-£22,166,258*

Notes:

1. NRT = nicotine replacement therapy; OTC = over the counter; Var = varenicline; Bup = bupropion; SSRI = selective serotonin reuptake inhibitors; SDI = self-determination intervention; MI = minimal intervention; CBT = cognitive behavioural therapy
2. Full intervention names and associated studies are provided in Table 5.4
3. IMD1 = most deprived; IMD5 = least deprived

The population health costs and health effects for each intervention by IMD quintile group are presented in Table 5.5. As these account for smoking population size and service utilisation rate, a pro-rich socioeconomic gradient in health benefits is observed. For the most deprived group, five interventions impose additional sector costs; the ICERs for these groups are all under £2,000 per QALY and lie well below the £20,000 cost-effectiveness threshold used by NICE.

5.3.2 Health inequality impact

The four alternative measures summarising the impacts of the interventions on population health and health inequality are shown in Table 5.6. These evaluate the change in the distribution of health for the general population in England, which includes non-smokers, former smokers and current smokers (as health opportunity costs also fall on non-smokers). In the base case analysis, all interventions marginally increase absolute health inequality across the whole population in England, as indicated by negative values for the slope index of inequality. The picture was more mixed in terms of relative inequality, for which 13/21 (62%) generated a reduction.

Social welfare effects are also reported in Table 5.6. All interventions increased relative and absolute social welfare. The amount by which the net health benefit exceeds the equally distributed equivalent shows the welfare loss of the health inequality increases in terms of population QALYs. For example, the varenicline, bupropion and SSRI intervention was estimated to add 241,659 QALYs to the population, but when the associated health inequality increase is accounted for, this is equivalent in value to 232,705 equally distributed QALYs (when using the Atkinson index). Thus an increase in SII of 0.0005 is therefore equivalent in value to a loss of 8,953 QALYs at the given level of health inequality aversion.

The reduction in the slope index of inequality is plotted against population net health benefit on the equity impact plane in Figure 5.3. The interventions all lie in the northwest quadrant of the plane, indicating that all interventions increase both population health and health inequality compared to no smoking cessation service.

The interventions would have the same rank order based on net health benefit alone or the change in social welfare as measured by the equally distributed equivalent health. That is, the intervention that is estimated to produce the biggest gain in total population health is also estimated to produce the biggest improvement in social welfare.

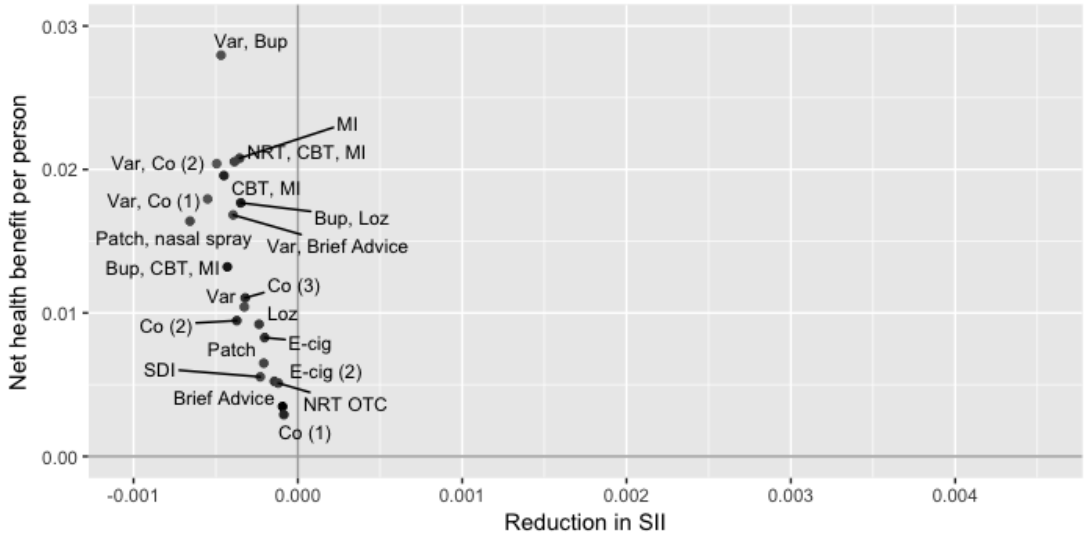
Table 5.6: Change in inequality indices from implementing each intervention

		Inequality reduction		**Social welfare impact**	
Intervention	**Population NHB**	**ΔSII**	**ΔRII**	**$\Delta EDE_{K,\alpha}$**	**$\Delta EDE_{A,\varepsilon}$**
NTC OTC	44,149	-0.00012	0.0000004	41,318	41,875
Co (1)	25,149	-0.00009	0.0000000	23,341	23,638
Var	90,042	-0.00033	-0.0000005	83,491	84,557
Co (2)	81,788	-0.00037	-0.0000016	75,039	75,926
Var, Co (1)	155,039	-0.00055	-0.0000007	144,171	146,072
Co (3)	95,453	-0.00032	-0.0000001	88,781	89,939
Var, Co (2)	176,310	-0.00049	0.0000011	165,381	167,693
Brief Advice	30,040	-0.00009	0.0000001	27,976	28,340
Var, Brief Advice	145,432	-0.00039	0.0000011	136,428	138,325
SDI	47,916	-0.00023	-0.0000011	43,804	44,303
Var, Bup	241,659	-0.00047	0.0000046	229,245	232,705
MI	179,616	-0.00035	0.0000033	169,993	172,499
CBT, MI	169,149	-0.00045	0.0000014	158,866	161,101
Bup, CBT, MI	114,162	-0.00043	-0.0000009	105,768	107,118
NRT, CBT, MI	177,496	-0.00039	0.0000028	167,616	170,054
Patch, nasal spray	141,699	-0.00066	-0.0000029	130,142	131,714
Patch	56,145	-0.00021	-0.0000004	51,970	52,622
Bup, Loz	152,707	-0.00035	0.0000022	143,951	146,015
Loz	79,617	-0.00024	0.0000003	74,344	75,338
E-cig	71,553	-0.00020	0.0000005	66,906	67,807
E-cig (2)	45,210	-0.00014	0.0000001	42,089	42,637

Notes:

1. ΔSII = change in slope index of inequality; ΔRII = change in relative index of inequality; EDE = equally distributed equivalent; $\Delta EDE_{K,\alpha}$ = change in Kolm Index EDE; $\Delta EDE_{A,\varepsilon}$ = change in Atkinson Index EDE; NHB = population net health benefit
2. Full intervention names and associated studies are provided in Table 5.4
3. SII and RII are reported as reduction, meaning that a positive value indicates a health inequality reduction Kolm and Atkinson Indices used inequality aversion parameters of 0.15 and 10.95, respectively

Figure 5.3: Equity impact plane showing the net health and health inequality impact of each intervention

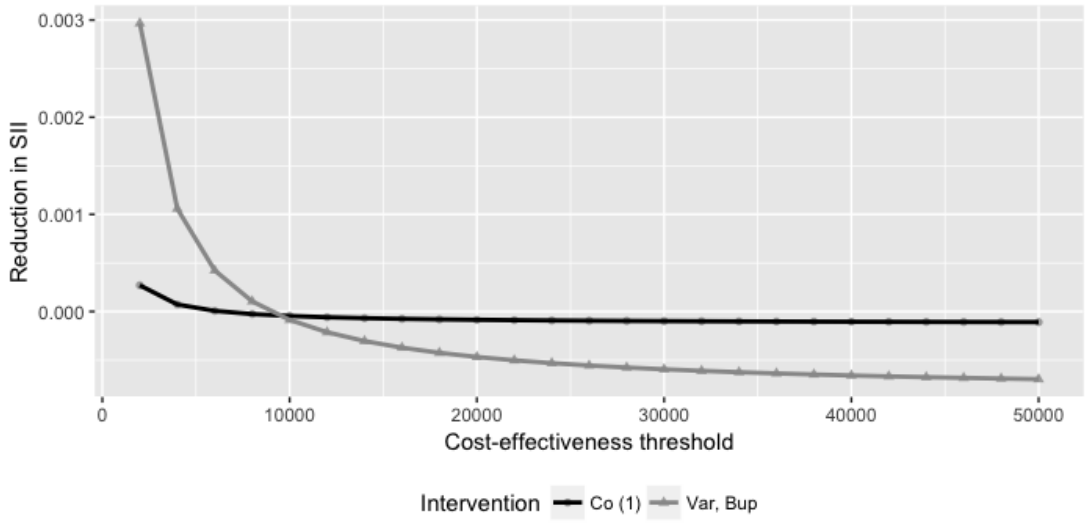


Note: Intervention abbreviations are provided in Table 5.4

5.3.3 Sensitivity analysis

Figure 5.4 shows, for the most and least cost-effective interventions, the effect of adjusting the cost-effectiveness threshold on the health inequality impact. When the threshold is £2,000 per QALY (which indicates that health opportunity costs are very high), we see an SII reduction of 0.003 from funding the varenicline, bupropion and SSRI intervention. As the cost-effectiveness threshold increases, SII reductions become smaller and become SII increases at a threshold of approximately £9,000. At £50,000 per QALY, the intervention causes an SII increase of 0.0007. For the counselling intervention, a similar trend is observed at a greatly reduced magnitude. The SII ranges from a reduction of 0.0003 to an increase of 0.0001 as the threshold goes from £2,000 to £50,000.

Figure 5.4: Impact of the cost-effectiveness threshold on the inequality impact of the most and least cost-effective interventions



Note: "Co (1)" = placebo + counselling; "Var, Bup" = varenicline and bupropion

Table 5.7: Reduction in slope index of inequality (SII) when using the quit odds ratios by index of multiple deprivation quintile group from (1) Dobbie et al. (2015) or (2) the intervention-specific odds Hiscock et al. (2013)

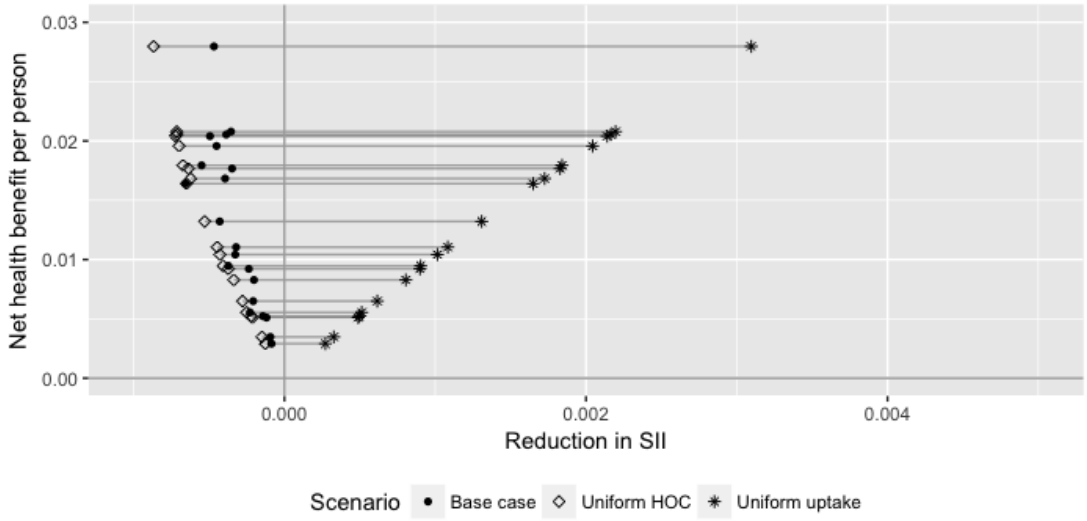
Intervention	Study	Intervention type	ΔSII (1)	ΔSII (2)	Difference
Placebo + counselling	Chengappa et al.	One-to-one	-0.00009	0.00004	0.00012
Placebo + counselling	Rigotti et al.	One-to-one	-0.00037	-0.00007	0.00031
Varenicline + counselling	Rigotti et al.	One-to-one	-0.00055	-0.00008	0.00047
Placebo + counselling	Jorenby et al.	One-to-one	-0.00032	0.00001	0.00033
Varenicline + counselling	Jorenby et al.	One-to-one	-0.00049	0.00000	0.00050
Brief advice	Heydari et al.	Closed group	-0.00009	-0.00018	-0.00009
Varenicline + brief advice	Heydari et al.	Closed group	-0.00039	-0.00063	-0.00024
Minimal intervention	Wittchen et al.	One-to-one	-0.00035	0.00013	0.00049
CBT + MI	Wittchen et al.	One-to-one	-0.00045	0.00003	0.00048
Bupropion + CBT + MI	Wittchen et al.	One-to-one	-0.00043	-0.00005	0.00038
NRT + CBT + MI	Wittchen et al.	One-to-one	-0.00039	0.00010	0.00049

Note:

4. NRT = nicotine replacement therapy; Var = varenicline; MI = minimal intervention; CBT = cognitive behavioural therapy
5. The quit odds informing both sets of estimates are provided in Table 5.2

The effect of using the intervention-specific odds of quitting by socioeconomic status on the expected reduction in health inequality is summarised in Table 5.7. Of the 11 interventions that could be mapped to the behavioural support categories, a majority (82%) are one-to-one interventions. For these, the discrepancy between IMD quintiles in terms of successful quit attempts is lower and they were estimated to be more successful in the most deprived groups compared to the base case. This has the effect of moving six interventions from health inequality increasing to reducing. For example, the SII reduction for a 'minimal intervention' changed from -0.0004 to 0.0001. The remaining two interventions are mapped to the closed group type. These are marginally less successful in the most deprived groups compared to our base case, which translated to a very small decrease in inequality reduction.

Figure 5.5: Effect on equity impact position of interventions of applying (i) uniform health opportunity costs over gender and socioeconomic groups and (ii) uniform uptake rates over socioeconomic groups

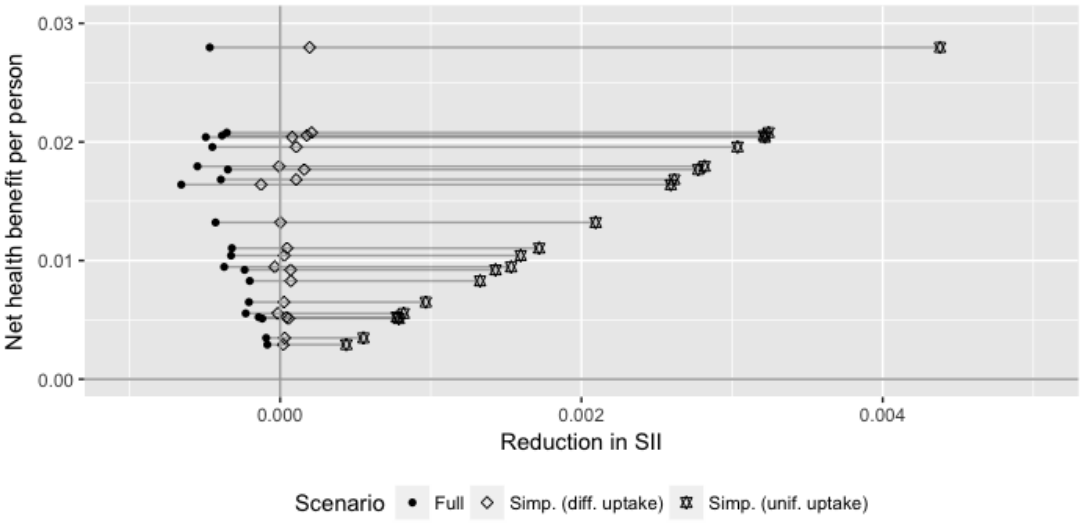


Note: HOC = health opportunity costs; SII = slope index of inequality

The sensitivity of an intervention's position on the health equity impact plane to the distribution of health opportunity cost is shown in Figure 5.5. If the health opportunity cost is equally distributed across all socioeconomic groups, each intervention shifts to the left, indicating a lesser reduction in health inequality. This effect is larger as per person net health benefit increases. All interventions remain in the northwest quadrant of the health equity impact plane. The effect of utilisation rate

on equity impact plane position is also shown in Figure 5.5, assuming an equal probability of uptake per eligible smoker of 6.8%. The interventions all move from the northwest to the northeast quadrant, with the extent of the shift again larger as net health benefit increases.

Figure 5.6: Equity impact position of smoking cessation interventions when conducting full and simplified DCEAs



Notes:

1. Full = base case analysis; Simp. = simplified analysis; Simp. (diff. uptake) = simplified analysis with socioeconomic variation in uptake; Simp. (unif. uptake) = simplified analysis with uniform uptake over socioeconomic status
2. SII = slope index of inequality

Figure 5.6 compares the results of our base case with those yielded from the simplified approach of chapter 4. When we assume uniform uptake over socioeconomic groups, all interventions lie in the northeast quadrant, improving population health and reducing inequality. The inequality reductions are greater than when uniform uptake is applied in the full DCEA analysis. When differential uptake is incorporated, the results are much closer to those of the base case, although only four (19%) move to the northwest quadrant. The direction of the health inequality impact is therefore incorrectly estimated for a vast majority of interventions when using the simplified approach, even when additional socioeconomic variation is accounted for. When no data on uptake is incorporated, the health inequality impacts

are substantially biased. The change in SII for the varenicline, bupropion and SSRI intervention, for example, goes from an increase of 0.004 to a reduction of 0.06. These differences are shown in Figure A5.9.

5.4 Discussion

5.4.1 Main findings

Our results show that smoking cessation interventions are expected to increase both population health and health inequalities, despite the much higher proportion of the smoking population coming from lower socioeconomic groups. This is driven by the lower expected net health benefits per smoker in more deprived groups, which are in turn driven by the socioeconomic variation captured in the adapted decision model: individuals with lower socioeconomic status have a lower probability of quit success, a lower probability of utilising cessation services, higher mortality and lower health-related quality of life.

These factors also compensated for the fact that all interventions are cost-saving, meaning that, in theory, resources would be freed up to be allocated to additional health services. As the results from chapter 3 indicate that additional funding at the margin would benefit the most disadvantaged more, these cost savings help to reduce health inequality in the population. This effect is demonstrated when we assume that the distribution of health opportunity costs is assumed to be equal over IMD groups, which generates greater inequality increases for all interventions.

A crucial factor in determining the magnitude and direction of the results is the differential probability of uptake of interventions per smoker by socioeconomic group. Our base case analysis bases probability of uptake by IMD quintile group on the proportion of eligible smokers utilising NHS stop smoking services, which indicate that smokers in the least deprived areas are more than twice as likely to use interventions than those in the most deprived areas (10% versus 4%, respectively). Assuming that the probability of uptake is equal for all smokers at the national average of 6.8% changes all interventions from inequality increasing to reducing. Our findings therefore demonstrate the value of eliminating socioeconomic variation

in service uptake, and offer further support for the targeting of interventions to more deprived areas, even though on a per recipient basis they are the least cost-effective.

An expected yet important methodological finding is that the simplified approach to DCEA outlined in chapter 4 overestimates the health inequality reductions of an intervention and, in this instance, does not estimate the same direction of effect. Accounting for differential uptake by socioeconomic group, which can be easily incorporated into the simplified analysis, corrected for a substantial proportion of this bias and, for several interventions, corrected the direction of effect. Since smoking is a health-related behaviour with a large socioeconomic gradient in terms of treatment effects, it is plausible to conclude that the differences we observe between simplified and full DCEA might not be so pronounced for other intervention areas. However, it clearly exemplifies the importance of accounting for socioeconomic variation in treatment efficacy and uptake where possible.

The process of retrospectively adapting a decision model to conduct a DCEA raises some important considerations for future studies. First, we show that it is possible to identify data with which to adapt the model and conduct a full DCEA within a short timeframe and with limited resource. However, it must also be noted that the socioeconomic variation in smoking behaviours and outcomes are a well-researched public health issue, and that data for other health and disease areas might not be so widely available.

The principal approach for considering health inequalities in economic evaluations in public health has been cost-consequence analysis (Trueman & Anokye 2013), which may have provided the ‘dashboard’ of results presented in Table 5.5. However, presenting only the distribution of QALY impacts for each intervention could be misleading as this would ignore the inequality in the distribution of health opportunity costs, which fall more substantially on the most deprived. The analysis we present also provides a way to model and evaluate the changes in population health inequalities. The health inequality impacts we observe are also highly correlated with cost-effectiveness. This is in part due to the similar pattern of uptake and efficacy over socioeconomic groups that we assume across all interventions.

5.4.2 Limitations

In addition to the assumptions made in this analysis when adapting the existing decision model, our results also reflect the modelling uncertainties and concerns about heterogeneity in the evidence base affecting the latter. Of those made in this study, a number are influential and require further discussion.

The first are the limitations in the evidence base that prevent us from systematically varying parameters by socioeconomic status. For a number of these parameters, capturing this socioeconomic variation would have favoured health inequality reductions and potentially influenced the direction of our results. For example, we may expect the relative risk reduction of all-cause mortality to be greater for more deprived groups, as evidence indicates that they are heavier smokers (ONS 2016). An exception to this is the socioeconomic variation in odds of a successful quit attempt at four weeks to represent quit rates at 52 weeks. A larger gradient may have been expected over time, generating higher relative benefits for the least deprived. A comparison of short- and long-term quit rates in the analysis by Dobbie and colleagues suggest that this bias does exist but is not substantial (Dobbie et al. 2015, p.68).

Another important factor is the ability to differentiate between different types of intervention in terms of uptake and efficacy. We sought intervention-specific evidence of socioeconomic variation in quit rates, but located only one study that described certain types of behavioural intervention. Further studies of smoking cessation interventions should therefore seek to report results by socioeconomic status, so that over time these estimates could be improved (Bravin et al. 2015). Likewise, data on uptake by intervention are only available for broad intervention type, and are contaminated by systematic variation in the availability of services in each local authority. The influence this has on our results depends on whether the equal provision of services across authorities would increase or decrease the pro-rich gradient currently observed.

We also use a cost-effectiveness threshold and health opportunity cost distribution that is appropriate for an analysis conducted from a health sector perspective. For

this particular case study that might be reasonable given the predominance of health sector cost savings in determining the intervention results. Furthermore, the results of our analysis are not sensitive to alternative assumptions about the size and distribution of the health opportunity costs. For public health interventions with a greater proportion of non-health sector costs, the lack of a measure of health opportunity cost specific to public health sector funding may be more important, but this is also true with respect to routine cost-effectiveness analysis.

We are not able to estimate the independent effects of socioeconomic status on a number of model inputs. For example, as we were unable to adjust mortality rates by socioeconomic status for smoking prevalence, lower socioeconomic groups have higher baseline mortality that is in part due to a higher proportion of smokers. Baseline mortality is then adjusted again for smoking status when a higher relative risk of mortality is applied to those in the smoker state of the model. This generates higher absolute mortality risk reductions from quitting for high deprivation groups and has likely led to smaller inequality increases being estimated in our base case analysis.

5.4.3 Conclusions

Although quantifying health inequality impacts did not affect the rank order of the smoking cessation interventions included in this study when compared to population health gain alone, our results demonstrate that the negative health inequality impact can be reversed through improved uptake in lower socioeconomic groups. It is shown, with limited additional resources, how the standard tools of economic evaluations can be adapted to provide useful quantitative information on aspects of benefit other than total health gain. Last, an informative DCEA should incorporate, where possible, evidence on the socioeconomic patterns of efficacy and uptake across the set of interventions being compared, as shown by the large differences in inequality impact when compared to the simplified DCEA approach.

Appendix 5

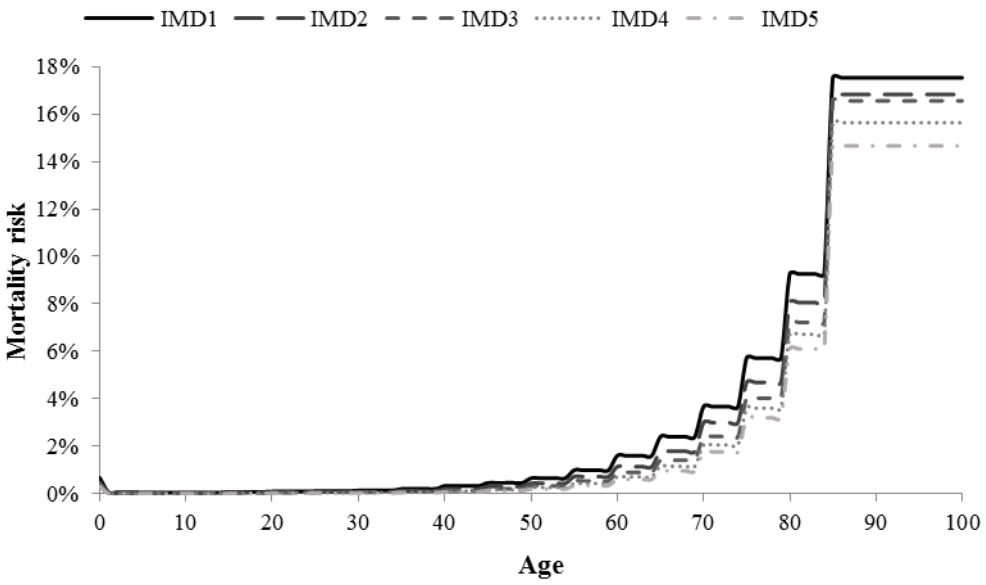
Table A5.8: Output from regression of health-related quality of life on age, smoking status and index of multiple deprivation (IMD) quintile group

Variable	**Coefficient**	**Standard error**
Constant	0.887***	(0.0131)
Age group		
16 to 24	Ref	
25 to 34	-0.00843	(0.0130)
35 to 44	-0.0565***	(0.0127)
45 to 54	-0.0817***	(0.0128)
55 to 64	-0.121***	(0.0130)
65 to 74	-0.130***	(0.0136)
75+	-0.171***	(0.0151)
Smoking status		
Former	Ref	
Current	-0.0423***	(0.00640)
IMD Quintile		
1 (Most deprived)	Ref	
2	0.0482***	(0.00905)
3	0.0486***	(0.00926)
4	0.0788***	(0.00938)
5 (Least deprived)	0.0994***	(0.00945)
N	5,447	
Adjusted R-Squared	0.048	

Notes:

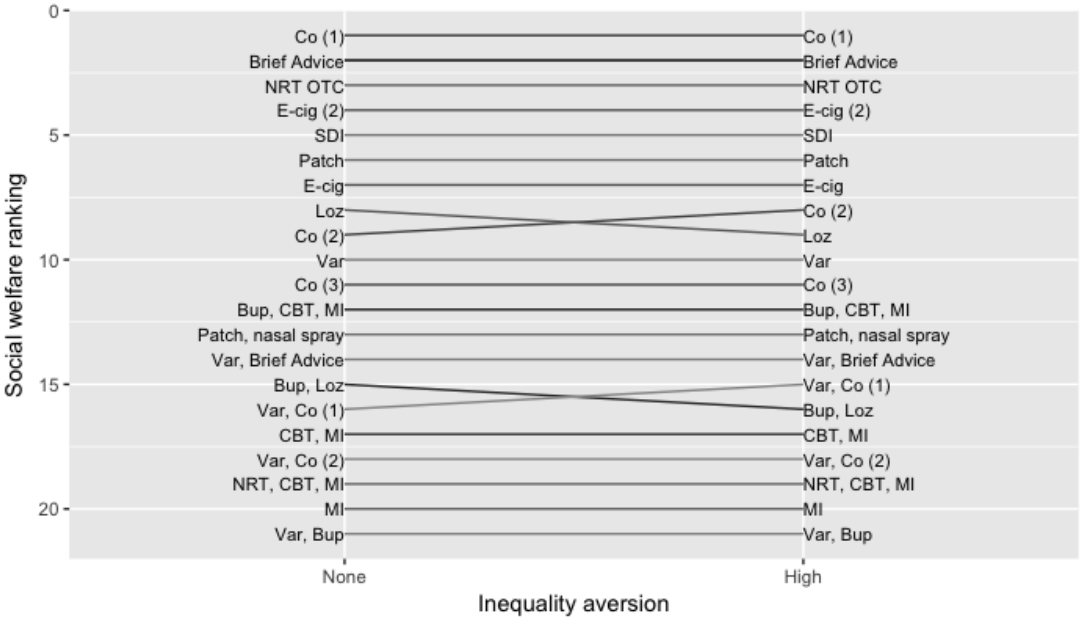
3. p<0.05 ** p<0.01 *** p<0.001
6. IMD quintiles 1 to 5 relate to the following IMD score ranges, respectively: 34.17 to 87.80; 21.35 to 34.17; 13.79 to 21.35; 8.49 to 13.79; 0.53 to 8.49.

Figure A5.7: Annual mortality rate by age and index of multiple deprivation (IMD) quintile group



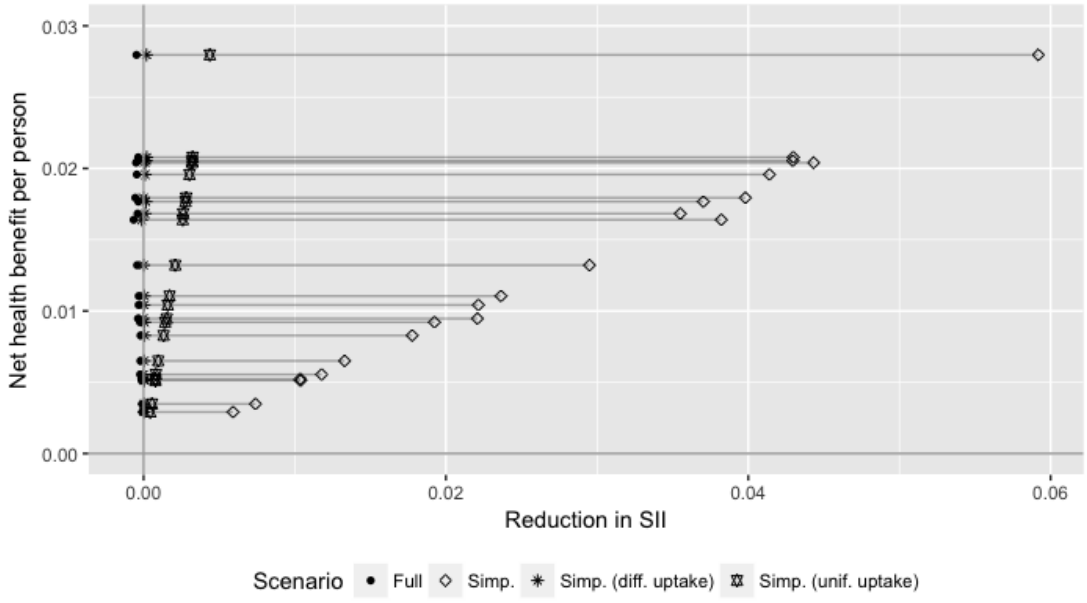
Note: IMD1 = most deprived, IMD5 = least deprived

Figure A5.8: Change in the Atkinson social welfare ranking of interventions when there is (i) no inequality aversion (ε=0) and (ii) high inequality aversion (ε=20)



Note: Social welfare ranking is determined by the change in the equally distributed equivalent health of the population

Figure A5.9: Comparison of equity impact plane locations of smoking cessation interventions when conducting full and simplified DCEA. This includes a simplified analysis in which interventions have 100% uptake across smokers



Notes:

1. Full = base case analysis; Simp. = simplified analysis; Simp. (diff. uptake) = simplified analysis with socioeconomic variation in uptake; Simp. (unif. uptake) = simplified analysis with uniform uptake over socioeconomic status
2. SII = slope index of inequality

Chapter 6: Conclusion

The objective of this thesis is to make empirical and methodological contributions to health economic evaluations that consider health inequality impacts. Throughout the preceding chapters, health inequalities are examined through disparities in lifetime health between gender and socioeconomic groups, which are both reflected in the univariate distribution of health. Although this is a limited definition, the techniques that are developed and presented in this thesis are flexible to any normative formulation of what constitutes unfair health inequality, and can therefore be redefined using any combination of social groups important to a decision-maker. Equally, they can be applied to alternative measures of health, such as quality-adjusted age at death or disability-adjusted life years. An opportunity for further research is to include additional social variables by which to stratify the population, such as ethnicity, to allow for more sources of inequality to be incorporated into distributional evaluations. However, there is a clear trade-off between greater stratification and practicality. Any increase in the number of social groups will have to be reflected at all stages of analysis: from the baseline distribution to the health opportunity costs and subgroup analysis of new treatments. Doing so would increase the data requirements and computational burden of conducting DCEAs.

In chapter 2, inequalities in quality-adjusted life expectancy are estimated for England. The distribution is designed to meet three important criteria: (i) to use a health metric consistent with cost-effectiveness analyses; (ii) to measure lifetime health; and (iii) to obtain estimates by equity-relevant subgroups. Differences between the most and least deprived fifths of the population are 11.9 QALYs. The univariate distribution, which incorporates gender as well as socioeconomic disparities, shows differences between the most and least healthy fifths of 11 QALYs.

The analysis makes several key contributions to the literature. It is the first to estimate health inequalities in terms of QALYs for the whole of the English population, using official national mortality statistics and a representative dataset of

over 25,000 observations to estimate quality of life weights. Unlike previous analyses (Collins 2013), this work also estimates the uncertainty around our estimates using Monte Carlo simulation. Furthermore, national inequality is estimated using a measure of morbidity that incorporates 245 health states, rather than the binary measures often used in health expectancy studies. This chapter also extends the traditional health expectancy analysis beyond measures of health inequality to estimate health-related social welfare. The equally distributed equivalent quality-adjusted life expectancy we estimate provides a measure of the social value of inequality expressed in terms of health.

The approach we outline has a number of uses in public policy and health technology assessment, and opens up avenues for further research. The methods we describe can be easily applied to new data and estimated for future years, generating a picture of health inequalities over time that more accurately reflects health experience than some other health indicators (ONS 2013). As is shown in subsequent chapters, the results can also be used as a baseline distribution of health, upon which health inequality impacts can be modelled. Due to the methods utilised, our results are likely to be more accurate than the approximations of QALE used in an earlier DCEA analysis (Asaria, Griffin, Cookson, et al. 2015). For example, our estimates are derived from Sullivan's method, a methodologically sound technique for estimating healthy life expectancy (Salomon et al. 2012), use IMD-specific mortality rates (rather than mapped rates from other socioeconomic measures) and estimate up-to-date health-related quality of life estimates from nationally representative data.

Chapter 3 addresses the other central methodological and empirical challenge associated with DCEA: the distribution of health opportunity costs. We utilise results from a major study (Claxton, Martin, et al. 2015) on the health effects of health care expenditure at the margin, which estimates the effects for each major disease area. We disaggregate these results by the age, gender and socioeconomic patterns observed in health care utilisation data. We find that the health of the poorest fifth is affected twice as much as the richest fifth by these marginal budget changes. Since expenditure reductions at the margin are a direct consequence of approving more costly new treatments, our findings represent the first estimates of the distribution of

health opportunity costs across social groups, and can feed directly into the evaluation of new technologies.

The results of this analysis are a novel contribution to the literature, since no study has looked at how health, by social group and in terms of both morbidity and mortality, is affected by national health care expenditures. The study is not without limitations, however. We make the strong assumption that an episode of health care generates the same health gain regardless of socioeconomic status. It is not clear what direction of bias this assumption will have on our results: the richer may be more effective at producing health from health care, but the poor are likely to be sicker before going to receive it. Another possibility for further research lies in estimating how the relationship between health and spending by disease area varies by deprivation level. The econometric models proposed by Claxton et al. (2015) offer a methodologically sound basis upon which to estimate interactions between mortality effects of expenditure and socioeconomic status, provided good quality data can be identified.

As the distribution we estimate relates to marginal expenditure *changes*, it can be applied to budget increases as well as the decreases. The results can effectively act as an 'equity benchmark' against which new interventions funded from expenditure increases can be judged in terms of their distributional consequences. The cost-per-EDE QALY thresholds we estimate offer an innovative way to adapt decision rules in economic evaluation to account for equity considerations. Previous work evaluating weighted or EDE QALYs in terms of benefits have not reflected these considerations in terms of health opportunity costs (Baeten et al. 2010; Lee et al. 2017). However, more work is needed on the practicality of summarising results in this way, in terms of both (i) communicating the intuition and meaning behind adjusted ICERs and thresholds, and (ii) presenting results in such a way so as not to embed a particular normative view about inequalities (through the inequality aversion parameter) into the decision making process.

In chapter 4 we propose an alternative way to conduct DCEAs that uses only the standard published estimates of mean per person incremental health benefits and

costs. The method is applied to 27 health technologies that have been approved by NICE. Five interventions increase population health and inequality, although in each case social welfare improves, indicating that the total health gains compensate for the inequality increases. We use manufacturer estimates of benefit and cost in our base case analysis. When a scenario analysis adjusts for potential bias in these values, we find the inequality impacts considerably less optimistic, with many more trade-offs occurring between the objectives of health maximisation and inequality reduction.

First and foremost, the simplified approach to DCEA offers a much less burdensome method of estimating inequality impacts, both in terms of development and computational time. This is clearly demonstrated through our being able to conduct 27 analyses in one study. The added value of the approach is that distributional analyses can be conducted in circumstances where (i) access to the original decision model is not available and (ii) data on how model inputs vary by equity-relevant groups.

A number of assumptions mean that our results cannot be interpreted as the health inequality impacts of NICE decisions. As mentioned, our base case analysis results use manufacturer estimates that regularly overstate health benefits and understate costs (Versoza et al. 2015). The adjustments we make in scenario analysis are through crude absolute changes to the ICER, which, although indicative, do not provide a robust and realistic alternative scenario. Second, a plethora of cancer drugs appraised by NICE over the timeframe of our analysis are not included in the sample due to the high volume of (i) therapies priced through confidential patient access schemes and (ii) Cancer Drugs Fund applications in that period for therapies that were rejected by NICE appraisal committees. A more complete picture would have to include such treatments, as they cover a substantial proportion of NICE appraisals and are more likely to have pro-rich distributions of patients.

The simplicity of the approach raises concerns over its validity. By not modelling differential mortality, quality of life, comorbidity risk or treatment efficacy, we do not account for a wide range of factors that can influence the distribution of costs and benefits. One aspect that we do not model but could be easily accommodated in the approach is variation in uptake. We make the optimistic assumption that new

interventions achieve 100% uptake in the relevant patient population. However, analyses adopting the simplified approach to evaluate single interventions could straightforwardly incorporate evidence or more realistic assumptions about market share and the associated socioeconomic variation. Doing so, as we show in chapter 5, leads to inequality impacts that are far closer to those yielded by a full DCEA.

Ultimately, the effect of limiting the sources of socioeconomic variation when estimating health inequality impacts is *a priori* uncertain. This is because treatment efficacy will likely benefit the affluent more, but relative mortality and comorbidity reductions will confer higher absolute benefits to the most deprived. We are able to examine these interactions in chapter 5, where a full DCEA is conducted on smoking cessation interventions. We find, despite substantially higher numbers of smokers coming from more deprived groups, that net health benefits are higher for more affluent groups and that smoking cessation interventions increase health inequalities. Differential uptake proves to be influential in these results: health inequality reductions are to be found by improving uptake across groups. All interventions generated population health benefits that again compensated for any increases to health inequality.

The simplified DCEA we conduct alongside this evaluation estimates an inequality impact in the opposite direction to the full analysis. We find that accounting for uptake was essential in order to have realistic estimates of the distributional effect. Ideally, analyses should also reflect differential uptake over socioeconomic groups. Doing so in our study substantially reduces the difference in inequality impact, and in several cases corrected the direction of effect. Furthermore, it should be noted that smoking cessation interventions are likely to involve larger socioeconomic variation in treatment efficacy than other types of treatments, with one study estimating that the least deprived have a 61% higher rate of quit success than the most deprived (Dobbie et al. 2015). As such large disparities in efficacy might not be common, we would not expect the disparity between the full and simplified approaches to be so great on average, particularly for interventions that do not require behaviour change.

Chapter 5 also demonstrates that it is possible to conduct full DCEAs using only a standard decision model and limited additional resource. We identified published and

freely available data on the socioeconomic variation of multiple key model inputs through pragmatic literature reviews. Given that smoking is health behaviour with strong associations with socioeconomic status, this could potentially be more challenging when analysing other interventions. To best identify heterogeneous treatment effects and other model inputs (including costs), however, individual patient-level data from UK-centred clinical trials is required, which can be linked to deprivation level through postcode. In this sense, the 'gold standard' of a randomised controlled trial holds true for DCEAs as well.

Uncertainty is a consideration that has not been feasible to explore in this thesis; only in the results for chapter 2 has it been appropriately characterised. For example, in chapter 3 the standard errors we calculate for the health opportunity cost distribution were negligible and only accounted for the uncertainty in the socioeconomic distributions of disease by age and gender. A full analysis would propagate the uncertainty quantified in the original analysis of Claxton et al. (2015) through Monte Carlo simulation. Doing so would require feeding each simulation iteration from their analysis into our own. Extensions and updates to this work should seek to address this technical challenge to show the impact of uncertainty in the distribution on the results of DCEAs of new interventions. Incorporating this and other sources of uncertainty into DCEAs represents an important area of future research, which should not only quantify uncertainty but also explore ways to present it to decision makers and develop appropriate graphical tools for visualisation.

Certain inputs have been shown to be more influential than others, and for which further research will invariably improve the robustness of DCEAs in future. In chapter 3, for example, we note how our estimates of health opportunity cost are quite dependent on the socioeconomic distribution of health care utilisation in respiratory disease, which accounts for over 30% of the health effects when the health sector budget is changed. As noted above, the uptake rates of interventions by socioeconomic status are also influential. Changes in the magnitude and distribution of uptake have a large influence in determining the health inequality impacts, which are intrinsically linked to the size of the recipient population. The analysis in chapter 4 highlights how crucial reliable estimates of incremental health and cost are when conducting a simplified DCEA. When making a simple adjustment to correct for

potential manufacturer bias, a substantial number of technologies change from being inequality reducing to increasing. Last, the analysis of chapters 4 and 5 show how the value of the cost-effectiveness threshold determines our results. At lower values, greater health opportunity costs are imposed, making the distribution we estimate in chapter 3 more influential in the overall inequality impact. The value of the threshold is even more influential when comparing cost-saving and cost-increasing interventions, which will have contrasting health inequality impacts.

Another issue common to all DCEAs that has yet to be addressed is what value to place on inequality reductions. Fundamentally this is a purely normative judgement as to what constitutes a sufficiently 'large' health inequality reduction. The health inequality impacts we estimate in chapters 4 and 5, in terms of the slope index of inequality, are mostly less than a hundredth of a QALY. This magnitude is a consequence of analysing the population as a whole and is *prima facie* difficult to weight relative to total health gain. However, a logical approach is to present unit changes in health inequality, such as changes in the SII, relative to cost, as is done for total health gain with the ICER. We also propose two alternative approaches to summarising inequality changes, though both involve specifying a normative judgement about inequality aversion. First, the change in EDE QALYs presents decision makers with direct comparison with population QALY impact; the societal value of the inequality change (the difference between the two) is therefore on same scale as total health gain. Second, the change in population EDE can be divided by total cost to obtain the cost-per-EDE QALY, a metric that integrates inequality reduction and health gain. Both of these statistics should, however, be presented for ranges of inequality aversion in order to prevent a specific value from being embedded in decision-making. How this is presented is another challenge for future research; throughout this thesis we present our base case results using an inequality aversion parameter elicited from the study of the general population in the UK (Robson et al. 2016).

Whilst these are all important aspects of DCEA that require attention in order to improve the robustness of the approach, this thesis provides a solid methodological and, in the case of the UK, empirical foundation that enables health inequality impacts to be explored in the quantitative analysis of health care.

Abbreviations

2PM	Two-part model
CCA	Cost-consequence analysis
CCG	Clinical Commissioning Group
DALE	Disability-adjusted life expectancy
DALY	Disability-adjusted life year
DCEA	Distributional cost-effectiveness analysis
DFLE	Disability-free life expectancy
DFLY	Disability-free life year
ECEA	Extended cost-effectiveness analysis
EDE	Equally distributed equivalent
GBD	Global Burden of Disease
HES	Hospital Episode Statistics
HLE	Healthy life expectancy
HSE	Health Survey for England
HRQL	Health-related quality of life
HRSWF	Health-related social welfare function
HTA	Health technology assessment
ICD	International Classification of Disease
ICER	Incremental cost-effectiveness ratio
IMD	Index of Multiple Deprivation
MAE	Mean absolute error
MAR	Missing at random
MCAR	Missing completely at random
MCDA	Multi-criteria decision analysis
MNAR	Missing not at random
MSE	Mean squared error
NICE	National Institute of Health and Care Excellence
NHB	Net health benefit
NHS	National Health Service
NS-SEC	National Statistics Socioeconomic Classification
OLS	Ordinary least squares
ONS	Office for National Statistics
PBC	Programme budgeting category
PCT	Primary Care Trust
QOF	Quality and Outcomes Framework
QAD	Quality-adjusted age at death
QALE	Quality-adjusted life expectancy
QALY	Quality-adjusted life year
RII	Relative index of inequality
SII	Slope index of inequality
SSRI	Selective serotonin reuptake inhibitors
TAC	Technology Appraisal Committee
WHO	World Health Organization

Bibliography

Adler, M., 2013. *Equity by the Numbers: Measuring Poverty, Inequality and Injustice*, Available at: http://ssrn.com/abstract=2263433.

Adler, M., 2011. *Well-being and fair distribution*, Oxford: Oxford University Press.

Ali, S. et al., 2017. How Robust Are Value Judgments of Health Inequality Aversion? Testing for Framing and Cognitive Effects. *Medical Decision Making*. Available at: http://journals.sagepub.com/doi/10.1177/0272989X17700842.

Asaria, M., Griffin, S., Cookson, R., et al., 2015. Distributional Cost-Effectiveness Analysis of Health Care Programmes - A Methodological Case Study of the UK Bowel Cancer Screening Programme. *Health Economics*, 24, pp.742–754.

Asaria, M. et al., 2016. How a universal health system reduces inequalities: lessons from England. *Journal of Epidemiology and Community Health*, 70(7), pp.637–643. Available at: http://jech.bmj.com/lookup/doi/10.1136/jech-2015-206742.

Asaria, M., Griffin, S. & Cookson, R., 2015. Distributional Cost-Effectiveness Analysis: A Tutorial. *Medical Decision Making*, pp.1–12. Available at: http://mdm.sagepub.com/cgi/doi/10.1177/0272989X15583266.

Asthana, S., Gibson, A. & Halliday, J., 2013. The medicalisation of health inequalities and the English NHS: the role of resource allocation. *Health Economics, Policy, and Law*, 8(2), pp.167–83. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22947257 [Accessed December 12, 2014].

Atkinson, A., 1970. On the measurement of inequality. *Journal of Economic Theory*, 2, pp.244–263.

Austin, P.C., Escobar, M. & Kopec, J.A., 2000. The use of the Tobit model for analyzing measures of health status. *Quality of Life Research*, 9(8), pp.901–910.

Baeten, S. et al., 2010. Incorporating equity-efficiency interactions in cost-effectiveness analysis - three approaches applied to breast cancer control. *Value In Health*, 13(5), pp.573–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20384980 [Accessed October 1, 2013].

Bajekal, M., 2005. Healthy life expectancy by area deprivation: magnitude and trends in England, 1994-1999. *Health Statistics Quarterly*, Spring(25), pp.18–

27.

Baltussen, R. et al., 2013. Balancing efficiency, equity and feasibility of HIV treatment in South Africa - development of programmatic guidance. *Cost Effectiveness and Resource Allocation*, 11(26), pp.1–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3851565&tool=pmcentrez&rendertype=abstract [Accessed August 10, 2014].

Baltussen, R. & Niessen, L.W., 2006. Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Effectiveness and Resource Allocation*, 4(14). Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1560167&tool=pmcentrez&rendertype=abstract [Accessed November 7, 2013].

Barnsley, P. et al., 2013. *Critique of CHE Research Paper 81: Methods for the Estimation of the NICE Cost Effectiveness Threshold*, London.

Barr, B., Bambra, C. & Whitehead, M., 2014. The impact of NHS resource allocation policy on health inequalities in England 2001-11: longitudinal ecological study. *British Medical Journal*, 348(7960), pp.1–10. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24865459 [Accessed September 25, 2014].

Barr, B., Higgerson, J. & Whitehead, M., 2017. Investigating the impact of the English health inequalities strategy: time trend analysis. *British Medical Journal*, p.j3310. Available at: http://www.bmj.com/lookup/doi/10.1136/bmj.j3310.

Basu, A. & Manca, A., 2012. Regression Estimators for Generic Health-Related Quality of Life and Quality-Adjusted Life Years. *Medical Decision Making*, 32(1), pp.56–69. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22009667%5Cnhttp://mdm.sagepub.com/cgi/doi/10.1177/0272989X11416988.

Beder, J.H. & Heim, R.C., 1990. On the use of ridit analysis. *Psychometrika*, 55(4), pp.603–616.

Bell, K. et al., 2007. *The Effectiveness of National Health Service Intensive Treatments for Smoking Cessation in England*, National Institute for Health and Care Excellence, London.

Bleichrodt, H., 1997. Health utility indices and equity considerations. *Journal of Health Economics*, 16(1), pp.65–91. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/10167345.

Bleichrodt, H., Diecidue, E. & Quiggin, J., 2004. Equity weights in the allocation of health care: the rank-dependent QALY model. *Journal of Health Economics*, 23(1), pp.157–71. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15154692 [Accessed August 8, 2014].

Bleichrodt, H., Doctor, J. & Stolk, E., 2005. A nonparametric elicitation of the equity-efficiency trade-off in cost-utility analysis. *Journal of Health Economics*, 24(4), pp.655–78. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15960990 [Accessed November 19, 2013].

Blondal, T. et al., 1999. Nicotine nasal spray with nicotine patch for smoking cessation: randomised trial with six year follow up. *British Medical Journal*, 318(7179), pp.285–288.

Bobinac, A. et al., 2012. Inquiry into the relationship between equity weights and the value of the QALY. *Value in Health*, 15(8), pp.1119–26. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23244815 [Accessed August 2, 2014].

Boniface, S. et al., 2012. *Health Survey for England 2011 - Volume 2: Methods and documentation*, London.

Bravin, J.I. et al., 2015. Socioeconomic predictors of smoking cessation in a worldwide online smoking cessation trial. *Internet Interventions*, 2(4), pp.410–418. Available at: http://dx.doi.org/10.1016/j.invent.2015.10.001.

Brazier, J. et al., 2013. *Eliciting societal preferences for burden of illness, therapeutic improvement and end of life for Value-Based Pricing: A report of the main survey*, Economic Evaluation Policy Research Unit, Sheffield.

Brazier, J., Roberts, J. & Deverill, M., 2002. The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics*, 21(2), pp.271–92. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11939242.

Brown, J. et al., 2014. Real-world effectiveness of e-cigarettes when used to aid smoking cessation: A cross-sectional population study. *Addiction*, 109(9), pp.1531–1540.

Brown, T., Platt, S. & Amos, A., 2014. Equity impact of European individual-level smoking cessation interventions to reduce smoking in adults: a systematic review. *The European Journal of Public Health*, 24(4), pp.551–556. Available at: http://eurpub.oxfordjournals.org/content/24/4/551?rss%3D1=.

Burger, S., 2010. A cost consequence analysis of outreach strategies for high risk

pregnant women. *Journal of Community Health Nursing*, 27(3), pp.137–45. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20694876 [Accessed October 3, 2014].

CADTH, 2006. *Guidelines for the Economic Evaluation of Health Technologies, Canada [3rd Edition]*, Ottawa.

Canadian Institute for Health Information, 2015. *Trends in Income-Related Health Inequalities in Canada*, Ontario.

Caponnetto, P. et al., 2013. EffiCiency and Safety of an eLectronic cigAreTte (ECLAT) as Tobacco Cigarettes Substitute: A Prospective 12-Month Randomized Control Design Study. *PLoS ONE*, 8(6), pp.1–12.

CDC, 2013. *CDC Health Disparities and Inequalities Report - United States, 2013*, Washington DC. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24264487.

Chengappa, K.N.R. et al., 2014. Varenicline for Smoking Cessation in Bipolar Disorder: A Randomized, Double-Blind, Placebo-Controlled Study. *Journal of Clinical Psychiatry*, 75(7), pp.765–772.

Chiang, C.L., 1972. On Constructing Current Life Tables. *Journal of the American Statistical Association*, 67(339), pp.538–541.

Chiang, C.L., 1983. *The Life Table And Its Applications*, Florida: Krieger.

Claxton, K. et al., 2010. *Appropriate Perspectives for Health Care Decisions*, York: Centre for Health Economics.

Claxton, K., Sculpher, M., et al., 2015. Causes for concern: Is NICE failing to uphold its responsibilities to all NHS patients? *Health Economics*, 24, pp.1–7.

Claxton, K., Martin, S., et al., 2015. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technology Assessment*, 19(14), pp.1–504. Available at: http://www.journalslibrary.nihr.ac.uk/hta/volume-19/issue-14.

Claxton, K. & Sculpher, M., 2015. *Response to the OHE critique of CHE Research paper 81*, York.

Cleary, S., Mooney, G. & Mcintyre, D.I., 2010. Equity and efficiency in HIV-treatment in South Africa: The contribution of mathematical programming to priority setting. *Health Economics*, 19, pp.1166–1180.

Coast, J., 2004. Is economic evaluation in touch with society's health values? *British Medical Journal*, 329(November), pp.1233–1236.

Cochrane, A., Leger, A.S. & Moore, F., 1978. Health service "input" and mortality

"output" in developed countries. *Journal of Epidemiology and Community Health*, 32, pp.200–205. Available at: http://jech.bmj.com/content/32/3/200.short.

Collins, B., 2017. Results from a Well-Being Survey in the North West of England: Inequalities in EQ-5D–Derived Quality-Adjusted Life Expectancy Are Mainly Driven by Pain and Mental Health. *Value in Health*, 20(1), pp.174–177. Available at: http://dx.doi.org/10.1016/j.jval.2016.08.004.

Collins, B., 2013. Using a survey to estimate health expectancy and quality-adjusted life expectancy to assess inequalities in health and quality of life. *Value in Health*, 16(4), pp.599–603. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23796294 [Accessed July 1, 2013].

Collins, M. & Latimer, N., 2013. NICE's end of life decision making scheme: impact on population health. *BMJ (Clinical research ed.)*, 346(March), p.f1363. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23518274 [Accessed November 4, 2014].

Cookson, R., Asaria, M., et al., 2016. Health Equity Indicators for the English NHS: a longitudinal whole-population study at the small-area level. *Health Services and Delivery Research*, 4(26), pp.1–224. Available at: https://www.journalslibrary.nihr.ac.uk/hsdr/hsdr04260/.

Cookson, R., Propper, C., et al., 2016. Socioeconomic inequalities in health care in England. *Fiscal Studies*, 37(3–4), pp.371–403.

Cookson, R. et al., 2017. Using Cost-Effectiveness Analysis to Address Health Equity Concerns. *Value in Health*, 20(2), pp.206–212.

Cookson, R., Drummond, M. & Weatherly, H., 2009. Explicit incorporation of equity considerations into economic evaluation of public health interventions. *Health Economics, Policy and Law*, 4(Pt 2), pp.231–45. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19216834 [Accessed May 22, 2013].

Crawford, R. & Emmerson, C., 2012. *NHS and social care funding: the outlook to 2021/22*, London.

Crémieux, P.Y., Ouellette, P. & Pilon, C., 1999. Health care spending as determinants of health outcomes. *Health Economics*, 8(7), pp.627–639.

Currie, J., 2009. Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. *Journal of Economic Literature*, 47(1), pp.87–122.

Daniels, N., 2000. Accountability for reasonableness. *British Medical Journal*, 321(November), pp.1300–1301.

Department of Health, 2013. *The NHS Constitution: The NHS belongs to us all*, London.

Department of Health and Ageing, 2009. *Review Health Technology Assessment in Australia*, Barton.

DH, 2006. *Health Equity Audit A Guide for the NHS*, London.

DH, 2013. *The Public Health Outcomes Framework for England, 2013-2016*, London.

DHSS, 1980. *Inequalities in Health: Report of a Research Working Group*, London.

Dillon, A., 2015. Carrying NICE over the threshold. Available at: https://www.nice.org.uk/news/blog/carrying-nice-over-the-threshold [Accessed August 1, 2017].

Dobbie, F. et al., 2015. Evaluating long-term outcomes of NHS stop smoking services (ELONS): A prospective cohort study. *Health Technology Assessment*, 19(95), pp.9–155.

Dolan, P. et al., 1995. *A Social Tariff for EuroQol: Results from a UK General Population Survey*, York.

Dolan, P. et al., 2005. QALY maximisation and people's preferences: a methodological review of the literature. *Health Economics*, 14(2), pp.197–208. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15386656 [Accessed October 30, 2014].

Dolan, P., 1998. The measurement of individual utility and social welfare. *Journal of Health Economics*, 17, pp.39–52.

Dolan, P., Edlin, R. & Tsuchiya, A., 2008. *The relative societal value of health gains to different beneficiaries*, Sheffield.

Dolan, P. & Olsen, J.A., 2001. Equity in health: the importance of different health streams. *Journal of Health Economics*, 20(5), pp.823–34. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11558650.

Dolan, P. & Tsuchiya, A., 2005. Health priorities and public preferences: the relative importance of past health experience and future health prospects. *Journal of Health Economics*, 24(4), pp.703–14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15960992 [Accessed October 30, 2014].

Dolan, P. & Tsuchiya, A., 2006. The elicitation of distributional judgements in the

context of economic evaluation. In A. Jones, ed. *The Elgar Companion to Health Economics*. Cheltenham: Edward Elgar Publishing, pp. 382–391.

Drummond, M. et al., 2005. *Methods for the Economic Evaluation of Health Care Programmes* 3rd ed., Oxford: Oxford University Press.

Drummond, M., 1989. Output Measurement for Resource Allocation Decisions in Heath Care. *Oxford Review of Economic Policy*, 5(1), pp.59–74.

Earnshaw, S.R. et al., 2002. Optimal Allocation of Resources Across Four Interventions for Type 2 Diabetes. *Medical Decision Making*, 22(5 suppl), pp.S80–S91. Available at: http://mdm.sagepub.com/cgi/doi/10.1177/027298902237704 [Accessed November 6, 2014].

Eberth, B. et al., 2014. Smoking-related disease risk, area deprivation and health behaviours. *Journal of Public Health (United Kingdom)*, 36(1), pp.72–80.

Epstein, D.M. et al., 2007. Efficiency, equity, and budgetary policies: informing decisions using mathematical programming. *Medical Decision Making*, 27(2), pp.128–37. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17409363 [Accessed October 20, 2014].

Feeny, D. et al., 2002. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Medical Care*, 40(2), pp.113–28. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11802084.

Fleurbaey, M. & Schokkaert, E., 2009. Unfair inequalities in health and health care. *Journal of Health Economics*, 28(1), pp.73–90. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18829124 [Accessed December 17, 2013].

Fleurence, R.L. & Hollenbeak, C.S., 2007. Rates and probabilities in economic modelling: transformation, translation and appropriate application. *PharmacoEconomics*, 25(1), pp.3–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17192114.

Gakidou, E.E., Murray, C.J. & Frenk, J., 2000. Defining and measuring health inequality: an approach based on the distribution of health expectancy. *Bulletin of the World Health Organization*, 78(1), pp.42–54. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2560605&tool=pmcentrez&rendertype=abstract.

Gallet, C.A. & Doucouliagos, C., 2015. *The impact of healthcare spending on health*

outcomes: A meta-regression analysis, Victoria.

Goetghebeur, M.M. et al., 2012. Bridging health technology assessment (HTA) and efficient health care decision making with multicriteria decision analysis (MCDA): applying the EVIDEM framework to medicines appraisal. *Medical Decision Making*, 32(2), pp.376–88. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21987539 [Accessed September 26, 2014].

Goetghebeur, M.M. et al., 2010. Combining multicriteria decision analysis, ethics and health technology assessment: applying the EVIDEM decision-making framework to growth hormone for Turner syndrome patients. *Cost Effectiveness and Resource Allocation*, 8(4), pp.1–15. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2856527&tool=pmcentrez&rendertype=abstract.

Goetghebeur, M.M. et al., 2008. Evidence and Value: Impact on DEcisionMaking: the EVIDEM framework and potential applications. *BMC Health Services Research*, 8, p.270. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2673218&tool=pmcentrez&rendertype=abstract [Accessed September 15, 2014].

Graham, H. & Kelly, M.P., 2004. *Health inequalities: concepts, frameworks and policy*, London.

Gravelle, H.S.E. & Backhouse, M.E., 1987. International cross-section analysis of the determination of mortality. *Social Science & Medicine*, 25(5), pp.427–441.

Green, C., 2009. Investigating public preferences on "severity of health" as a relevant condition for setting healthcare priorities. *Social Science & Medicine*, 68(12), pp.2247–55. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19406545 [Accessed October 30, 2014].

Grossman, M., 1972. On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, 80(2), pp.223–255.

Guindo, L.A. et al., 2012. From efficacy to equity: Literature review of decision criteria for resource allocation and healthcare decisionmaking. *Cost Effectiveness and Resource Allocation : C/E*, 10(1), p.9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3495194&tool=pmcentrez&rendertype=abstract [Accessed August 14, 2014].

Harberger, A.C., 1978. On the Use of Distributional Weights in Social Cost-Benefit

Analysis. *Journal of Political Economy*, 86(2), pp.S87–S120.

Health and Social Care Information Centre, 2015. Quality and Outcomes Framework. Available at: http://www.hscic.gov.uk/qof [Accessed November 27, 2015].

Heydari, G. et al., 2012. Quitting smoking with varenicline: parallel, randomised efficacy trial in Iran. *International Journal of Tuberculosis and Lung Disease*, 16(2), pp.268–272.

Hiscock, R. et al., 2013. Behavioural therapy for smoking cessation: The effectiveness of different intervention types for disadvantaged and affluent smokers. *Addictive Behaviors*, 38, pp.2787–2796. Available at: http://dx.doi.org/10.1016/j.addbeh.2013.07.010.

Hiscock, R., Dobbie, F. & Bauld, L., 2015. Smoking cessation and socioeconomic status: An update of existing evidence from a national evaluation of english stop smoking services. *BioMed Research International*, 2015.

HM Treasury, 2011. *The Green Book: Appraisal and Evaluation in Central Government*, London.

House of Commons Health Committee, 2009. *Health Inequalities: Third Report of Session 2008-09*, London.

HSCIC, 2014. *Statistics on smoking: England 2014*, Leeds.

HSCIC, 2016. *Statistics on smoking: England 2016*, Leeds.

Issa, J.S. et al., 2013. Effectiveness of coadministration of varenicline, bupropion, and serotonin reuptake inhibitors in a smoking cessation program in the real-life setting. *Nicotine and Tobacco Research*, 15(6), pp.1146–1150.

Jia, H., Zack, M.M. & Thompson, W.W., 2011. State Quality-Adjusted Life Expectancy for U.S. adults from 1993 to 2008. *Quality of Life Research*, 20(6), pp.853–863. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21210226 [Accessed September 20, 2013].

Johri, M. & Norheim, O.F., 2012. Can cost-effectiveness analysis integrate concerns for equity? Systematic review. *International Journal of Technology Assessment in Health Care*, 28(2), pp.125–32. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22494637 [Accessed August 7, 2014].

Jorenby, D.E. et al., 2006. Efficacy of Varenicline, an alpha-4-beta-2 Nicotinic Acetylcholine Receptor Partial Agonist, vs Placebo or Sustained-Release Bupropion for Smoking Cessation: A Randomized Controlled Trial. *Journal of*

the American Medical Association, 296(1), pp.56–63.

Jull, J. et al., 2017. When is a randomised controlled trial health equity relevant? Development and validation of a conceptual framework. *BMJ Open*, 7(9).

Kamm, F.M., 2013. Rationing and the disabled. In N. Eyal et al., eds. *Inequalities in Health*. New York: Oxford University Press, pp. 240–259.

Lee, T. et al., 2017. Strategic Distributional Cost-Effectiveness Analysis for Improving National Cancer Screening Uptake in Cervical Cancer : A Focus on Regional Inequality in South Korea. *Cancer Research and Treatment*, pp.1–10.

Lindholm, L. & Rosen, M., 1998. On the measurement of the nation's equity adjusted health. *Health Economics*, 7, pp.621–628.

Lindholm, L., Rosén, M. & Emmelin, M., 1998. How many lives is equity worth ? A proposal for equity adjusted years of life saved. *Journal of Epidemiology & Community Health*, 52, pp.808–811.

Little, R. & Rubin, D., 2002. *Statistical Analysis with Missing Data* 2nd ed., Hoboken, New Jersey: Wiley-Interscience.

Lorenc, T. et al., 2013. What types of interventions generate inequalities? Evidence from systematic reviews. *Journal of Epidemiology and Community Health*, 67(2), pp.190–193. Available at: http://jech.bmj.com/lookup/doi/10.1136/jech-2012-201257.

Lowdell, C., Bardsley, M. & Morgan, D., 1999. *Acheson Report: The inquiry into inequalities in health - implications for London*, London.

Maestad, O. & Norheim, O.F., 2009. Eliciting people's preferences for the distribution of health: A procedure for a more precise estimation of distributional weights. *Journal of Health Economics*, 28(3), pp.570–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19368983 [Accessed August 29, 2014].

Maheswaran, H. et al., 2013. Estimating EQ-5D utility values for major health behavioural risk factors in England. *Journal of Epidemiology & Community Health*, 67(2), pp.172–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22844084 [Accessed July 1, 2013].

Manuel, D.G. & Schultz, S.E., 2004. Health-related quality of life and health-adjusted life expectancy of people with diabetes in Ontario, Canada, 1996-1997. *Diabetes Care*, 27(2), pp.407–414. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14747221.

Marmot, M. et al., 2008. *Closing the gap in a generation: health equity through action on the social determinants of health.*, Geneva.

Marmot, M. et al., 1997. Social inequalities in health: next questions and converging evidence. *Social Science & Medicine*, 44(6), pp.901–10. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9080570.

Marmot, M. et al., 2010. *The Marmot Review: Fair Society, Healthy Lives*, London.

Martin, S., Rice, N. & Smith, P.C., 2012. Comparing costs and outcomes across programmes of health care. *Health Economics*, 21, pp.316–337.

Martin, S., Rice, N. & Smith, P.C., 2008. Does health care spending improve health outcomes? Evidence from English programme budgeting data. *Journal of Health Economics*, 27(4), pp.826–842.

Marton, J., Sung, J. & Honore, P., 2015. Does More Public Health Spending Buy Better Health? *Health Services Research and Managerial Epidemiology*, 2(0). Available at: http://hme.sagepub.com/lookup/doi/10.1177/2333392815580750.

Mathers, C.D. et al., 2001. World Health Report 2000 Healthy life expectancy in 191 countries , 1999. *The Lancet*, 357, pp.1685–1691.

Mathes, T. et al., 2013. Methods of international health technology assessment agencies for economic evaluations- a comparative analysis. *BMC Health Services Research*, 13(371), pp.1–10. Available at: BMC Health Services Research.

Mauskopf, J.A. et al., 1998. The role of cost-consequence analysis in healthcare decision-making. *PharmacoEconomics*, 13(3), pp.277–88. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10178653.

Mbuagbaw, L. et al., 2017. Considerations and guidance in designing equity-relevant clinical trials. *International Journal for Equity in Health*, 16(93).

McCabe, C., Claxton, K. & Culyer, A.J., 2008. The NICE cost-effectiveness threshold: What it is and what that means. *PharmacoEconomics*, 26(9), pp.733–744.

McLennan, D. et al., 2011. *The English Indices of Deprivation 2010: Technical Report*, London.

Minet Kinge, J. & Morris, S., 2010. Socioeconomic variation in the impact of obesity on health-related quality of life. *Social Science & Medicine*, 71(10), pp.1864–71. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20932623 [Accessed July 2, 2013].

Miot, J. et al., 2012. Field testing of a multicriteria decision analysis (MCDA) framework for coverage of a screening test for cervical cancer in South Africa. *Cost Effectiveness and Resource Allocation : C/E*, 10(1), p.2. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3330006&tool=pmcentrez&rendertype=abstract [Accessed October 27, 2014].

Mirelman, A. et al., 2012. Decision-making criteria among national policymakers in five countries: a discrete choice experiment eliciting relative preferences for equity and efficiency. *Value in Health*, 15(3), pp.534–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22583464 [Accessed July 25, 2014].

Morris, S., Carr-Hill, R. & Vallejo-Torres, L., 2007. *Combining Age Related and Additional Needs (CARAN) Report*,

Morton, A., 2014. Aversion to health inequalities in healthcare prioritisation: a multicriteria optimisation perspective. *Journal of Health Economics*, 36, pp.164–73. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24831800 [Accessed September 30, 2014].

Murray, C.J.L., 1994. Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bulletin of the World Health Organization*, 72(3), pp.429–445.

NICE, 2009. *Appraising life-extending, end of life treatments*, London.

NICE, 2014. *Developing NICE guidelines: the manual*, London.

NICE, 2004. *Guide to the Methods of Technology Appraisal*, London.

NICE, 2013. *Guide to the methods of technology appraisal 2013*, London.

NICE, 2012a. *Health inequalities and population health*, London.

NICE, 2006. *Smoking: brief interventions and referrals: NICE Guideline [PH1]*, London.

NICE, 2008a. *Social Value Judgements: Principles for the development of NICE guidance*, London.

NICE, 2008b. *Stop smoking services: NICE Guideline [PH10]*, London.

NICE, 2012b. *The NICE public health guidance development process (third edition)*, London.

Nixon, J. & Ulmann, P., 2006. The relationship between health care expenditure and health outcomes: Evidence and caveats for a causal link. *The European Journal of Health Economics*, 7(1), pp.7–18.

Noble, M. et al., 2004. *The English Indices of Deprivation 2004: Summary (revised)*,

London.

Nord, E., 2005. Concerns for the worse off: fair innings versus severity. *Social Science & Medicine*, 60(2), pp.257–63. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15522483 [Accessed October 30, 2014].

Nord, E. et al., 1999. Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Economics*, 8(1), pp.25–39.

Nord, E. et al., 1995. Maximizing health benefits vs egalitarianism: An Australian survey of health issues. *Social Science & Medicine*, 41(10), pp.1429–1437. Available at: http://linkinghub.elsevier.com/retrieve/pii/027795369500121M.

Nord, E., 1993. The trade-off between severity of illness and treatment effect in cost-value analysis of health care. *Health Policy*, 24(3), pp.227–238. Available at: http://linkinghub.elsevier.com/retrieve/pii/016885109390042N.

Norheim, O.F., 2013. Atkinson's Measure of Inequality: Can Measures of Economic Inequality Help Us Understand Trade-Offs in Healthcare Priority Setting? In N. Eyal et al., eds. *Inequalities in Health*. Oxford: Oxford University Press, pp. 214–231.

Norwegian Ministry of Health and Care Services, 2007. *National strategy to reduce social inequalities in health*, Oslo.

Office for National Statistics, 2014. *Population Estimates for UK, England and Wales, Scotland and Northern Ireland, Mid-2013*, London.

ONS, 2016. *Adult smoking habits in Great Britain, 2014*, London. Available at: http://www.ons.gov.uk/ons/rel/ghs/opinions-and-lifestyle-survey/adult-smoking-habits-in-great-britain--2014/stb-adult-smoking-habits-in-great-britain--2014.html.

ONS, 2013. *Inequality in Disability-Free Life Expectancy by Area Deprivation: England, 2003-06 and 2007-10*, London.

Or, Z., 2001. *Exploring the effects of health care on mortality across OECD countries*,

Paulden, M. et al., 2014. Some Inconsistencies in NICE's Consideration of Social Values. *PharmacoEconomics*. Available at: http://link.springer.com/10.1007/s40273-014-0204-4 [Accessed August 25, 2014].

Peacock, S.J. et al., 2007. Priority setting in health care using multi-attribute utility theory and programme budgeting and marginal analysis (PBMA). *Social*

Science & Medicine, 64(4), pp.897–910. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17141930 [Accessed October 10, 2014].

Petrou, S. & Kupek, E., 2008. Social Capital and its relationship with measures of Health Status: Evidence from the Health Survey of England 2003. *Health Economics*, 143(17), pp.127–143.

PHAA, 2012. *Public Health Association of Australia: Policy-at-a-glance – Health Inequities Policy*, Canberra.

Powers, M. & Faden, R.R., 2000. Inequalities in Health, Inequalities in Health Care: Four Generations of Discussion about Justice and Cost-Effectiveness Analysis. *Kennedy Institute of Ethics Journal*, 10(2), pp.109–127. Available at: http://muse.jhu.edu/content/crossref/journals/kennedy_institute_of_ethics_journal/v010/10.2powers.html [Accessed July 27, 2014].

Public Health England, 2014. *Cancer by Deprivation in England: Incidence, 1996-2010 and Mortality, 1997-2011*, London.

Rabin, R. & de Charrao, F., 2001. EQ-5D: a measure of health status from the EuroQol Group. *Annals of Medicine*, 33, pp.337–343.

Rabin, R., Oemar, M. & Oppe, M., 2011. *EQ-5D-3L User Guide*, Rotterdam.

Raftery, J.P., 2014. NICE's cost-effectiveness range: Should it be lowered? *PharmacoEconomics*, 32, pp.613–615.

Raphael, D., 2000. Health inequalities in Canada: current discourses and implications for public health action. *Critical Public Health*, 10(2), pp.193–216.

Rawlins, M., Barnett, D. & Stevens, A., 2010. Pharmacoeconomics: NICE's approach to decision-making. *British Journal of Clinical Pharmacology*, 70(3), pp.346–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2949905&tool=pmcentrez&rendertype=abstract [Accessed November 4, 2014].

Rigotti, N.A. et al., 2010. Efficacy and safety of varenicline for smoking cessation in patients with cardiovascular disease: A randomized trial. *Circulation*, 121(2), pp.221–229.

Robberstad, B. & Norheim, O.F., 2011. Incorporating concerns for equal lifetime health in evaluations of public health programs. *Social Science & Medicine*, 72(10), pp.1711–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21531492 [Accessed August 29, 2014].

Robine, J.-M., Romieu, I. & Cambois, E., 1999. Health expectancy indicators.

Bulletin of the World Health Organization, 77(2), pp.181–5. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2557596&tool=pmcentrez&rendertype=abstract.

Robson, M. et al., 2016. Eliciting the level of health inequality aversion in England. *Health Economics*.

Roset, M. & Gaminde, I., 1999. Quality adjusted life expectancy. In J. M. Cabasés & I. Gaminde, eds. *Discussion Papers / 17th Plenary Meeting of the Euroqol Group*. Pamplona: Euroqol Group, pp. 173–182.

Salomon, J.A. et al., 2012. Healthy life expectancy for 187 countries, 1990-2010: a systematic analysis for the Global Burden Disease Study 2010. *Lancet*, 380(9859), pp.2144–62. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23245606 [Accessed May 28, 2013].

Salti, N., Brouwer, E. & Verguet, S., 2016. The health, financial and distributional consequences of increases in the tobacco excise tax among smokers in Lebanon. *Social Science & Medicine*, 170, pp.161–169.

Sassi, F., Archard, L. & Le Grand, J., 2001. Equity and the economic evaluation of healthcare. *Health Technology Assessment*, 5(3), pp.1–138. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11207451.

Sen, A., 2002. Why health equity? *Health Economics*, 11(8), pp.659–66. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12457367 [Accessed August 23, 2014].

Shah, K.K. et al., 2011. *NICE's social value judgements about equity in health and health care*, York.

Shah, K.K., 2009. Severity of illness and priority setting in healthcare: a review of the literature. *Health Policy*, 93(2–3), pp.77–84. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19712990 [Accessed June 10, 2014].

Shah, K.K., Tsuchiya, A. & Wailoo, A.J., 2014. Valuing health at the end of life: an empirical study of public preferences. *The European Journal of Health Economics*, 15(4), pp.389–99. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23657476 [Accessed December 13, 2014].

Shiell, A., 2009. Still waiting for the great leap forward. *Health Economics, Policy and Law*, 4(Pt 2), pp.255-60–3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19216832 [Accessed October 1, 2013].

Singh, S.R., 2014. Public Health Spending and Population Health. *American Journal of Preventive Medicine*, 47(5), pp.634–640. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0749379714002359.

Smith, S.S. et al., 2009. Comparative effectiveness of 5 smoking cessation pharmacotherapies in primary care clinics. *Archives of Internal Medicine*, 169(22), pp.2148–55. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2891174&tool=pmcentrez&rendertype=abstract.

Sullivan, D.F., 1971. A single index of mortality and morbidity. *HSMHA Health Reports*, 86(4), pp.347–54. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1937122&tool=pmcentrez&rendertype=abstract.

Sutton, M. et al., 2002. *Allocation of Resources to English Areas: Individual and small area determinants of morbidity and use of healthcare resources. Report to the Department of Health*, Edinburgh.

Tanios, N. et al., 2013. Which criteria are considered in healthcare decisions? Insights from an international survey of policy and clinical decision makers. *International Journal of Technology Assessment in Health Care*, 29(4), pp.456–65. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24290340 [Accessed October 23, 2014].

Taylor, L. & Blair-Stevens, C., 2002. *Introducing health impact assessment (HIA): Informing the decision-making process*, London.

Thokala, P. & Duenas, A., 2012. Multiple criteria decision analysis for health technology assessment. *Value in Health*, 15(8), pp.1172–81. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23244821.

Tobin, J., 1958. Estimation of Relationships for Limited Dependent Variables. *Econometrica*, 26(1), pp.24–36.

Tony, M. et al., 2011. Bridging health technology assessment (HTA) with multicriteria decision analyses (MCDA): field testing of the EVIDEM framework for coverage decisions by a public payer in Canada. *BMC Health Services Research*, 11(1), p.329. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3248909&tool=pmcentrez&rendertype=abstract [Accessed September 15, 2014].

Tranvåg, E.J., Ali, M. & Norheim, O.F., 2013. Health inequalities in Ethiopia:

modeling inequalities in length of life within and between population groups. *International Journal for Equity in Health*, 12, p.52. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3716728&tool=pmcentrez&rendertype=abstract [Accessed October 15, 2014].

Trueman, P. & Anokye, N.K., 2013. Applying economic evaluation to public health interventions: the case of interventions to promote physical activity. *Journal of Public Healthublic health*, 35(1), pp.32–39. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3580051&tool=pmcentrez&rendertype=abstract [Accessed October 3, 2014].

Tsuchiya, A., 2000. QALYs and ageism: Philosophical theories and age weighting. *Health Economics*, 9, pp.57–68.

Turrell, G. et al., 2006. *Health inequalities in Australia: morbidity, health behaviours, risk factors and health service use*, Canberra.

Ubel, P. a, 1999. How stable are people's preferences for giving priority to severely ill patients? *Social Science & Medicine*, 49(7), pp.895–903. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0277953699001744.

Valkonen, T., Sihvonen, A.-P. & Lahelma, E., 1997. Health expectancy by level of education in Finland. *Social Science & Medicine*, 44(6), pp.801–808. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0277953696001906.

Vallejo-Torres, L. et al., 2016. On the Estimation of the Cost-Effectiveness Threshold: Why, What, How? *Value in Health*, 19(5), pp.558–566.

Verguet, S. et al., 2013. Public finance of rotavirus vaccination in India and Ethiopia: an extended cost-effectiveness analysis. *Vaccine*, 31(42), pp.4902–10. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23871824 [Accessed February 4, 2014].

Verguet, S., Laxminarayan, R. & Jamison, D.T., 2014. Universal public finance of tuberculosis treatment in India: An extended cost-effectiveness analysis. *Health Economics*.

Versoza, L. et al., 2015. Do Evidence Review Groups Bias NICE Decisions? In *International Society for Pharmacoeconomic Outcomes Research 18th Annual European Congress*. Milan.

Vogl, M. et al., 2012. Smoking and health-related quality of life in English general population: implications for economic evaluations. *BMC Public Health*, 12(1), p.203. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3352300&tool=pmcentrez&rendertype=abstract [Accessed July 3, 2013].

Wagstaff, A., 1991. QALYs and the equity-efficiency trade-off. *Journal of Health Economics*, 10(1), pp.21–41. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10113661.

Wailoo, A., Tsuchiya, A. & McCabe, C., 2009. Weighting Must Wait. *PharmacoEconomics*, 27(12), pp.983–989.

Weatherly, H. et al., 2009. Methods for assessing the cost-effectiveness of public health interventions: key challenges and recommendations. *Health Policy*, 93(2–3), pp.85–92. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19709773 [Accessed June 6, 2014].

West, R., 2006. *Background smoking cessation rates in England*,

White, C. & Butt, A., 2015. *Inequality in health and life expectancies within Upper Tier Local Authorities : 2009 to 2013*,

Whitehead, S.J. & Ali, S., 2010. Health outcomes in economic evaluation: the QALY and utilities. *British Medical Bulletin*, 96, pp.5–21. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21037243 [Accessed November 19, 2013].

Wille, N. et al., 2010. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Quality of Life Research*, 19(6), pp.875–86. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2892611&tool=pmcentrez&rendertype=abstract [Accessed June 17, 2013].

Williams, A., 1997. Intergenerational Equity: An Exploration of the "Fair Innings" Argument. *Health Economics*, 6, pp.117–132. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9158965.

Williams, A., 2001. The "fair innings argument" deserves a fairer hearing! Comments by Alan Williams on Nord and Johannesson. *Health Economics*, 10(7), pp.583–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11747042.

Williams, A. & Cookson, R., 2006. Equity-efficiency trade-offs in health technology assessment. *International Journal of Technology Assessment in Health Care*, 22(1), pp.1–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16673674.

Williams, G.C. et al., 2006. Testing a self-determination theory intervention for motivating tobacco cessation: Supporting autonomy and competence in a clinical trial. *Health Psychology*, 25(1), pp.91–101. Available at:

http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2006-01035-012&site=ehost-live%5Cngeoffrey_williams@urmc.rochester.edu%5Cnhttp://psycnet.apa.org/journals/hea/25/1/91.pdf.

Wittchen, H.U. et al., 2011. Smoking cessation in primary care – a randomized controlled trial of bupropione, nicotine replacements, CBT and a minimal intervention. *International Journal of Methods in Psychiatric Research*, 20(1), pp.28–39.

Wood, R. et al., 2006. Measuring inequalities in health: the case for healthy life expectancy. *Journal of Epidemiology & Community Health*, 60(12), pp.1089–92. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2465513&tool=pmcentrez&rendertype=abstract [Accessed May 21, 2013].

World Health Organization, 2015. Disease and injury country estimates - United Kingdom. Available at: http://www.who.int/healthinfo/global_burden_disease/estimates_country/en/ [Accessed July 20, 2015].