Framework for cross-sectoral economic evaluation of public health interventions

A case study on a brief alcohol intervention

Francesco Ramponi

Ph.D. Thesis

University of York
Health Sciences

April 2019

Abstract

Public health interventions are often characterised by costs and outcomes across multiple sectors, with benefits that might only be realised far in the future. Further, such interventions frequently aim to address health equity concerns. For these reasons, the economic evaluation of public health interventions raises methodological and technical challenges. The lack of a consistent framework to evaluate their cost-effectiveness has been emphasised in the literature.

This thesis has two main objectives. First, to develop an analytical framework for the economic evaluation of public health interventions. The second aim is to show how to operationalise the proposed framework, using a brief alcohol intervention to reduce alcohol consumption among criminal offenders as a case study.

The proposed framework extends cost-effectiveness analysis (CEA) methods which are widely used for the evaluation of health care interventions, and consists of a cross-sectoral analysis, with the potential incorporation of health equity concerns and cross-temporal impacts.

The case study intervention impacts Health Care (HC) and Criminal Justice System (CJS). Conclusions and recommendations differ according to the perspective adopted for the evaluation. Analyses provide different results when conducted from the following perspectives: naïve HC (i.e. considering exclusively health-related costs and health effects on offenders); CJS; full HC (i.e. including also spill-over effects from criminal justice on victims' health and additional costs falling on the HC budget to treat victims); and joint HC and CJS. Alternative value judgments and equity considerations incorporated in the economic evaluation also affect the results of the analysis.

The proposed framework can provide support to decision making for local authorities, and could be potentially employed for the economic evaluation of all public health interventions. Results of the case study demonstrate the important implications of this work for future evaluations.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my supervisor Gerry Richardson and co-supervisor Mona Kanaan for their expert guidance and support throughout my time as a Ph.D. student. Special thanks go to Simon Walker, a precious addition to my thesis advisory panel (TAP), for providing expert advice and for his invaluable comments. I would like to thank also the other members of my TAP Maria Horne and Steve Parrott for their time and feedback.

An immense thank you is due to my friends in Trieste, York and elsewhere, for their encouragement and emotional support especially in the most difficult days of my Ph.D. journey. I am grateful to my fellow Ph.D. students and research fellows in the Centre for Health Economics and the Department of Economics & Related Studies. I would like to thank in particular those with whom I shared my office during these years for being not only amazing friends, but also source of stimulating discussions and advice.

Thanks to María José Aragón, James Gaughan, Vigyan, Dina Jankovic, Christopher Lübker, Matthew Robson and David Glynn for proofreading my thesis.

I dedicate this thesis to my family: Mom, Dad, Anna and Giacomo. This and any other achievement of mine would have not be possible without their love and support.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Parts of this work have been presented and discussed in conferences. A working paper titled "*Theoretical framework for a cross-sectoral and cross-temporal economic evaluation of public health interventions*" has been discussed at the 2017 Health Economists' Study Group (HESG) Summer Meeting (University of Aberdeen). The working paper "*Cross-sectoral economic evaluation of public health interventions: a case study on alcohol consumption*" has been discussed at the 2018 HESG Summer Meeting (University of Bristol) and presented at the 2018 European Health Economics Association (EuHEA) Conference (Maastricht). A third working paper titled "*Incorporating health equity concerns in the cross-sectoral cost-effectiveness analysis of public health interventions: a case study on alcohol consumption*" has been discussed at the 2018 EuHEA PhD Student-Supervisor and Early Career Researcher Conference (University of Catania).

The main findings of the case study developed in this thesis have been included in the Public Health Research Consortium (PHRC) report *"The economic evaluation of public health programmes with costs and effects falling outside the NHS on multiple budget holders"* (available at: http://phrc.lshtm.ac.uk).

This PhD project was funded and supported by the White Rose network and the Health Economics and Outcome Measurement (HEOM) theme of National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care, Yorkshire and Humber (CLAHRC YH) programme.

Chapter 1: Introduction

1.1. Overview

Public health interventions aim to promote health, prevent ill health and reduce health inequalities. Such interventions are not focused at the individual, but at a group level or above, potentially affecting the whole population (NICE, 2015, 2012). Resources available to fund health interventions are scarce, and the role of economic evaluation is to assist decision makers in their decisions about which intervention(s) to choose (Drummond, M. et al., 2015). The complex nature of public health interventions raises methodological and technical challenges in their evaluation (Weatherly et al., 2009).

Economic evaluations of health care interventions are typically conducted from the health (and social) care perspective and are only focused on health care costs and health consequences. Such a narrow perspective potentially underestimates the full impact of a public health intervention. It may be necessary to assess a broader range of non-health costs and effects falling on other sectors (e.g. education, criminal justice, employment and environment). Further, the impacts may be far in the future. Moreover, in this context, additional concerns about health equity might become critically important (Weatherly, Cookson & Drummond, 2014, Weatherly et al., 2009, Chalkidou et al., 2008, Drummond, M. et al., 2007).

The omission of relevant cross-sectoral and long-term impacts might cause imprecise and misleading results, and lead to inconsistent decision making against decision makers' aims. The inclusion of concerns about the distribution of health (and other outcomes) across the population and health inequalities might also impact the preferred intervention.

1.2. Research questions

Within this context, the main aim of this thesis is to define appropriate methods for the economic evaluation of public health interventions with impacts on multiple sectors.

More pragmatically, the objective is to develop and implement an analytical framework that extends cost-effectiveness analysis (CEA) methods which are widely used for the evaluation of health care interventions. The proposed framework could consider costs and outcomes across sectors, long-term costs and effects, and health inequity concerns in the analysis. Such framework can provide support to decision making for local authorities, and could be potentially employed for the economic evaluation of all public health interventions.

To demonstrate the important implications of this work for future evaluations, the developed framework of analysis will be explored using a brief intervention for offenders with alcohol related problems as a case study. The case study will be focused on addressing the cross-sectoral and health inequity aspects of the analysis. Nonetheless, the proposed framework could theoretically accommodate also the inclusion of cross-temporal aspects, and this aspect will be discussed as a potential extension of the evaluation.

1.3. Structure of the thesis

This thesis consists of three parts. The first part (Chapters 2 and 3) aims at addressing the research questions from a more theoretical point of view and is focused on the development of the framework. The second part (Chapters 4, 5, 6 and 7) of the thesis is based on a case study. Lastly, in Chapter 8, theoretical approach, findings and limitations of the case study are discussed and conclusions are drawn; recommendations for further research are also provided. The next three sub-sections describe each part of the thesis in further detail.

1.3.1. Theoretical part

After this first introductory chapter, the second chapter explores the theoretical framework of the economic evaluation of health care interventions. In the first part of Chapter 2, theoretical underpinnings and analytical methods typically employed for the assessment of health care interventions are reviewed and categorised. The final output is a taxonomy which aims at clarifying the characteristics and differences of the various potential alternative evaluation choices. The second part of Chapter 2 describes the nature of public health interventions and illustrates the theoretical and practical challenges in their analysis. Limitations of the traditional techniques are highlighted, and alternative potential methods of analysis to overcome these issues are briefly reviewed. At the end of the chapter, the main characteristics of the proposed approach to the analysis are delineated.

In the third chapter a framework which extends a 'traditional' CEA to include an evaluation across sectors, an inequity analysis, and an assessment of cross-temporal impacts is proposed. To capture cross-sectoral impacts, outcomes falling on different sectors are assessed and trade-offs among them are estimated. A compensation scheme across decision makers in different sectors in order to address resource allocation problems – generated when costs fall on a sector and benefits on another – is proposed. The cross-sectoral analysis is then extended in order to incorporate health inequity considerations in the evaluation. The objective is to assess impacts on health inequities across population subgroups, and alternative solutions for aggregating outcomes across individuals and dimensions are investigated. To reflect intertemporal issues, extrapolation techniques could be employed to estimate long-term costs and consequences of the intervention. Moreover, the impact of budgetary policies could be assessed.

1.3.2. Case study

From Chapter 4 to Chapter 7, a brief alcohol intervention to reduce alcohol consumption among offenders is used as a vehicle to illustrate how to operationalise the proposed framework. Even though the scope of the theoretical framework is potentially broader, the focus of the case study is on the operationalisation of the cross-sectoral and health equity analyses.

Each chapter covers one step of the analysis. In Chapter 4, the cost-effectiveness of the intervention is investigated from a 'naïve' health care system perspective, with decision makers solely focused on costs falling on the health and social care budget, and effects on trial participants. A 'traditional' CEA is thus conducted to evaluate health-related costs and health effects only on the offenders.

In Chapter 5, a CEA is carried out from the criminal justice system perspective (i.e. with decision makers solely focused on criminal justice). Costs falling on the budget of the criminal justice system are estimated, and reductions in reconviction frequency are used as a measure of effectiveness.

Chapter 6 builds on the evaluations conducted in Chapters 4 and 5. In the first part of Chapter 6, the CEA conducted from the health care perspective is revised and spill-over effects from criminal justice on victims are included in the evaluation. A 'full' health care system perspective is therefore adopted, with decision makers still solely focused on health. In the second part of the chapter, another CEA is conducted from a broader and integrated perspective, where decision makers are focused on both health and criminal justice.

Chapter 7 investigates the effects of incorporating health equity concerns in the cross-sectoral analysis.

1.3.3. Discussion and conclusions

Chapter 8 briefly summaries objectives and results of the previous chapters, and highlights the contributions of this thesis to the literature. Further, it covers the discussion of both the theoretical analytical framework and the analysis of the case study. In this chapter, the potential cross-temporal extension of the evaluation is discussed. Moreover, policy implications and recommendations for future research are pointed out, and conclusions of the work are drawn.

Chapter 2: Theoretical background

2.1. Introduction to the economic evaluation of health care interventions

Economic evaluation can be defined as the comparative analysis of alternative courses of action in terms of both their costs and consequences (Drummond, M. et al., 2015, p. 4). The two main features of economic evaluation are thus highlighted: firstly, it deals with both inputs (costs) and outputs (consequences) related to an action; secondly, economic evaluation concerns itself with choices.

Because resources are limited, it is not possible to produce all desired outputs and it is necessary to undertake decisions regarding which course(s) of action to follow. The purpose of economic evaluation is to inform such decisions by providing judgments and comparisons of alternative states of the world. As a consequence, it is necessary to define which are the theoretical normative principles that support these judgements, on the basis of agreed assumptions about what constitutes values (Drummond, M. et al., 2015).

The ‘underlying normative principles’ are also known as ‘theoretical underpinnings’, ‘normative viewpoint’, ‘theoretical approach’ or ‘underlying methodology’. All these terms emphasise the value judgments (i.e. the ‘norms’) on which the economic evaluation is grounded. In the health care context, different theoretical approaches can be adopted, based on either the concepts of the welfarist approach (Tsuchiya & Williams, 2001, Hurley, 2000), or on alternative value judgments (Culyer, 2012, Brouwer et al., 2008, Claxton, Sculpher & Culyer, 2007).

Another aspect of the economic evaluation regards the analytical technique employed. Alternative analytical techniques are available, and the choice of employing one or another is typically related to the theoretical approach which underpins the analysis.

There is no right or wrong decision in the choice of theoretical approach and analytical technique. Nevertheless, methodological and theoretical aspects in the analysis should be justified by an informed choice. Analysts and decision makers should be aware of the underlying value assumptions and potential limitations of the results. This is also necessary because results of analyses based on alternative normative principles and analytical techniques might diverge and therefore potentially yield different resource allocations and recommendations (Buchanan & Wordsworth, 2015, Bala, Zarkin & Mauskopf, 2002).

Awareness about the value assumptions behind the economic evaluation and methods of analysis should be of paramount importance. Nevertheless, not only are these aspects frequently left implicit, but also in the literature there is inconsistency in the definition of the theoretical approaches, and in the use of the terminology. A review of the literature was therefore undertaken in order to identify and describe the alternative theoretical underpinnings and analytical methods currently employed in the UK for the economic evaluation of health care interventions. On the basis of these findings, a taxonomy which categorises the normative principles and analytical techniques was developed.

2.1.1. Structure of the chapter

Section 2.2 classifies and describes characteristics, potentialities and limitations of the different theoretical approaches to the evaluation of health care interventions typically employed in the UK. Sub-section 2.2.1 illustrates the welfarist approach, sub-sections 2.2.2 and 2.2.3 explore the extra-welfarist and social decision making approaches, respectively. At the end of the Section 2.2, a taxonomy of the normative principles is provided.

In Section 2.3, the analytical techniques (i.e. the methods used to practically implement the economic analysis) are discussed. The link between normative principles and analytical methods is also investigated. A table which classifies and summarises the analytical methods currently employed for the economic evaluation of health care interventions is provided at the end of the section.

Lastly, Section 2.4 explores more specifically the evaluation of public health interventions. Challenges in the economic evaluation of public health interventions and the appropriateness of employing the more traditional analytical techniques are investigated. At the end of Section 2.4, the proposed approach for the evaluation is briefly introduced.

2.2. Normative principles

2.2.1. Welfarist approach to economic evaluation

Welfare economics is a normative framework used to judge alternative states of the world (Drummond, M. et al., 2015). The welfarist approach to welfare economics – or equivalently 'welfarist economics', as defined by Brouwer et al. (2008) – argues that individuals are the best judges for themselves and act rationally in order to maximise their welfare, which is based on outcomes only, and not on processes. In summary, welfarist economics is based on the following four tenets (Hurley, 2000, p. 60):

1) Utility maximisation. In economics the term has tended to be synonymous with preference; the more preferable an outcome, the more utility associated with it. Utility maximisation principle implies that individuals are rational, rank the available options and choose with consistency the most preferred among them (i.e. the option associated with the highest utility);
2) Individual sovereignty. Individuals are the best judges of their own preferences;
3) Consequentialism. What matter are the resulting effects and outcomes of an action or a choice. The process is not relevant;
4) Welfarism. All the non-utility aspects of a situation are not considered in the judgment. The 'goodness' of a state is determined only by the utility levels (i.e. welfare) attained by individuals in that state (Drummond, M. et al., 2015, Hurley, 2000).

'Welfarism' and 'welfarist approach' are therefore not synonyms and they should not be used interchangeably. Welfarism regards only one aspect of welfarist economics and is the tenet which excludes all the non-utility aspects from the evaluation (Culyer, 2012, p. 67, Hurley, 2000). However, inconsistency in the use of the terminology can be found in the literature. For instance, according to the Encyclopedia of Decision Making (2009), 'welfarism' (and not 'welfarist economics') is *"often used to describe the dominant ethical framework used in welfare economics to judge states of the world with"* (Kattan & Cowen, 2009, p. 1189).

2.2.1.1. Allocation of scarce resources to maximise social welfare

In reality, the 'basic economic problem' is that unlimited 'wants' of the society are hindered by the limited resources available, and choices need to be made between alternative uses of resources (Drummond, M. et al., 2015). Therefore, welfarist economics aims to define their optimal allocation, which is defined as the one that maximizes the welfare of the society (Kattan & Cowen, 2009).

In the classical tradition of welfarist economics, utility measures were considered cardinal and interpersonally comparable: it was deemed possible to quantify the magnitude of increases or decreases in utility for the individuals, and make comparison among them. As a consequence, under these assumptions, the overall social welfare can be defined as a function of the levels of utility attained by members of the society (Kattan & Cowen, 2009, Brouwer et al., 2008, Hurley, 2000).

The modern welfarist economics theory is instead characterised by the abandon of the conception of cardinal utilities, in favour of ordinal utilities, and the rejection of intra-individual comparability of welfare gains and losses (Brouwer et al., 2008). In other words, the only information available is that the utility of an individual increases or decreases (Kattan & Cowen, 2009).

2.2.1.2. Pareto and potential Pareto principles

In the strictest interpretation of welfarist economics theories, changes in ordinal utilities are judged using the Pareto principle. According to the Pareto principle, a state is preferred to an alternative one if at least one individual is better off (i.e. has higher utility) and no one is worse off (i.e. experiences a decrease in utility) (Brouwer et al., 2008, Tsuchiya & Williams, 2001). An alternative principle to judge alternative states of the world is the potential Pareto criterion, which can be also called compensation test, or Kaldor-Hicks criterion (Hurley, 2000, p. 61). Potential Pareto principle states that if the gainers are hypothetically able to compensate the losers, and after the compensation is paid losers are at least as well off as originally, then there is an improvement for society (Tsuchiya & Williams, 2001).

When judgements are grounded in the potential Pareto principle, 'potential' compensations can be estimated and comparisons of situations where some individuals gain and other lose are allowed. In a perfectly competitive market, compensation can be based on prices, which represent the social value of alternative activities. In the presence of distortions (e.g. due to monopoly, taxation or externalities), observed prices do not reflect the social value, and 'shadow prices' (i.e. prices as if there was a competitive and undistorted market) can be estimated. In absence of markets, revealed preferences or hypothetical valuations can be exploited to estimate the willingness-to-pay (WTP) to reveal individuals' value (Claxton, Sculpher & Culyer, 2007). As explained previously, these compensations are only hypothetical, and do not actually have to be made (Tsuchiya & Williams, 2001).

2.2.1.3. Welfarist approach to health economic evaluation

The concepts and the framework of welfarist economics apply also to the health care context. Nevertheless, the literature shows a lack of consistency in the terminology used. To avoid misunderstandings, in this thesis the term 'welfarist approach' will be used to describe the approach to health economic evaluation based on the concepts of welfarist economics.

Following Hurley (2000), the main characteristic of the welfarist approach (as opposed to the extra-welfarist approach, described in the next section) is that value is assessed using utility. According to the welfarist approach to the economic evaluation of health care, the purpose of health care is to increase individual utility, and WTP can be used as a proxy for changes in utility. Nevertheless, it can be argued that WTP is affected by the initial income distribution, and consequently by the ability to pay. This is because the marginal utility of income declines as income increases, affecting compensation, offered and required (Drummond, M. et al., 2015, Claxton, Sculpher & Culyer, 2007). Moreover, even including adjustments to avoid the problems related to the ability to pay, the use of compensation can be considered incoherent in the economic evaluation of health care, in particular when dealing with interventions that may cause death (Broome, 1978). According to Broome (1978), compensation would be based on an ex-ante valuation (i.e. made before the distribution of its costs and benefits is exactly known), as a proxy of the unknown and correct ex-post valuation. When analysing a project causing death (which would require infinite compensation), ex-ante valuation would be different from ex-post valuation. More precisely, the former would have finite costs and the latter infinite ones. Therefore, it would be incorrect to take social decisions valuing health on the basis of ex-ante individual preferences measured through compensation.

Further, according to the welfarist approach, the 'evaluative space' (i.e. the relevant elements and components to be included in the evaluation) is confined to individual utilities only and there is no need for a decision maker to use their own values (Brouwer et al., 2008). In other words, the welfarist approach avoids the specification of an exogenous (i.e. determined by an external authority) objective function, because the outcomes are valued by individuals through the compensations. Equity concerns can be therefore introduced in the evaluation only if they are incorporated into the individuals' utility measure, and not if they reflect exogenous equity concerns (e.g. weights on outcomes imposed by an external authority) (Claxton, Sculpher & Culyer, 2007).

2.2.2. Extra-welfarist approach

Given the previously mentioned limitations of the welfarist approach when applied to the health care context, an alternative approach which tries to specify an objective function which is not solely based on individual utility has become a popular alternative, the extra-welfarist approach (Drummond, M. et al., 2015). According to the extra-welfarist approach, individuals are no more the unique source of valuation because what is relevant is something beyond the particular outcomes the individuals choose for themselves (Brazier et al., 2007). The determination of the content of the evaluative space (i.e. elements and components to be included or excluded) can be performed also by authoritative decision makers, experts, representative samples of the public or representatives of the affected individuals (Brouwer et al., 2008).

Sometimes the extra-welfarist approach is referred to as 'non-welfarist' approach (Drummond, M. et al., 2015, Culyer, 2012). This is to stress that this approach does not expand the evaluative space to include more ('extra') than only individual utility: it just uses a different evaluation space ('extra-ordinary' rather than 'extra'). Utility information can be replaced by other characteristics or capabilities, or complemented with other information, which can regard for example the quality of utility or equity weights (Brouwer et al., 2008).

In the extra-welfarist approach the outcome of interest becomes no more the welfarist concept of welfare based on utility only, but a less defined 'well-being' (Brouwer et al., 2008). In fact, 'welfare' *"refers to well-being as assessed specifically in utility terms; in contrast 'well-being' is used more generally and can be assessed in terms other than utility"* (Hurley, 2000, p. 59).

2.2.2.1. The origins of the extra-welfarist approach

According to Culyer (2012), Sen's theories (1993) - combined to previous concepts such as merit goods developed by Musgrave (1959) and basic goods by Tobin (1970) and Rawls (1971) - contributed crucially to the development of the extra-welfarist approach in the context of the economic evaluation, based on the centrality of health. In fact, it can be argued that the extra-welfarist approach builds on Sen's *"view of living as a combination of various 'doings and beings', with quality of life to be assessed in terms of the capability to achieve valuable functionings"*. In Sen's view, 'functionings' are defined as *"the various things that he or she manages to do or be in leading a life"* (Sen, 1993, p. 271).

In contrast with traditional evaluations based on commodities, income, or material resources, according to Sen, resources are just instruments to enhance people's well-being. Health in particular is a fundamental characteristic affecting people's capability to 'flourish' as human beings

(Culyer, 2012). In other words, well-being should be measured according to what individuals could in principle do (i.e. capabilities) and not what they actually do (i.e. functionings), and ill health reduces the ability to achieve functionings (Edwards, Charles & Lloyd-Williams, 2013).

2.2.2.2. Current implementation of the extra-welfarist principles

The perception of peculiarity and uniqueness of health in the economic literature fostered the spread of evaluations where health (and not welfare) is perceived as the main focus (Drummond, M. et al., 2015). According to the current interpretation of the extra-welfarist principles, health is pursued by policy makers for its own sake and not because it yields utility or merely to the extent that it yields utility (Brouwer et al., 2008). As a consequence, maximising health gains is considered the main objective of a health care programme, subject to a single exogenous budget constraint, which is given by the resources available for the health care system (Coast, Smith & Lorgelly, 2008). However, the extra-welfarist approach still requires the definition of a social objective function to be maximised. When the main goal of an intervention can be considered to be health, it is however reasonable to ignore the other aspects in the social objective function (Sculpher & Claxton, 2012).

The interpretation of the extra-welfarist approach and its implementation in the economic evaluation of health care is still discussed (Coast, Smith & Lorgelly, 2008, Birch & Donaldson, 2003). Health is typically measured in terms of length and quality of life, using the Quality-Adjusted Life Years (QALYs), which are valued equally across the population (Drummond, M. et al., 2015). It has been argued that the recommendations from the National Institute for Health and Care Excellence (NICE) played a fundamental role in establishing the common extra-welfarist practice of using QALYs as the standard measure of benefits (Coast, Smith & Lorgelly, 2008). Coast et al. (2008) pointed out that originally Sen's capability approach was actually characterised by a broader perspective, with a particular emphasis on equity and distributional concerns. Moreover, the implicit assumption of social well-being depending only on health maximisation has been criticised. Such a practice retains the idea of maximisation (no more of utility, but of health), which is still related to the welfarist approach (Coast, Smith & Lorgelly, 2008). Evidence also suggests that maximisation of health gain is not all society and experts are concerned about in relation to health care decision making. Policymakers and health care professionals involved in the decision making base their preferences also on factors such as severity of the disease, target age, and magnitude of the individual health benefit from the interventions (Baji et al., 2016).

Furthermore, for Birch and Donaldson (2003), the extra-welfarist approach acknowledges and accepts the risks of paternalism and dictatorship caused by imposing decision makers' preferences on individuals. The two authors argued that actually in Sen's theory the source of valuation was still the individual, on the basis of its set of functionings and capabilities. Moreover, they claimed that welfarist economics theoretically could be able to incorporate all the concerns raised by proponents of extra-welfarist approach, by adapting the concept of utility (Birch & Donaldson, 2003). Nevertheless, even if theoretically the welfarist approach might be broadened, the feasibility of such adaptations (e.g. integrating equity aspects and evaluation of intangible social goods in a welfarist context analysis on individual utilities) is not taken into consideration. Furthermore, Broome's (1978) argument about the compensation of risks rather than of health itself would remain still unsolved.

2.2.2.3. Issues in the specification of a social objective function

Economic evaluations based on extra-welfarist principles still have the objective of maximising a specified social objective (based on well-being) function. The extra-welfarist approach is thus still prescriptive, and one of the main issues in its implementation regards the difficulties in defining explicitly a socially legitimate objective function to be maximised (Sculpher & Claxton, 2012).

In reality, any specification of a social objective function would be controversial. Given the many conflicting and contradictory claims on what is socially valuable, each stakeholder would have its own subjective view with regard to the factors that must be taken into consideration and how to value them (Walker et al., 2019). The issues about reaching a general consensus about a social objective function were discussed by Arrow (1950) and Sen (1970) in their impossibility theorems.

Because both welfarist and extra-welfarist approaches aim to define social value using an explicit social objective function (based either on welfare or wellbeing), the difficulty in achieving a consensus among decision makers about a complete objective function is a common issue for both approaches (Walker et al., 2019).

For this reason, in the UK, the influence of the extra-welfarist theories on the more traditional welfarist concept of an individualistic social welfare, together with the recognition of the practical limitations of the extra-welfarist approach, led to the development of the more pragmatic social decision making approach (Hurley, 2000).

2.2.3. Social decision making approach

The social decision making approach has the purpose of informing social decisions rather than making claims about social welfare. Compared to welfarist and extra-welfarist approaches, the main difference is that social decision making approach avoids investigating the social values, and therefore does not aim at specifying the structure (i.e. the arguments) of a social objective function.

According to the social decision making approach, the specification of a complete, explicit and legitimate expression of social value is not possible (Drummond, M. et al., 2015, Paulden & Claxton, 2012, Williams, 1981, Sugden & Williams, 1978).

According to this approach, decision making bodies (e.g. government departments) can be seen as the agent of a socially legitimate higher authority, which cannot express an explicit and complete social welfare function, but allocates resources and gives the agent a responsibility to pursue explicit objectives (e.g. to improve health). In these circumstances the delegated authority cannot be asked to improve social welfare, since it cannot be specified (Claxton et al., 2011, Claxton, Sculpher & Culyer, 2007). The role of social decision making is therefore more humble than welfarist and extra-welfarist approaches: this approach can be used only to inform decisions rather than prescribe them (Drummond, M. et al., 2015, p. 116).

The authority of the decision making bodies is legitimated by the socio-democratic process, which is how society tries to balance conflicting and contradictory claims on resources and social objectives (Drummond, M. et al., 2015, Paulden & Claxton, 2012). The allocation of budgets available for each decision maker can be viewed as an expression of social preferences (Sculpher & Claxton, 2012) and decision makers are informed about the characteristics and consequences (in terms of costs and benefits) of alternative courses of action.

2.2.3.1. Implementation of social decision making approach in the health care context

Social decision making can help facilitate decisions in the absence of consensus about an explicitly defined social objective function. Health may be considered one of the arguments of an unspecified underlying social objective function. In the UK, the Department of Health and Social Care (DHSC) is given the responsibility of pursuing health maximisation, following the recommendations of NICE (Drummond, M. et al., 2015, Sculpher & Claxton, 2012, Claxton et al., 2011).

In the absence of consensus on principles, 'accountability for reasonableness' becomes fundamental in order to agree on what is legitimate and fair (Daniels, 2000). Daniels and Sabin (1998) argued that, to assure procedural fairness, a decision process must involve: transparency about the grounds for decisions; agreement of stakeholders regarding the relevance of reasons and rationales; existence of procedures for revising decisions in light of new evidence and arguments. The UK health system is now explicitly applying this framework of accountability in its decision making process, and accountability for reasonableness helped shape thinking about how NICE should incorporate social value judgments into its recommendations to the National Health Service (NHS) in England and Wales (Daniels & Sabin, 2008).

2.2.3.2. The crucial role of opportunity costs

When alternative mutually exclusive courses of action are available, the decision rule to allocate resources to a specific intervention is based on the comparison of the resources necessary to fund the intervention and their opportunity costs, which are the impacts associated with what could have been done with the best alternative use of the same resources (Sculpher, Claxton & Pearson, 2017). In other words, opportunity costs are represented by what is forgone in order to accommodate the resources to provide new services (Sculpher et al., 2014).

There is often no explicit consideration of which services might need to be displaced to generate the funds to pay for the new intervention (i.e. which interventions should be removed, delayed, or downscaled to generate the necessary funding). An estimate of the system's marginal productivity (i.e. the relationship between changes in over-all health system expenditure and changes in the relevant measure of benefit) can be therefore used to measure of what is given up as financial resources are drawn away from other services (Sculpher, Claxton & Pearson, 2017).

Estimates of health opportunity costs using the estimated relationship between changes in health expenditure and changes in health outcomes for health care in UK have been recently generated (Claxton, Martin, et al., 2015). This provides an empirical estimate of the marginal productivity of the NHS (i.e. the health impacts of small changes in spending across the NHS budget) (Drummond, M. et al., 2015).

2.2.4. Taxonomy of the normative principles

Table 2.1 summarises the aforementioned approaches, and provides a taxonomy in terms of the relevant outcomes taken into considerations, sources of validation, the way of determining the social objective function and the criteria to inform judgements. A brief description of the implementation of each approach in the health care context is also included.

	Welfarist approach	Extra-welfarist approach	Social decision making approach
Description	Based on the tenets of welfarist economics: - **Utility maximisation;** - **Individual sovereignty;** - **Consequentialism;** - **Welfarism.**	Grounded on Sen's theories, extra-welfarist approach broadens, or narrows, the **evaluative space** to include outcomes other than individual utility. A '**well-being**' concept (which may include **utility information as well as other extra measures and indicators**) replaces the welfare based on utility only.	More **pragmatic approach** where the principal (socially legitimate higher authority) devolves responsibility and resources to an **agent** (decision making body), who **pursues explicit objectives.**
Relevant outcome(s)	Only **utility.** Centrality of **welfare** (defined as well-being measured in utility terms).	Utility information can be **replaced by other outcomes, or complemented with other information**, which can regard for example the quality of utility or equity weights. The determination of the specific relevant outcomes depends on the source of validation in the specific context. Usually health is the main focus.	Depends on the **objective** related to the decision maker authority.
Source of valuation	Affected **individuals.**	The determination of the content of the evaluative space can be performed by the individuals, but also by **authoritative decision makers, experts, representative samples of the public or representatives of the affected individuals**.	The authority of the decision making bodies is **legitimated by the democratic political process.** **The allocation of budgets available to** decision makers is expression of social preferences.
Social objective function	The **social welfare function** is a function of individual welfares.	**Social well-being function** where outcomes can be weighted in order to incorporate equity and other considerations.	**No explicit, complete and legitimate social objective function** is fully specified.

Criteria to inform judgments	Optimal allocation of resources to **maximise social welfare.** Judgments usually based on the **potential Pareto principle**, which allows for hypothetical compensations. This requires the assumption of feasibility of benefits monetarisation.	**Maximisation of social well-being.**	**Decision makers are informed** about the characteristics and consequences (in terms of costs and benefits) of alternative courses of action. Adoption or rejection decisions are related to **opportunity costs** (i.e. impacts associated with what could have been done with the best alternative use of the same resources).
Implementation in the health care context. Weaknesses and criticisms	**WTP** and compensations (offered and required) can be affected by the ability to pay and income. Even including adjustments regarding the ability to pay, the **use of compensation can be considered incoherent** when dealing with health. Confining the evaluative space to individual utilities only could be a limit for the evaluation: for example, **equity concerns imposed by an external decision maker cannot be included.**	Current interpretation is based **on health** (measured in QALYs) **maximisation**, subject to a single exogenous **budget constraint**, given by the resources available for the health care system. Criticisms to this interpretation relate to the **narrow perspective focusing only on health maximisation** (as opposed to the broader capability approach by Sen), with **no emphasis on equity and distributional concerns.** Risks of **paternalism and dictatorship** caused by imposing others' preferences on individuals. The main issue in the implementation regards the **difficulties in defining explicitly a socially legitimate objective function** to be maximised. Impossibility theorems demonstrated that a general consensus about such function cannot be reached.	More humble approach: objective function is not defined, and the role is to **inform decisions** rather than making prescriptive claims about social welfare or well-being. In the UK, decision power is assigned to the DHSC, which is given the responsibility by the whole social democratic process to pursue health maximisation (typically measured using QALYs).
Key references	(Drummond, M. et al., 2015, Brouwer et al., 2008, Tsuchiya & Williams, 2001, Hurley, 2000, Broome, 1978)	(Drummond, M. et al., 2015, Brouwer et al., 2008, Coast, Smith & Lorgelly, 2008, Birch & Donaldson, 2003, Sen, 1993, 1970, Arrow, 1950)	(Sculpher, Claxton & Pearson, 2017, Drummond, M. et al., 2015, Culyer, 2012, Paulden & Claxton, 2012, Sculpher & Claxton, 2012, Williams, 1981, Sugden & Williams, 1978)

Table 2.1 Normative principles for the economic evaluation of health care interventions

2.3. Analytical techniques

Section 2.2 illustrated how judgements regarding alternative states or courses of action can be underpinned by different normative principles. In this section, the analytical methods commonly employed to perform the analysis are described and classified. Because analytical techniques usually depend also on the specific approach chosen for the evaluation, the relationship between the techniques employed for the evaluation and the theoretical approach is also investigated.

All the methods of full economic evaluation involve the assessment of both costs and outcomes of the interventions to be compared. While the estimation of costs is common to all of them, the feature that distinguishes alternative techniques is how the consequences of health care programmes are valued. In the UK, there are three main methods for the economic evaluation of health care interventions: Cost-Consequence Analysis (CCA),[1] Cost-Benefit Analysis (CBA), and Cost-Effectiveness Analysis (CEA). As explained later in Sub-section 2.3.3, in this thesis the CEA category includes also CEA that measures consequences using QALYs, even if frequently referred to as Cost-Utility Analysis (CUA)[2] (Drummond, M. et al., 2015). Another potential analytical technique is the Cost-Minimisation Analysis (CMA); nevertheless, this is not a full economic evaluation. In fact, the use of CMA is sometimes suggested when the consequences of the alternatives under consideration are judged broadly equivalent, and the difference between the alternatives boils down to a simple comparison of costs (Drummond, M. et al., 2015). Moreover, it has been argued that the use of CMA cannot be determined in advance because of the uncertainty around costs and consequences (Briggs & O'Brien, 2001).

2.3.1. Cost-consequence analysis

CCA is an analytical method based on the scepticism about the possibility of condensing the result of an economic evaluation into one number, such as the net benefit or the cost per outcome. CCA does not rely on any assumption about what, how and by whom costs and consequences must be valued. No outcome aggregation is undertaken and results are shown in the form of a table (Brazier et al., 2007, p. 14, Coast, 2004).

---

[1] Due to its limitations in informing decision making (explained in Sub-section 2.3.1), some might argue that CCA does not meet the full criteria to be classified as a method of economic evaluation for making decisions.
[2] QALYs are here considered a preference based measure of health that do not necessarily reflect utility.

CCA might be easier to understand for the decision makers with respect to other types of analysis (Coast, 2004). Nevertheless, because of the absence of synthesis of costs and outcomes, policy makers' weights and criteria of decisions are left implicit (Kelly et al., 2005). The weakness of CCA is therefore its potential lack transparency and consistency across decisions (Drummond, M. et al., 2015).

As CCA is just a description of costs and consequences and is theoretically not dependent on any value assumption,[3] it could be used theoretically in all analyses, independently of the underlying normative principle of the evaluation. Nonetheless, the main drawback of the CCA regards its limits in informing decision makers, because it does not provide any guidance as to how the different outcomes should be weighed against each other (Greco, Lorgelly & Yamabhai, 2016). In fact, probably the best use of CCA is as a preliminary technique to describe interventions systematically, and it is particularly useful when dealing with a broad range of costs and benefits (Weatherly et al., 2009, Kelly et al., 2005).

2.3.2. Cost-benefit analysis

If the normative viewpoint of the analysis is the welfarist approach, CBA is the usual method implemented (Brazier et al., 2007, Gafni, 2006, Bala, Zarkin & Mauskopf, 2002). In the CBA, both costs and health outcomes are valued in monetary terms and become therefore comparable. Usual techniques to value the monetary equivalent of health outcomes are the human capital approach and the estimation of the WTP for the non-marketed goods, through the elicitation of stated preferences in an hypothetical contingent market (Brazier et al., 2007).

CBA aims to address questions of allocative efficiency, such as whether a specific goal is worth achieving, or what should be the amount of resources to be allocated in order to achieve the goal (Donaldson, 1998). Due to its theoretical foundation in welfarist economics and in the use of Kaldor-Hicks criterion as a hypothetical compensation test, CBA holds significant conceptual appeal for economists: it can potentially allow decision makers to assess health and non-health outcomes, and perform comparisons between programmes in all sectors of economy (Grosse, Wordsworth & Payne, 2008, Gafni, 2006).

---

[3] Pragmatic assumptions are however still necessary, for example, regarding the choice of which outcomes and costs to include in the analysis.

The fact that CBA is underpinned by the principles of welfarist economics enables such prescriptive power for CBA and allows converting and describing everything in commensurable terms. However, this can also be regarded as the main limitation and challenge of this method (Culyer & Chalkidou, 2019).

In practice, the implementation of CBA can be difficult (Culyer & Chalkidou, 2019, Donaldson, 1998). CBAs have usually large data requirements, such as estimates of all the relevant costs, and surveys to estimate WTP values (NICE, 2012). Moreover, traditionally, there have been measurement issues concerning how health and non-health impacts can monetarised (see Sub-section 2.2.1) (Culyer & Chalkidou, 2019). Furthermore, the use itself of WTP can be perceived in contrast with NICE's equity objectives: individual WTP is a measure of demand rather than of need, whereas in the NHS, health care is allocated according to need (NICE, 2012, Weatherly et al., 2009). Nevertheless, even assuming the correctness of the monetarisation of health, there would still be a risk of limiting the included costs and benefits to those that are measurable. For example, the value of equity and social cohesion may not be easily quantified in a CBA performed with a welfarist approach (NICE, 2012, Brazier et al., 2007, Kelly et al., 2005).

2.3.3. Cost-effectiveness analysis

The limitations of CBA's welfarist principles in the health care context (Culyer & Chalkidou, 2019) contributed to the adoption of alternative methods for the economic evaluation of health care programmes. In the UK, and in many other countries, the use of CEA has generally prevailed (Drummond, M. et al., 2015).

CEA avoids the monetarisation of health because effects are measured in 'natural' units (e.g. number of life-years saved or deaths prevented). Results are typically presented in terms of incremental cost per unit of effect. Such estimates are called Incremental Cost-Effectiveness Ratios (ICERs), and can be combined with a Cost-Effectiveness (CE) threshold to inform whether a new policy should be undertaken or not (Brazier et al., 2007). By incorporating the threshold value in the results of the analysis, results of the CEA can be also expressed using Incremental Net Benefits (INBs). INBs can be presented in monetary or natural units: Incremental Net Monetary Benefits (INMBs) and Incremental Net Health Benefits (INHBs), respectively (Drummond, M. et al., 2015).

The nature and the value of the threshold has been widely discussed in the literature (Drummond, M. et al., 2015). The perspective that the threshold should reflects health opportunity cost (i.e. the health that could have been generated if the money required to fund the intervention had been spent on something else) has come to be known as a 'supply side' approach. In contrast, according to the ‘demand side’ approach, the threshold should reflect the consumption value of health and empirical research aim to represent societal WTP for additional health gains (i.e. what individuals are willing to forego in non-healthcare/private consumption for gains in healthcare) (Thokala et al., 2018). It has been argued that CE thresholds used across the world might be considered overestimates and are based on historical estimates, heuristics or judgements. Empirical estimates of the supply side threshold – based on the opportunity costs of an investment, measured as the health that could have been generated if the money required to fund it had been spent on something else – could be considered more appropriate for judging the cost-effectiveness of new technologies if the aim is to maximise population health in a budget constrained health care system. In the UK, the supply-side threshold has been found to be much lower than the threshold typically recommended by NICE (Sculpher, Claxton & Pearson, 2017, Claxton, Martin, et al., 2015).

2.3.3.1. Utilities, QALYs and cost-utility analysis

CEA allows evaluating only one outcome for each analysis. CEA is in fact particularly suitable for decision making when no trade-off between outcomes is required. This can be a limitation for decision making purposes when multiple outcomes must be assessed (Brazier et al., 2007). For this reason, in the past it has been recognised that it would have been extremely useful to describe the total improvement in health using a single numerical measure of effectiveness, and a variant of the CEA has been introduced: the CUA (Torrance & Feeny, 1989).

In the CUA the multidimensional health change is condensed into a single number called ‘utility’. QALYs are usually the generic outcome chosen to describe the health consequences of the treatment to be evaluated, by incorporating both quantity and quality of life. Since the consequences of the intervention are captured by a generic measure of health gain, CUA provides better comparability across both different health outcomes and interventions (Drummond, M. et al., 2015, Brazier et al., 2007).

Some confusion arises in the literature about the definition of economic analysis employing QALYs: these are referred to as both CEAs and CUAs. In the usual practice, the terms ‘utilities’ and ‘QALYs’ have been frequently used interchangeably, even though originally Torrance and Feeny (1989) were already clear in distinguishing conceptually these two ways of addressing the objective of using a unique measure of the health consequences.

In general terms, utilities are preferences that an individual or society may have with respect to any particular set of health outcomes. Nevertheless, in the economic literature, the term ‘utility’ has a very precise meaning and is associated with the theories on decision making under uncertainty developed by von Neumann and Morgenstern (vNM). Since QALYs are based on preferences on health states, it has long been debated whether QALYs were a measure of utility or not. It has been concluded that QALYs can be utilities only when formed from preferences measured with standard gambles. Nevertheless, this is only one possible technique to construct them, and even in this case, restrictive and unrealistic assumptions must hold. As a consequence, QALYs are generally considered not consistent with the vNM concept of utility (Drummond, M. et al., 2015).

For this reason, in this thesis I will use the term ‘preferences’ instead of utility, and I will avoid the use of the label ‘CUA’. Since QALYs are not an expression of utility, analyses employing QALYs will be referred to as CEAs, as suggested by Drummond, M. et al. (2015). The term ‘CEA’ will include also CEAs that use QALYs, even if in the literature these are frequently referred to as CUAs.

2.3.3.2. Use of CEA and its theoretical underpinnings

Because the welfarist approach requires a measurement of utility, CEA is not justifiable on strictly welfarist underpinnings. As previously highlighted, health (and not utility) is the main focus of the analysis, and it is assumed that decision makers aim to maximise health subject to resource constraints. CEA techniques are usually chosen when the analysis is based on extra-welfarist or social decision making principles (Grosse, Wordsworth & Payne, 2008, Brazier et al., 2007, Bala, Zarkin & Mauskopf, 2002, Dolan & Edlin, 2002).

When the analysis is based on extra-welfarist or social decision making approach principles, whether the QALYs can or cannot describe utility is not crucial for the purpose of the analysis. QALYs are usually judged a good approximation of health gains and that is necessary and sufficient in terms of decision making. QALYs are a generic measure of health gain and they offer the potential to compare programmes in different areas of health. No alternative method for the measurement of health gains showed significant improvement in the quality of the analysis or facilitations in the calculation of the effects (Drummond, M. et al., 2015).

The role of opportunity costs is fundamental. Recommendations of a CEA depend crucially on the explicit comparison of its results with a threshold that should represent the alternative use of the resources if this policy were not implemented (Sculpher, Claxton & Pearson, 2017, Sculpher & Claxton, 2012). Actually, even following a more welfarist approach (e.g. assuming that health is measurable in monetary terms), estimation of opportunity costs would be crucial to answer the question whether the value of the health gained exceeds the value of the health forgone (Sculpher & Claxton, 2012).

With regard to the policy implications of the evaluation, economic analysis based on CBA allows making statements about social welfare. Nevertheless, that requires the specification of an explicit social welfare function based on the aggregation of peoples' willingness to pay. In contrast, CEA can be operationalised even in absence of an explicit social objective function , but can be used only for questions of technical or operational efficiency (Claxton et al., 2011). As a consequence, role of CEA is more modest, claiming to inform social decisions in health rather than prescribing social choice (Culyer, 2012).

2.3.4. Taxonomy of the analytical techniques

Table 2.2 summarises the aforementioned analytical techniques, and provides a taxonomy in terms of the outcome measurement and the associated normative principles. A brief description of advantages and disadvantages of each method is also reported. In Table 2.2 the term 'CEA' encompasses also CEAs that use QALYs, even if in the literature these are frequently referred to as CUAs.

	Cost-consequence analysis	Cost-benefit analysis	Cost-effectiveness analysis
Outcome measurement	Natural units. Multiple (including non-health) outcomes.	Monetary units.	Natural units or a single numerical measure of effectiveness (usually QALYs for health).
Normative principles	Theoretically no underlying value assumptions.	Usually welfarist approach.	Usually extra-welfarist approach or decision making approach.
Advantages	Might be easier to understand. Useful tool for a preliminary description of interventions, especially when the range of costs and benefits is broad.	Able to address questions of allocative efficiency. Can potentially allow decision makers to assess health and non-health outcomes, and perform comparisons between programmes in all sectors of economy.	Can be operationalised even in absence of an explicit social objective function. Aims at informing social decisions rather than prescribing social choice.
Disadvantages	Limits in informing decision makers because of the absence of synthesis of costs and outcomes. Potential lack transparency and consistency across decisions.	In practice, the implementation can be difficult due to data requirements and issues in the monetarisation of health and non-health impacts. WTP and compensations can be affected by income and ability to pay. Value assumptions on which welfarist approach is grounded are frequently considered not appropriate in the health care context.	To inform decision making, it requires to be combined with an estimate of the opportunity costs of investment. Cannot solve issues of allocative efficiency. Only one outcome (even if multi-dimensional) can be used for each analysis.
Key references	(Drummond, M. et al., 2015, Brazier et al., 2007, Coast, 2004)	(Culyer & Chalkidou, 2019, Drummond, M. et al., 2015, Brazier et al., 2007, Gafni, 2006, Donaldson, 1998)	(Sculpher, Claxton & Pearson, 2017, Drummond, M. et al., 2015, Sculpher & Claxton, 2012, Brazier et al., 2007, Torrance & Feeny, 1989)

Table 2.2 Analytical techniques for the economic evaluation of health care interventions

2.4. Economic evaluation of public health interventions

The previous sections explored the main normative principles (Section 2.2) and the analytical methods (Section 2.3) for the economic evaluation of health care programmes. In this section, the suitability and limitations of these approaches and techniques in assessing public health interventions are explored.

Theoretical approach and practical analytical technique of the analysis are a matter of choice. When conducting an economic evaluation of a health care intervention, the choice of the analytical technique depends mostly on the underlying normative underpinnings, and there is not an approach which is conclusively more correct than any other. Nevertheless, the choice depends also on characteristics of the intervention that is being evaluated (Culyer & Chalkidou, 2019, Drummond, M. et al., 2015). For example, when health care interventions yield a broad range of non-health benefits, such as in the case of evaluation of public health programmes, some methods might capture more accurately their full impacts (Buchanan & Wordsworth, 2015). In fact, in the last version of the NICE manual for developing guidelines, interventions with health and non-health outcomes in public sector and other settings were distinguished from interventions with health outcomes in NHS settings, and different recommendations (e.g. use of CBA) were also provided for alternative reference cases (NICE, 2015).

In sub-sections 2.4.1 and 2.4.2 the context of public health interventions is briefly delineated and the main challenges associated with the economic evaluation of such interventions are summarised. Sub-section 2.4.3 explores the suitability and limitations of traditional economic evaluation methods in assessing public health interventions. In this last section, before moving to Chapter 3, the normative principles and analytical techniques of the proposed framework are briefly illustrated.

2.4.1. Public health context

Acheson (1988) defined public health as *"the art and science of preventing disease, prolonging life and promoting health through the organized efforts of society"*. The target of public health interventions is therefore shifted from the individual patient to a group level. In other words, the 'public health service' has the role of identifying and responding to health problems in order to protect populations' health (Holland, 2004).

Examples of key public health areas are: conditions and diseases such as cardiovascular disease, cancer, obesity, diabetes, vaccine-preventable infections; behaviours such as smoking, drug or alcohol misuse, sexual and physical activity; other factors affecting health such as environment, working environment, housing and transport; accidents and injuries; child and maternal health; mental and oral health. Public health often deals with aspects of health and wellbeing broader than specific diseases, and public health policies take into account a wide spectrum of determinants of health, which work through four vectors: population, environment, society and organisations (NICE, 2012).

2.4.2. Challenges in the economic evaluation of public health interventions

In the context of public health interventions, economic evaluation should support decision makers in allocating the available resources to best meet public health aims (Edwards, Charles & Lloyd-Williams, 2013). Nevertheless, as highlighted in the Wanless Reports (2004), in the UK *"although there is often evidence on the scientific justification for action and for some specific interventions, there is generally little evidence about the cost-effectiveness of public health and preventative policies or their practical implementation"* (Wanless, 2004).

This general lack of cost-effectiveness evidence in public health may be due to the complex nature of public health interventions (Weatherly et al., 2009, Chalkidou et al., 2008). Research aiming to inform evidence-based decision making in public health can be technically difficult, and present methodological and technical issues (Wanless, 2004). Chalkidou et al. (2008) discussed these difficulties and pointed out several areas of controversy and uncertainty. Weatherly et al. (2009) identified in particular four main challenges in the economic evaluation of public health interventions: attribution of effects, measuring and valuing outcomes, identifying inter-sectoral costs and consequences, and incorporating equity considerations.

2.4.2.1. Quality of research evidence and attribution of effects

Public health interventions are targeted at populations or communities rather than specific individuals. Moreover, given the nature of public health interventions (e.g. preventative), any differences in outcomes between interventions may occur far into the future. As a consequence, in general, the available evidence is weaker with respect to clinical medicine, and sometimes is even lacking. For example, traditional randomised control trials (RCTs) are sometimes either impossible or excessively costly to perform, because long follow-ups would be required for the outcomes to be shown and large numbers of patients needed to demonstrate an effect (Deidda et al., 2019, Weatherly et al., 2009, Chalkidou et al., 2008). Further, generalisability issues might be faced when moving from evidence based on a sample (selected into RCTs) to the population, especially if the outcomes of individuals or groups of individuals change the behaviour of others (which is common in social examples and in public health whenever there is a possibility of contagion) (Deaton & Cartwright, 2018). In such a context, it is more common for researchers to use evidence from observational studies and find solutions to selection and measurement biases, reverse causality and omitted variables (Deidda et al., 2019).

2.4.2.2. Identifying inter-sectoral costs and consequences

Public health interventions may generate wider consequences and ripple effects in the economy than those relating to health alone. For example, an intervention to reduce alcohol consumption may affect both health and criminal behaviour. Similarly, a programme providing free school meals to disadvantaged children might have not only health effects, but also consequences for their education. Assessing the intervention from a narrow perspective would likely underestimate their impact, and potentially lead to unfair decisions about funding (Coast, Smith & Lorgelly, 2008).

Similarly to public health interventions, interventions in other sectors (e.g. education, criminal justice, housing, transport) may have effects not only on important outcomes for those sectors, but also on health. For example, an intervention to improve housing may also improve health levels, through a reduction of illness and injuries (Weatherly et al., 2009). Responsibility for public health in England recently moved away from the NHS to local authorities (LAs), which provided an opportunity to integrate public health with other LA functions such as education, housing and crime (Frew, 2017, Department of Health, 2012).

The presence of multiple sectors involved in the analysis might generate issues related to how to account for costs of the interventions. Since costs of public health interventions can fall on different parts of the public and private sector, budgetary impacts on different stakeholders should be identified separately (Sculpher et al., 2014, Claxton, Sculpher & Culyer, 2007). However, there is a lack of methods for properly accounting for costs when more than one government department or local government is involved in the delivery of an intervention. Moreover, issues might arise during the decision making process, especially when multiple decision makers are involved, and one government department gets the benefits of an intervention but another has to implement it (NICE, 2015, 2012).

2.4.2.3. Measuring and valuing outcomes

Methods based on standard health gain measurements may be inadequate as they may not capture the broader set of (non-health) benefits generated by public health interventions (Squires et al., 2016, Marsh et al., 2012, Weatherly et al., 2009).

It can be argued that non-health outcomes should be included in the economic evaluation. If so, it would become crucial to measure (i.e. estimate quantities) and value (i.e. choose which outcome to measure, and how to use results to inform decisions) them. Even though strongly related, measuring and valuing outcomes are distinct issues to be addressed singly, and imply the introduction of additional value judgments (i.e. how to value an outcome, and what is relevant and what is not worth to be recorded) in the evaluation (Squires et al., 2016, Marsh et al., 2012).

2.4.2.4. Incorporating equity considerations and concerns about health inequalities

Health inequalities are differences in health care access, quality of care or health outcomes present within a population, and are not necessarily unfair or unjust: they are a matter of fact (NICE, 2012). By contrast, as defined in the NICE guidelines (2012), *"a health inequity is an unnecessary, avoidable, unfair and unjust difference in someone's health or healthcare"* and is therefore a 'normative' judgment (whereas an inequality is a 'positive' statement).

Following the introduction of the Equality Act (2010), NICE has the duty to consider the effect of its activities in relation to equality in order to avoid unlawful discrimination and to identify opportunities for promoting equity (NICE, 2015, 2012, 2008). Moreover, NICE public health guidance considers not only the aspects protected by the Equality Act (age, disability, gender reassignment, pregnancy and maternity, race, religion or belief, sex and sexual orientation), but also other characteristics related to socioeconomic differences and particular situations (e.g. asylum seekers, homeless, gypsies and travellers) (NICE, 2012).

In many cases tackling unfair inequalities in health is the primary goal of a public health intervention or programme (Chalkidou et al., 2008). Health maximisation cannot be the only rationale for decision making, and the inclusion of social concerns related to the distribution of health become crucial (Weatherly et al., 2009).

2.4.3. Analytical techniques for the evaluation of public health interventions

2.4.3.1. Limitations of the 'traditional' CEA

CEAs are focused on the efficiency goal of health maximisation and struggle to capture non-health benefits (Edwards, Charles & Lloyd-Williams, 2013, Weatherly et al., 2009, Kelly et al., 2005). The use of CEA appears to be appropriate in performing more clinical-based assessments, whereas potentially underestimates the full impact of large-scale public health interventions (Claxton, Sculpher & Culyer, 2007). CEA's narrow perspective might in fact misrepresent the actual impact (in terms of costs and effects) of an intervention, and the omission of relevant cross-sectoral outcomes from a broader perspective might lead to misleading results that jeopardise the accuracy of the final funding decision.

Another limitation regards the availability and use of opportunity costs estimates. As explained in Section 2.2.3, opportunity costs of public health interventions are represented by what is forgone in order to accommodate the resources to provide new services, and are crucial for decision making (Sculpher et al., 2014). In CEAs of public health interventions, sometimes all costs are accrued and results are compared to the usual threshold for health care sector. Such a procedure is suggested also in the NICE guidance for public health evaluation (NICE, 2012), and leads to two main issues.

First of all, it has been argued that the £20,000-30,000 range used by NICE in the evaluation of public health interventions might be too high (Chalkidou et al., 2008). It has been demonstrated that the threshold does not reflect marginal productivity of the health care system (Claxton, Martin, et al., 2015). This criticism seems to be supported also by the empirical evidence. Owen et al. (2012) found that almost 90% of public health interventions assessed by NICE between 2006 and 2010 were cost-effective at a threshold of £30,000 per QALY, and the majority of them were highly cost-effective.

Secondly, even if it is assumed that the NICE threshold is acceptable for judging cost effectiveness, public health decisions can involve different decision makers' budgets. As a consequence, the resources displaced to accommodate a new investment correspond to a different amount of opportunity costs, and having multiple funders in place creates challenges with assigning opportunity cost for the intervention (Frew, 2017). If costs falling on other sectors are accrued, it would be incorrect to compare them to the health care threshold. Such a threshold would be not sufficient anymore to inform decisions in public health with effects falling on multiple sectors, and it would become necessary to estimate opportunity costs of each budget involved in the analysis, valuing from a broader perspective the resources displaced in each sector included (Sculpher et al., 2014, Claxton, Sculpher & Culyer, 2007).

2.4.3.2. Cost-consequences analysis

In the NICE guidelines for the appraisal of public health interventions it is recognised that public health involves aspects other than health, and this requires more inclusive methods of analysis and a change in the perspective (NICE, 2015, 2012). The NICE guideline for public health builds on and adapts the previous guidance for the evaluation of clinical interventions (Edwards, Charles & Lloyd-Williams, 2013), and the main change is the higher emphasis put on the use of CCA and CBA for interventions in this area (NICE, 2015).

As illustrated in Section 2.3.1, CCA provides a 'balance sheet' of cost and consequences of an intervention, and allows to capture multiple and also non-health outcomes. Examples of CCA in evaluation of public health interventions can be found in the literature (e.g. the analysis of an intervention to promote physical activity). Nevertheless, CCA is simply used just as a transparent way of reporting impacts of the intervention (Trueman & Anokye, 2013). In fact, the main drawback of the CCA regards its limits in informing decision makers, because it does not provide any guidance as to how the different outcomes should be weighed against each other (Greco, Lorgelly & Yamabhai, 2016). For this reason, Kelly (2005) and Weatherly et al. (2009) suggested that CCA might be a suitable pragmatic technique to systematically describe interventions characterised by a broad range of costs and benefits. Once completed a CCA, it would be possible to identify some narrower questions to be answered through CEAs within the CCA framework, or at the very least, the inter-sectoral impacts of interventions should be described qualitatively. In case it would be feasible to perform the CEAs, such an approach would permit to identify with separate analyses which programme is the best to improve health, which performs better in reducing inequalities, and which works in changing behaviour or promoting the uptake of the intervention. On the basis of these results, the estimation of the trade-offs between different outcomes across the interventions could be undertaken (Kelly et al., 2005).

2.4.3.3. Cost-benefit analysis

In the most recent version of NICE guidance, it is acknowledged that CBA would be able to take account of wider costs and benefits of public health interventions, and avoid the difficulties of aggregating data because CBA provides outcome measures and costs that are directly comparable (NICE, 2012). For the same reasons, CBA is recommended by HM Treasury's Green Book (2003) as well.

By using CBA in the public health framework and valuing everything in monetary and commensurable terms, comparisons between alternatives would be potentially straightforward, even when they include non-health related benefits. Conceptually, CBA could incorporate also impacts in other sectors of the economy (e.g. individuals' ability to return to work, impacts on education and criminal justice system), and would allow to perform inter-sectoral comparisons of how resources can be used (Weatherly et al., 2009, Gafni, 2006, Donaldson, 1998).

However, as already discussed in Section 2.3.2, the typical theoretical underpinnings of CBA might be deemed not appropriate when dealing with health, and in practice the implementation of CBA might be challenging. In particular, as distributional goals are often an explicit motivation for public health programmes, CBA may fall short in incorporating equity concerns in the evaluation (Vining & Weimer, 2010)

2.4.3.4. A brief overview of other potential methods

In the UK, the most common technique employed for the economic evaluation of health care programmes is CEA, where maximising health gains is considered the primary objective of the intervention (Drummond, M. et al., 2015, Brouwer et al., 2008, Coast, Smith & Lorgelly, 2008). Nevertheless, the widespread use of CEA does not mean that this method is *per se* superior to CBA, or other alternative analytical techniques. In fact, in other fields, such as environmental economics, the usual practice is to use CBAs (Gafni, 2006). CEA is usually implemented in the health care sector for feasibility reasons and with the aim of meeting decision-makers' requirements (Drummond, M. et al., 2015).

Particularly for the evaluation of public health interventions, traditional CEA showed several shortcomings. Alternative methods have been thus proposed for the evaluation of public health interventions, and in particular for the assessment of broader effects. A brief summary of the most relevant ones is provided here.

Social return on investment

Social return on investment (SROI) is a technique that shares many similarities with CBA and has been also proposed for the evaluation of public health interventions (Edwards, Charles & Lloyd-Williams, 2013). SROI measures social, environmental and economic outcomes using monetary values to represent them. Such a procedure enables a ratio of benefits to costs to be calculated (Nicholls et al., 2009). Similarly to CBA, SROI has its roots in welfarist economics, so it shares the same limitations (Edwards, Charles & Lloyd-Williams, 2013).

Capabilities

In CEA, the usual practice is to measure outcomes of health care interventions using QALYs. Nevertheless, QALYs are a health related outcome measure and do not include all the aspects that may matter to decision makers (Brazier & Tsuchiya, 2015), especially when evaluating public health interventions (Weatherly, Cookson & Drummond, 2014). In an ideal framework of analysis, local authorities should take into account not only health consequences of public health activities, but also wider outcomes and externalities (Sculpher et al., 2014).

Capability approach (Sen, 1993) can be employed as an encompassing measure of both health and non-health outcomes and is now considered one of the most relevant alternatives to the conventional methods (Greco, Lorgelly & Yamabhai, 2016, Marsh et al., 2012, Lorgelly et al., 2010). One of the most widely known suite of measures to value outcomes for use in economic evaluation is the ICEpop CAPability (ICECAP) (Coast, Kinghorn & Mitchell, 2015, Coast, Smith & Lorgelly, 2008). Capabilities offer a rich set of dimensions for evaluation rather than focusing solely on health status. Moreover, their theoretical extra-welfarist underpinnings make capabilities appropriate for use with public health interventions due to the emphasis put on equity concerns (Lorgelly et al., 2010).

Capabilities could in principle be used as a universal outcome measure in a CEA. Nevertheless, capability approach is a development paradigm (Greco, Lorgelly & Yamabhai, 2016). The implementation of the capability approach for use in economic evaluations is indeed hindered by the requirement of defining a legitimate capability space (i.e. identify a legitimate authority that can determine the dimensions to be considered) and the difficulties in measuring the relative preferences for each capability (Lorgelly et al., 2010). Moreover, pragmatic issues related to the lack of opportunity cost estimates and to the need of anchoring the measures at a value of zero for death prevent their use for resource allocation problems in health care (Brazier & Tsuchiya, 2015, Coast, Kinghorn & Mitchell, 2015). The implementation of the capability approach for decision making purposes is in fact still at a relatively early stage of development and there is a lack of consistency about how the measurement of capabilities can be used to aid decision-making (Mitchell et al., 2017, Coast, Kinghorn & Mitchell, 2015).

Subjective wellbeing

The use of capabilities in the analysis is usually backed by an extra-welfarist approach. On the contrary, a welfarist approach to the analysis would emphasise individual preferences, and criticise capabilities for relying on expert opinion and their inability to generate a monetary estimate of benefit (Dolan & White, 2007). For this reason, the use of subjective wellbeing (SWB) has been suggested as a potential alternative (Dolan & Metcalfe, 2011, Dolan & White, 2007).

The SWB approach measures the differences in people's self-assessment of their well-being as a consequence of the intervention analysed (Marsh et al., 2012). SWB is underpinned by traditional welfarist principles, where individuals are regarded as the best judge of their own conditions, and public policy has the goal to maximise the sum of everybody's happiness (or utility) (Greco, Lorgelly & Yamabhai, 2016). Consequently, an analysis based on SWB would suffer the usual welfarist limitations. For example, it might be not suitable when social concerns and distributional issues must be addressed (Coast, Smith & Lorgelly, 2008).

Even though meta-outcomes of happiness or well-being are already developed – examples are the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS) and the Wellbeing Adjusted Life Year (WELBY) (Coast, Smith & Lorgelly, 2008) - SWB is still in a developing stage, and there are few examples of its use to generate evidence to inform policy making (Greco, Lorgelly & Yamabhai, 2016, Marsh et al., 2012).

2.4.4. The proposed approach

In 2004, a HM Treasury's report recommended the development of a consistent framework to evaluate the cost-effectiveness of public health interventions, but this framework still does not exist (Wanless, 2004).

The aim of the next chapter is to address this gap in the literature and develop a framework for the evaluation of public health interventions that integrates, at different stages, different analytical techniques. The proposed framework is underpinned by a social decision making approach, even though alternative value judgments can be introduced at different stages of the analysis.

Combining emerging methods with those traditionally used for health economic evaluation has been already suggested in the literature (Briggs, 2016, Wildman et al., 2016). For example, Wildman et al. (2016) provided suggestions about how to develop a novel approach to the analysis which combines existing methods from economic evaluation beyond health, while maintaining the extra-welfarist approach to valuing health outcomes. The solution proposed here is based on a framework that is flexible (regarding both analytical techniques to be used, and theoretical underpinnings), but transparent and explicit.

With regard to the analytical techniques, the first stage of the framework consists of an initial descriptive CCA based on the 2nd Cost-Effectiveness Panel's 'impact inventory' (Sanders et al., 2016), but with the aim of disentangling budget impacts and direct effects according to decision makers' responsibilities. Afterwards, CEAs from different perspectives are conducted. These CEAs have the role of assessing the impacts of the public health intervention in all the relevant sectors. For this reason, an outcome measure commonly accepted for decision making purposes in that sector is used, and opportunity costs are made explicit.

Potentially, the CEAs within the sectors (or dimensions) can be adjusted in order to include equity concerns in the aggregation function. This step of the framework would imply the introduction in the analysis of principles that might be deemed more extra-welfarist. Furthermore, In order to obtain an overall cross-sectoral estimate, an aggregation function across sectors based on individual WTP is proposed. It is acknowledged that the use of WTP to aggregate across dimensions might be associated with a more welfarist approach to the evaluation. However, the proposed framework is still grounded in the social decision making approach, and no explicit social welfare function is defined.

It might be argued that the proposed approach resembles a CBA that acknowledges the existence of fixed budgets and opportunity costs, and allows for the incorporation of equity concerns. From a pragmatic point of view, results of such a CBA might be in fact similar to those obtained using the proposed framework. However, from a rigorous theoretical point of view, the framework proposed in this thesis is grounded in the social decision making approach. The underlying normative principles are not welfarist and a complete and explicit objective function is not specified. By contrast, efforts are made to link the analysis to the real decision making processes. Decision makers have their remits, objectives, perspectives and budgets constraints, and these features must be acknowledged in the evaluation. For these reasons, the starting point of the proposed framework is a CEA conducted from a narrow health care perspective, which reflects health care decision maker's remits. Afterwards, in order to expand the perspective and introduce alternative value judgments, the framework uses an amalgamation of approaches rather than focusing on the use of one method (Neumann et al., 2018, Sculpher & Claxton, 2012).

Details about the proposed theoretical framework can be found in Chapter 3.

Chapter 3: Analytical framework

3.1. Introduction to the framework

As illustrated in the previous chapter, the role of economic evaluation is to assist decision makers in their decisions about which intervention to fund from scarce resources (Drummond, M. et al., 2015). Nevertheless, the complex nature of public health interventions raises methodological and technical challenges in their evaluation (Weatherly et al., 2009). For example, a brief intervention for alcohol related problems might affect not only alcohol misusers' health, but also their probability of becoming involved in criminal activity and their employability. Moreover, expenditures in other sectors may have health effects. For example, investments in the housing sector (e.g. spending money on home safety) could reduce medical costs and improve lives (e.g. by reducing the risk of falls especially among the elderly) (Butler, 2018). Furthermore, frequently an important aim of public health interventions is to reduce health inequalities (NICE, 2012).

Analyses of public health interventions have typically focused on their impact on health, but such a narrow perspective potentially underestimates the full impact of public health interventions. It may be necessary to assess a broader range of costs and effects, including those falling on other sectors. Further, significant impacts might occur far in the future. Moreover, additional concerns about equity might become critically important (Weatherly et al., 2009). Within this context, the aim of this thesis is to define appropriate methods for the economic evaluation of public health interventions. In particular, this chapter illustrates the proposed theoretical framework for a cross-sectoral economic evaluation of public health interventions, with the potential incorporation of health equity considerations. The framework can theoretically accommodate also a cross-sectoral extension of the analysis. Such an extension is briefly introduced in this chapter, and is discussed as potential further work in Chapter 8.

A brief overview of the three main stages of the analytical framework is provided in the next sub-section, together with the normative underpinnings of the framework. The stages are then illustrated in detail in in the following sections. The last section of this chapter describes the case study selected to operationalise the analytical framework. In Chapters 4, 5, 6 and 7, a public health intervention will be used as a case study to illustrate how to implement the proposed framework.

3.1.1. Stages of the framework

The proposed framework consists of a three-stage analysis. In its first stage, it aims at extending a 'traditional' CEA to include the impact of the introduction of an intervention on different sectors (and different decision makers). The second stage concerns the incorporation of health equity implications in the analysis. In the third stage, the framework is extended to include the impacts of long-term costs and effects, and to reflect different budgetary policies. This last stage is discussed as a potential extension of the analysis in Chapter 8.

3.1.1.1. Cross-sectoral analysis

The cross-sectoral analysis is based on a framework originally formulated by Claxton et al. (2007), and then developed by Claxton et al. (2010), Sculpher et al. (2014) and Walker et al. (2019). The latest version of the framework, as described by Walker et al. (2019), can be seen as a broader, extended version of the 'impact inventory' proposed by the 2nd Cost-Effectiveness Panel (Sanders et al., 2016).

The cross-sectoral analysis aims at informing decisions to be taken by legitimate decision makers about interventions that have multiple effects which impact on different decision makers, and can be used to suggest compensation schemes across decision makers. Such compensation schemes could address resource allocation problems (also known as 'wrong pocket problems' (Butler, 2018, Erickson, Galloway & Cytron, 2012)), generated when costs fall on a sector and benefits on another (Walker et al., 2019).

First of all, the cross-sectoral analysis requires the specification of the perspective employed for the analysis. Afterwards, the first step of the cross-sectoral analysis consists of assessing direct effects (i.e. what is gained) and opportunity costs (what could have been gained with an alternative use of the resources) for each sector (or dimension). The second step is to explore ways to aggregate across dimensions (Walker et al., 2019).

Methods to conduct the cross-sectoral analysis and use its results to inform decision makers are described in Section 3.2.

3.1.1.2. Incorporation of health equity concerns

The cross-sectoral analysis can be extended in order to include an assessment of the impacts of the intervention on health inequalities that might be deemed unfair and unjust (i.e. health inequities). Equity impact and equity trade-off analyses can be used to incorporate concerns for the distribution of impacts across population subgroups in the aggregation within dimensions (Cookson et al., 2017).

Methods to assess health inequalities and to incorporate equity concerns in the cross-sectoral analysis are illustrated in Section 3.3.

3.1.1.3. Cross-temporal extension

Extrapolation techniques can be employed to estimate the full impact of the intervention, including long-term costs and consequences (Bojke et al., 2017). To reflect intertemporal issues, the impact of budgetary policies (i.e. whether annual deficits are allowed or not) can then be assessed. Rigidity of budget constraints over time may limit the scope to implement interventions which would be cost-effective if decision makers had more control over inter temporal budgets (McKenna et al., 2010). Moreover, constraints on the time horizon for the analysis affect the relevant costs and benefits included in the evaluation (Frew, 2016).

The cross-temporal extension of the framework is only discussed as a potential extension of the analysis in Chapter 8.

3.1.2. Underlying normative principles of the framework

In the literature, it has been stressed how any evolution in the analytical techniques for economic evaluation should be accompanied by more debate about the theoretical and value propositions underlying the various forms of economic evaluation (Marsh et al., 2012, Weatherly et al., 2009). In Section 3.1.1, the analytical techniques employed in the framework have been briefly described. This sub-section clarifies the value assumptions behind the framework.

It is generally accepted that the economic evaluation approach should fit with the priorities of the decision makers (Frew, 2017). Therefore, with the aim of reflecting the actual decision making process and meeting decision maker requirements, the proposed framework is grounded in a social decision making approach.

As illustrated in Chapter 2, Section 2.2.3, the assumption behind the social decision making approach is that each decision maker has its own objective(s) and its own constraints (potentially related to the available budget and a fixed time period), and is legitimised (by the social-democratic process) to take decisions in its area of responsibility. This approach is not based on particular value assumptions and avoids the specification of a complete social objective function (e.g. based on welfare or wellbeing). Nevertheless, decisions still need to be made. Current budgetary allocations (i.e. budgets available for decision makers) is thus seen as a legitimate expression of some unknown underlying social value function, and budgets themselves contribute to determining the opportunity costs of investments. The aim of the social decision making approach is therefore to inform a series of decision makers, each with a remit for a particular area, by reporting impacts of alternative decisions in a transparent way (Walker et al., 2019, Drummond, M. et al., 2015, Paulden & Claxton, 2012, Claxton et al., 2011).

Nonetheless, in subsequent stages of the economic evaluation, additional value assumptions will be required in order to perform the evaluation. Normative judgments are necessary in the pragmatic choice of sectors and dimensions[4] to be considered in the analysis, and for choosing the methods of aggregation of the outcomes within and across dimensions. The framework is sufficiently flexible to accommodate the inclusion of principles that might reflect a more welfarist or extra-welfarist approach to the evaluation. Nevertheless, the framework forces these normative judgements to be explicit.

In summary, the proposed framework is based on CEAs grounded on a social decision making approach. The aggregation function[5] within dimensions also allows the introduction of health inequality considerations. The aggregation function across dimensions allows for the use of social valuation based on WTP to aggregate outcomes. The incorporation of equity considerations and the use of social valuation based on WTP might be deemed more related to an extra-welfarist and a welfarist approach, respectively. However, the analysis avoids the specification of a complete social objective function and aims at reflecting decision making processes and context. The framework is also flexible and potentially compatible with other methods (and therefore other theoretical underpinnings), such as SROI, standard CBA, and capability approach (see Section 2.4.3). All these methods can be captured using the proposed framework, exposing the set of value judgements which is inherent in each (Walker et al., 2019).

---

[4] Sectors might not necessarily equal dimensions.

[5] See Sub-section 3.2.2.1

3.2. Cross-sectoral analysis

Public health interventions are challenging to evaluate because of their costs and effects falling on different sectors. Typically, health economic evaluations focused on a single-sectoral payer that seeks to maximise health, typically through interventions delivered by the health care system. Nevertheless, focusing on health only would underestimate the total impact of the intervention on society. Moreover, multiple sectors contribute to the production of health, and the health care sector might generate benefits besides health (Remme, Martinez-Alvarez & Vassall, 2017).

The framework aims to extend the 'traditional' health economic evaluation across sectors. Before exploring how to manage and aggregate cross-sectoral costs and effects, it is firstly necessary to have a clear and explicit definition of the perspective employed for the analysis.

3.2.1. Definition of a broader perspective

Public health interventions might affect multiple decision makers with different objectives and responsibilities. To evaluate these interventions, it is frequently recommended in the literature to use a 'societal perspective' for the evaluation (Jonsson, 2009, Weinstein, 1990).

However, the way the term 'societal perspective' is interpreted and operationalised within economic evaluations is largely subjective. Analyses conducted from a 'societal perspective' often merely include productivity costs in addition to health effects and costs (Drost et al., 2017). Moreover, even if a clear societal perspective was defined, it would still be unclear how to use this information to inform choices across different settings and decision makers (Claxton et al. 2010).

For all these reasons, as suggested by Walker et al. (2019), in this analysis the term 'societal perspective' is avoided. However, the need of broadening the perspective from that of health is acknowledged, and the framework considers an integrated perspective that involves multiple decision makers with different objectives and budgets.

3.2.1.1. The 'impact inventory' and its dimensions

The concept of a broader and integrated perspective can be operationalised using the 'impact inventory' recommended by the 2nd Cost-Effectiveness Panel, which is a structured table that contains consequences both inside and outside the formal health care sector (Sanders et al. 2016).

Potentially, the impact inventory could include all consequences in their natural units, and without any intrinsic value judgments. However, for pragmatic reasons, the impact inventory cannot be exhaustive and the choice of the dimensions to be included is based on a series of value judgements (Walker et al., 2019). Moreover, other value judgments will be necessary in the selection, for each sector, of an outcome measure deemed appropriate to inform decisions about which interventions to fund. Nevertheless, these value judgments simply reflect the existing decision making structure and process, and the current resources allocations.

Using the 'impact inventory' tool, an intervention can be described in terms of its impacts on a set of dimensions of interest, determined by value judgements and institutional arrangements of the decision makers to be informed (Walker et al., 2019). The list proposed by the 2nd Cost-Effectiveness Panel assumes correspondence between sectors and dimensions, and includes: Formal and Informal Health Care sector; Non−Health Care sectors (Social Services, Criminal Justice, Education, Housing, and Environment); and Private sector (Productivity and Consumption). Other dimensions can also be specified (Sanders et al., 2016).

Economic evaluations are usually conducted from the perspective of the decision maker allocating public funds, therefore informal care is rarely valued for inclusion in the analysis (Weatherly, Faria & Van Den Berg, 2014). The dimension 'Formal and informal Health Care' identified by the 2nd Cost-Effectiveness Panel would already imply an extension of the conventional perspective usually employed in the UK for the economic evaluation of health care interventions (NICE, 2013). For this reason, the two dimensions are here distinguished.

When comparing the list of dimensions proposed in the 'impact inventory' to the conventional approach proposed by NICE guidelines (2013) in the UK, the distinction of formal and informal health care is not the only difference. Social care is in fact usually included in the analysis from a health care perspective. This separation is maintained here, but this aspect will be discussed in more detail in the next chapter (Sub-section 4.4.4) and in Chapter 8, Sub-section 8.3.

In order to show the importance of using a broad perspective when assessing public health programmes, a few illustrative examples are briefly outlined using the impact inventory.

3.2.1.2. Examples of cross-sectoral impacts of hypothetical public health interventions

The following examples are based on actual public health interventions reviewed in the process of identification of a case study for this thesis.

Programme providing free school meals to disadvantaged children

The Free School Meals (FSM) pilot was a two-year programme operating in three local authorities between the autumn of 2009 and summer of 2011, to extend entitlement to free school meals (Kitchen et al., 2013).

The programme provided free lunches for eligible pupils and therefore affected private consumption of these individuals. Moreover, pupils were provided with healthier types of food at lunchtime, and impacts of the FSM programme did not fall only on educational attainment (which was the main objective of the study), but also on health care. With regard to the costs of the intervention, the programme was funded by two local authorities that made free school meals available to all primary school children, and a third local authority that made them more available to both primary and secondary school children by increasing the number of families entitled to them.

Intervention to tackle substances misuse

The Randomised Injectable Opiate Treatment Trial (RIOTT) compared supervised injectable heroin or injectable methadone with optimised oral methadone for chronic refractory heroin addiction in patients not responding to current oral maintenance treatment (Byford et al., 2013).

The costs of RIOTT fell on NHS and pharmacies providing the treatment. Impacts ranged from health (deaths, QALYs, levels of substances attributable diseases, changes in hospitalisations attributable to substance misuse), to criminal behaviour and nights in prison, social services and housing (residential care, hostels and shelters) (Byford et al., 2013, Hunter & Hasan, 2013). Furthermore, even if not assessed in this evaluation, such an intervention might also affect transportation (due to road accidents) and private consequences (productivity, work absence and money spent on drugs) related to addiction.

Table 3.1 provides a summary of the potential important impacts of the programmes on the identified sectors.

Sectors (dimensions)		Free school meals	Substances misuse
Public	Health Care	X	X
	Social services		X
	Criminal justice		X
	Education	X	
	Housing		X
	Environment		
	Other (e.g. Transport)		X
	Other (e.g. Local Authority)	X	X
Private	Informal Health Care		X
	Employers (e.g. Productivity)		X
	Individuals (e.g. Consumption)	X	X

Table 3.1 Impacts on health care and non-health care sectors

The objective of these examples is to show that conducting an analysis from a narrow perspective tailored to one decision maker risks omitting important outcomes for other decision makers (Walker et al., 2019).

In the 2nd Cost-Effectiveness Panel's approach, opportunity costs falling on different sectors are not made explicit. On the contrary, the proposed framework aims at extending the 'impact inventory' and obligates the distinction between direct impacts and opportunity costs (i.e. impacts associated with what is forgone) (Walker et al., 2019). In fact, the economic evaluation involves multiple decision makers (from multiple sectors) with their own budget, and expenditures must be allocated correctly, in order to reflect the different opportunity costs (implied by the resources displaced in that sector) appropriately (Sculpher, Claxton & Pearson, 2017, Sculpher et al., 2014, Claxton et al., 2010, Claxton, Sculpher & Culyer, 2007).

3.2.2. Direct effects and opportunity costs falling on multiple sectors and decision makers

According to the recommendations of the 2nd Cost-Effectiveness Panel, items listed in the impact inventory should be presented in the form of disaggregated consequences across different sectors. This is because *"there are no widely agreed on methods for quantifying and valuing some of these broader effects in cost-effectiveness analyses"* (Sanders et al., 2016).

The aim of the work by Walker et al. (2019) was precisely to go beyond this stage of the evaluation and define an appropriate framework, which is built on the cross-sectoral resource allocation scheme developed by Claxton et al. (2007) and Sculpher et al. (2014). The framework captures the effects and the opportunity costs in each sector, then considers whether sectors are winners or losers and can be used to suggest potential compensations. The framework aims at informing decisions to be taken by legitimate decision makers about interventions characterised by multiple objectives and resulting in impacts on multiple budget constraints. A resource allocation scheme based on compensation schemes across decision makers can be proposed to make decisions, but other decision making criteria can also be employed (e.g. dominance criteria). The proposed cross-sectoral framework reflects the exogenous budget allocations across sectors, which are determined by the political debate and the socio democratic process. The amount of resources allocated to each sector and its production function determine the marginal productivity of the sector and consequently the opportunity costs and the monetary value associated with each outcome. The framework reflects therefore, the current resource allocation and fits decision makers' objectives and the approaches already taken to inform decisions (Walker et al., 2019).

While grounded in the work by Walker et al. (2019), the version of the analytical framework proposed here is more focused on the pragmatic operationalisation of a cross-sectoral analysis, and is thus less abstract and more prescriptive. While the original framework aims to illustrate all the potential alternative options and solutions for conducting the evaluation, here some choices in the way of conducting the analysis have been taken (e.g. with regard to the outcomes to be evaluated and the method to aggregate them – see Sub-section 3.2.2.1). Nevertheless, these choices are explicitly stated, and the analytical framework proposed retains the aim of being generic, flexible, adaptable and implementable in various public health contexts.

To operationalise the resource allocation scheme, for each sector (dimension) it is necessary to define an outcome measure that is deemed appropriate for decision making purposes. Moreover, an estimate of the marginal productivity (k) of the sector for the selected outcome measure is necessary. The marginal productivity of the sectors represents the opportunity costs of the investment (i.e. what could be alternatively done with those resources).

3.2.2.1. Criteria for aggregation within and across dimensions

Walker et al. (2019) suggested two potential alternative general aggregation approaches: a within-individual approach and a within-dimension approach. The first approach consists of aggregating first within-individuals across all dimensions, and then to aggregate across individuals. The latter approach aggregates first across individuals for each dimension, then across dimensions at the population level.

In order to develop this analytical framework, it was necessary to choose between these aggregation approaches. Because the objective of this work is to start from an evaluation from the health care perspective, then add dimensions to the analysis, the framework proposed here follows the within-dimension approach. Moreover, the within-dimension approach might be considered more appropriate given decision makers' remit to focus on a particular sector.

The first aggregation is therefore within dimensions. As a result, the proposed framework is slightly more prescriptive when compared to the work by Walker et al. (2019). However, methodological choices are made explicit and they pave the way to further extensions of the analysis (i.e. the incorporation of health equity considerations and cross-temporal extension). Considerations (in particular in terms of impacts on equity) about using the alternative within-individual aggregation approach are further discussed in Chapter 7, Sub-section 7.4.1.

3.2.3. Aggregation within dimensions

When aggregating impacts on each dimension across individuals, many functional forms (at dimension level) are possible, each representing alternative normative judgements (Walker et al., 2019). In this first cross-sectoral stage of the framework, the aggregation is based on the sum of the unweighted changes in dimensions across individuals. In the next stage of the framework (described in Section 3.3), concerns about health inequality will be introduced and the changes across individuals will be weighted, for instance by their current allocations or other equity judgments.

3.2.3.1. The health care sector

With regard to the health sector, it is assumed that the objective of the decision maker is to maximise health (H), and that health is measured in QALYs. Therefore, for a new intervention to be funded, the health benefits associated with the new intervention must be higher than the opportunity costs (OCs) of the investment. This can be assessed using the INHB. The INHB of an intervention that costs ΔC_H to the health system is given by the difference between the direct effects (DEs) on health (ΔH) and the health OCs. Consequently, if the DEs on health are higher than the health OCs, then the intervention is cost-effective and is associated with positive INHB. To estimate the OCs of an intervention, an estimate of the marginal productivity of the health care sector (k_H) is required. In fact, by knowing that at the margin one QALY will be lost as a result of every k_H pounds not spent, direct health care costs (ΔC_H) can be converted into health OCs (Drummond, M. et al., 2015).

The formula to calculate the INHB is described in Equation 3.1. If NHB is positive, the decision maker would want to recommend the intervention.

$$INHB = Direct\ effects\ on\ health - Health\ opportunity\ costs$$

$$INHB = \Delta H - \frac{\Delta C_H}{k_H} \qquad (3.1)$$

The health OCs of changes in spending across the NHS budget have been recently estimated, and it was found that for approximately every £ 13,000 increase or decrease in spending, one QALY is gained or lost elsewhere (Claxton, Martin, et al., 2015). In other words, £ 13,000 is the (central)[6] estimate of k_H, the marginal productivity of the health sector (Sculpher, Claxton & Pearson, 2017, Woods et al., 2016).

Details about the methods to conduct a cost-effectiveness analysis from the narrow health care perspective are illustrated in Chapter 4.

3.2.3.2. Private consumption

The impact on private consumption (ΔC_C, measured in £) can be assessed by measuring the differential costs falling on individuals in terms of individuals consumption, employers productivity and informal health care.

From the narrow perspective of the individual's out of pocket expenditures, the Incremental Net Consumption Benefit (INCB) can be computed and valued in monetary terms as shown in Equation 3.2.

$$INCB = \Delta C_c \qquad (3.2)$$

3.2.3.3. Any other potential non-health care sector

The same approach used to assess Incremental Net Benefits (INB) from the health care perspective can be extended to other sectors. For illustrative purposes, a generic non-health sector is defined here. The same structure can be then applied to specific sectors (e.g. education, criminal justice…) and expanded to multiple dimensions (i.e. more than one non-health sector).

Let us assume that the same intervention to be assessed has not only direct effects (ΔH) and costs (ΔC_H) on health, but also a direct non-health effect (ΔN) and direct non-health costs (ΔC_N). From the

---

[6] The estimation was subject to parameter and structural uncertainty.

narrow perspective of the non-health sector decision maker, the same adoption criterion can be used, and an Incremental Net Non-health Benefit (INNB) can be computed as illustrated in Equation 3.3.

$$INNB = \Delta N - \frac{\Delta C_N}{k_N} \qquad (3.3)$$

If the INNB is positive, the recommended decision for the non-health decision maker will be to implement the intervention.

As previously stressed, the method is therefore based on a social decision making approach, and to perform a CEA outside the health care sector, it is necessary to define a non-health outcome which is appropriate to inform decision in the relevant sector. Variations in the outcome in each sector must be compared to a comparable estimate of opportunity costs based on what is displaced to accommodate the additional costs.

The lack of opportunity costs estimates (i.e. what benefits are displaced elsewhere by diverting resources to a certain intervention, or what benefits could be generated if the resources were used for another purpose) outside the health care sector, is well known (Hill et al., 2017). There is a gap in the literature regarding the estimation of the marginal productivity (k_N) for public sectors other than health (Sculpher et al., 2014)[7]. Nevertheless, in this thesis pragmatic solutions are proposed, and details about the specific methods to conduct a cost-effectiveness analysis from the criminal justice system perspective are illustrated in Chapter 5.

3.2.3.4. Spill-over effects

With regard to what is forgone to implement the intervention, costs generated in the health care sector have OCs within the health sector that can be computed using k_H. However, OCs can also fall outside the health care sector (i.e. interventions displaced in the health sector may have impacts on the non-health sector and on private consumption). To be calculated, these cross-sectoral OCs require an estimate of the marginal productivity of the health sector for non-health outcomes and for consumption, called k_{HN} and k_{HC}, respectively.

The same reasoning applies to the non-health sector. Costs generated in the non-health sector have OCs within the same sector, due to potential alternative use of the same resources for other non-health policies. These OCs can be calculated using the marginal productivity of the non-health

---

[7] Even in health there are actually limited estimates for countries based on their own data. Nevertheless, estimates exist based on cross-country methods (Woods et al., 2016)

sector (k_N). Nevertheless, non-health OCs can also fall outside the non-health sector. The health that is forgone as a result of non-health resources not being used for other non-health policies can be calculated using the marginal productivity of the non-health sector for health (k_{NH}). Additionally, the consumption that is forgone as a result of the non-health care costs is given by the marginal productivity of the non-health sector for consumption (k_{NC}). It is assumed here that changes in private consumption are one of the potential direct effects of the intervention. There are no opportunity costs (in health and non-health) generated by consumption.

Figure 3.1 describes the structure of the cross-sectoral DEs and OCs. DEs fall on health and non-health sectors, and on consumption. Costs falling on the health sector (highlighted in red in Figure 3.1) generate OCs not only within the health sector, but also in non-health sector and consumption. Similarly, costs falling in the non-health sector (highlighted in blue in Figure 3.1) generate OCs in the non-health and health sector, and in consumption. Marginal productivities (k's) can be used to express costs in appropriate units of measurement for each sector.

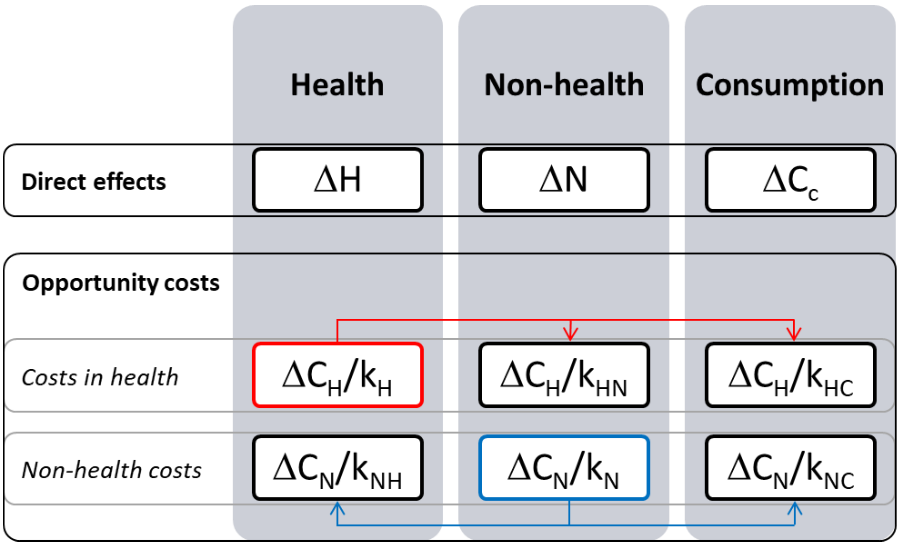


Figure 3.1 Scheme of Cross-Sectoral Effects and Opportunity Cost

Equation 3.1 for INHB can therefore be extended in order to include all the actual opportunity costs.

$$INHB = \Delta H - \left(\frac{\Delta C_H}{k_H} + \frac{\Delta C_N}{k_{NH}}\right) \tag{3.4}$$

The same extension can be applied to the INNB and INCB formulae to give:

$$INNB = \Delta N - \left(\frac{\Delta C_N}{k_N} + \frac{\Delta C_H}{k_{HN}}\right) \tag{3.5}$$

$$INCB = \Delta C_c - \left(\frac{\Delta C_N}{k_{NC}} + \frac{\Delta C_H}{k_{HC}}\right) \tag{3.6}$$

INHBs, INNBs and INCBs can be computed as illustrated in Equations 3.4, 3.5 and 3.6, and are expressed in different units. INHBs are measured in health units (usually QALYs), INNBs in the specific non-health outcome considered in the analysis (e.g. educational outcome or criminal justice outcome) and INCBs in monetary terms. To inform cross-sectoral decision making, these benefits must be somehow aggregated into a single comprehensive INB measure.

3.2.4. Aggregation across dimensions

In order to perform the aggregation across dimensions, a common metric is needed. For this reason monetary values in consumption terms are used as a common numeraire across the valuation of dimensions. In this way each attribute can be valued, allowing the identification of trade-offs and heterogeneity in valuation (Wildman et al., 2016). For each dimension, the willingness to give up consumption V_i for the outcome produced by sector *i* can be estimated (Woods et al., 2016). This estimate corresponds to the consumption value for the outcome of the selected sector. The monetary value of a unit of health output in terms of consumption (V_H) (i.e. the amount of consumption an individual would trade for one unit of health), and the monetary value of a unit of non-health output in terms of consumption (V_N) (i.e. the amount of consumption an individual would trade for one unit of the non-health outcome) can therefore be used to aggregate outcomes across dimensions.

The introduction of the individual WTP V_i in the analysis might be seen as a move away from the social decision making approach. It might be argued that an analysis that uses monetary valuation of the benefits based on individuals' consumption reflects a more welfarist approach. Nevertheless, the framework is still grounded in the assumptions of the social decision making approach. The valuation of the OCs is determined by the current allocation of resources among sectors and their production functions. Because of the social democratic process, this allocation is accepted as it is, without arguing it is optimal or correct. It is recognised and accepted that the individual valuations V_i of the outcomes might differ from the marginal productivity of the sector k_i (which is determined by the budget allocations) (Claxton, Sculpher & Culyer, 2007).

In the formulation of the framework proposed here, once impacts in each dimension are measured in terms of individual consumption, the aggregation approach is additive across dimensions. Nevertheless, the framework is flexible and allows for the use of other aggregation functions that might be deemed more appropriate according to the value judgements of the decision makers involved (Walker et al., 2019). It is actually not even guaranteed that a consensus between the different decision makers on the method for aggregation and the values used could be reached.

The framework can however provide evidence on the cross-sectoral impacts of the intervention and present results based on different value judgments, and therefore help to inform decisions (Walker et al., 2019, Remme, Martinez-Alvarez & Vassall, 2017, Claxton et al., 2010).

As illustrated in Equation 3.7, the cross-sectoral Incremental Net Monetary Benefit (INMB) can be obtained by adding INHB and INNB when expressed in monetary consumption units. For example, the net monetary benefit of the health element of an intervention can be derived by multiplying the INHB with the monetary value of a QALY to give health benefits measured in monetary consumption units. When INHB and INNB are converted into monetary terms using individual WTP, it becomes straightforward to also include impacts on individual consumption (INCB).

$$INMB = V_H * INHB + V_N * INNB + INCB \qquad (3.7)$$

The resulting cross-sectoral INMB can be computed as shown in Equation 3.8.

$$INMB = V_H \left(\Delta H - (\frac{\Delta C_H}{k_H} + \frac{\Delta C_N}{k_{NH}}) \right) + V_N \left(\Delta N - (\frac{\Delta C_N}{k_N} + \frac{\Delta C_H}{k_{HN}}) \right) + \left(\Delta C - (\frac{\Delta C_N}{k_{NC}} + \frac{\Delta C_H}{k_{HC}}) \right) \qquad (3.8)$$

The practical operationalisation of the cross-sectoral framework is hindered by data availability issues. Estimates of marginal productivity within the health sector and across health and consumption are already available (Claxton, Martin, et al., 2015, Claxton, Sculpher, et al., 2015). Claxton et al. (2015) found that for approximately every additional £13,000 spent, a QALY is displaced elsewhere within the health care system (Claxton, Martin, et al., 2015). Moreover, it was found that for every additional £1.11 spent in health care, £1 of consumption is lost[8] (Claxton, Sculpher, et al., 2015, Roberts, G., 2015). These findings provide the estimates for k_H and k_{HC} in Equation 3.8. In contrast, marginal productivities (k's) of other public health expenditures have not been estimated. The lack of estimates of *k's* outside the health care sector requires the introduction of pragmatic assumptions. Details about implementation of the cross-sectoral analysis for the health and criminal justice sectors are described in Chapter 5 and Chapter 6.

3.2.5. Informing cost-effectiveness from an integrated perspective

Theoretically, from an integrated perspective, if the INMB resulting from Equation 3.8 is positive, then the new intervention should be introduced. In practice, all the decision makers involved may

[8] This is assumed to be equivalent to patient's net production (i.e. their contribution or production of resources, net of their consumption or utilisation of resources)

only agree about the introduction of the new programme if all the components of the overall INMB (i.e. the INBs for each sector) are positive.

It might well happen that INHB and INNB have opposite signs. In such a situation one sector would be bearing the costs of the intervention, and another benefitting from it (de Salazar et al., 2007). In other words, one sector wins and one loses. Such a situation leads to a conflict between the decision makers and has been referred to as the ‘wrong pocket problem’ (Butler, 2018, Erickson, Galloway & Cytron, 2012)’.

A wrong pocket problem undercuts the incentives for collaboration, but the proposed framework would help decision makers in address such resource allocation decisions. Nevertheless, in order to ensure the implementation of programmes that are cost-effective from an integrated perspective, but not cost-effective from a narrower perspective, it may be necessary to enforce adequate compensations among sectors. If the sector that benefits from the intervention is able to compensate the loss in the other sector, then the intervention can be implemented and can generate an overall positive benefit (Claxton, Sculpher & Culyer, 2007). However, as discussed in Chapter 8, Sub-section 8.6, if actual compensation were to be paid, issues about misreporting and gaming the system might arise.

3.3. Incorporation of health equity concerns in the analysis

As illustrated in Section 3.2.3, in the cross-sectoral stage of the framework, the aggregation at dimension level was simply based on the sum of the unweighted changes in dimensions across individuals. Nevertheless, the aggregation within dimensions could potentially take into account distributions of impacts across population subgroups (Walker et al., 2019).

Including distributional concerns in the cross-sectoral analysis is particularly important when dealing with health outcomes, because one of the aims of public health interventions is specifically to reduce unfair inequalities (i.e. inequities) in health (NICE, 2012, 2008). Therefore, the aim of this section is to illustrate how to incorporate health equity considerations into the analysis.

3.3.1. Rationale for the inclusion of health equity considerations into the evaluation

As mentioned in Chapter 2, Section 2.4.2, whereas health inequalities are differences in health or health care and are not necessarily unfair or unjust, *"a health inequity is an unnecessary, avoidable, unfair and unjust difference in someone's health or healthcare"* (NICE, 2012). Health inequities are therefore related to social justice and value judgements, and NICE has the duty to address health inequity issues (NICE, 2015, 2012, Marmot et al., 2010, NICE, 2008).

In the light of this, the distribution of health impacts between population sub-groups assumes particular relevance. Cost-effectiveness might not be the only rationale for decision making, and the inclusion of social concerns related to the distribution of health might become relevant (Weatherly et al., 2009).

3.3.2. Health inequities and value judgments

The policy objective underpinning a conventional CEA is to maximise total health in the general population, subject to a budget constraint. CEAs are typically grounded on the value judgement that "a QALY is a QALY", irrespective of who is the beneficiary. Population-level health gain thus corresponds to the unweighted sum total of all individual health gains (Cookson et al., 2017, Drummond, M. et al., 2015). However, even though the "QALY is a QALY" principle is widely assumed in the practice of economic evaluation, it is not universally accepted. It has been suggested that there are occasions when differential consideration should be given to health gains based on the characteristics of those receiving care (Round & Paulden, 2018).

Alternative value judgements could therefore be endorsed in the evaluation. For instance, diminishing returns of health benefits (Dolan, Shaw, et al., 2005) could be incorporated into the analysis (Walker et al., 2019). This is because it has been argued that social value does not increase linearly in marginal increments of quality and length of life (and consequently a QALY maximisation approach would actually be based on wrong assumptions). Moreover, other attributes affecting the social value attached to health gains could be introduced in the evaluation, such as the desire to reduce health inequities, the dispersion of benefits as widely as possible and the compensation of certain groups for their disadvantaged lifetime health prospects (Dolan, Shaw, et al., 2005).

The rationale for introducing adjustments (based on alternative value judgments) in the assessment of health variations is related to the concept of vertical equity. According to the concept of vertical equity, different needs of individuals should correspond to different levels of health care. In other words, morally relevant factors can justify unequal health care treatments. Nevertheless, the focus on vertical equity should not however divert attention from ensuring horizontal equity as well. According to the principle of horizontal equity, people with the same health care needs should receive equal treatment (Culyer, 2015, Wonderling, 2011).

3.3.2.1. Normative principles and potential value judgements

As explained in Chapter 2, Sub-sections 2.2.1 and 2.4.3.3, according to the welfarist approach, the individual should be the one and only source of valuation. For this reason, the analysis could be grounded in a welfarist approach only if value judgments and distributional considerations can be included in individuals' utility function. Such an approach to the evaluation would not allow for the inclusion of exogenous equity concerns in the evaluation.

In the proposed framework it is recognised that tackling unfair and unjust health inequalities is one explicit target of public health decision makers. Health equity implications must be taken into account in the analysis. For this reason, this stage of the evaluation allows external authorities to impose weights to particular populations or outcomes.

The aim of including the health inequity assessment in a formal analytical framework is to make social value judgements explicit and allow decision makers to explore the impacts of alternative value judgments and investigate potential trade-offs. Such an analysis can be used to determine the strategy that best addresses the dual objectives of maximising population health and minimising variations in health which are considered unfair on the basis of agreed value judgments (Asaria et al., 2015).

If impacts on individuals' health level or other alternative equity considerations are incorporated into the evaluation, it becomes fundamental to assess where, in the general population, the health impacts fall. A potential grouping of the population can be based on socioeconomic status (SES) (Cookson et al., 2017).

The unequal (and potentially inequitable) distribution of health

Using a categorisation based on the Index of Multiple Deprivation (IMD)[9], Love-Koh et al. (2015) investigated the quality-adjusted life expectancy (QALE) at birth for each wealth quintile group of the population in England. As reported in Table 3.2, the authors found a positive correlation between SES and QALE: individuals in the lower IMD quantiles groups have lower QALE at birth.

IMD quantile	Quality-adjusted life expectancy (QALE)
Q1 - poorest	64.66
Q2	68.55
Q3	70.58
Q4	73.57
Q5 - richest	75.63

Table 3.2 Population QALE according to IMD quantiles, adapted from Love-Koh et al. (2015)

The current unequal distribution of health might raise equity issues. The association between SES (here measured through the IMD) and QALE might be deemed unnecessary, avoidable, and unjust. Therefore, health inequalities across IMD subgroups might be judged inequitable (Marmot et al., 2010).

In order to reduce health inequity, decision makers might want to prioritise interventions that benefit disadvantaged (and therefore less healthy) groups. If concerns about deprivation and the distribution of health are included in the economic evaluation, it becomes necessary to assess whether the intervention affects health inequalities (in terms of differential QALEs) as well as assessing cost-effectiveness (Asaria, Griffin & Cookson, 2016).

Other potential value judgments

SES can be used as the equity-relevant characteristic in determining population subgroups. Nevertheless, the perceived fairness of health inequalities might also be based on a broad range of other factors, such as: burden, rarity or severity of disease or condition; age; responsibility or culpability, or other personal characteristics (also in relation with other people or societal positions) (Bobinac et al., 2012).

[9] IMD is the official measure of relative deprivation for small areas in England, and ranks every small area in England from 1 (most deprived area) to 32,844 (least deprived area) (Department for Communities Local Government, 2015).

In the UK, factors beyond cost-effectiveness already influence the decision-making process (Briggs, 2016). For example, NICE introduced 'end-of-life criteria', which are based on the assumption that the general public would put extra value on health outcomes achieved at the end of life. Appraisal Committee are asked to consider the impact of giving greater weight to the treatment gains achieved in the later stages of disease if: the treatment is indicated for patients with short life expectancy; the treatment offers an extension to life compared to current NHS treatment; the treatment is licensed for small patient populations (Shah, Tsuchiya & Wailoo, 2015, NICE, 2013). As a result, higher priority can be given to life-extending treatments than to other types of treatments.

3.3.3. Analytical techniques to incorporate value judgements in the CEA

Sub-section 3.3.2 described potential value judgments that could be included in economic evaluations. The objective of this sub-section is to explore the available analytical techniques to integrate social value judgments into economic evaluations. In other words, alternative functional forms to aggregate health impacts are investigated.

Various methods are available, and can be classified as 'indirect' and 'direct' approaches. Indirect approaches report fairness considerations alongside the economic analysis, allowing for discrete comparisons within the final fairness informed economic evaluation. Examples of indirect approaches are various types of Multi-Criteria Decision Analysis (MCDA)[10], based on quantitative, qualitative or mixed comparison. By contrast, direct approaches incorporate fairness considerations into the primary economic analysis (Dukhanin et al., 2018).

The ultimate objective of the proposed framework is to incorporate equity concerns into the cross-sectoral cost-effectiveness analysis. For this reason, the focus will be on direct approaches. Among the direct approaches, the two methodologies most commonly implemented according to the review by Dukhanin et al. (2018) are equity weighting and distributional cost-effectiveness analysis (DCEA).

[10] MCDA is based on the definition and measurement of the 'criteria' (what is being taken into account) and the 'weights' to be put on each element. By combining the criteria scores using the weighting scheme, MCDA can produce an overall assessment of each intervention and has the ability to capture, in a systematic manner, a broader range of policy outcomes (Marsh et al., 2012). MCDA can thus make the impacts of criteria and weights on the decision explicit, but it cannot avoid the need for value judgements (Devlin & Sussex, 2011).

Equity weighting is an instrument that can be used to perform an equity trade-off analysis, where the aim is to quantify trade-offs between improving total health and other equity objectives (Cookson et al., 2017). Alternatively, DCEA can be employed to conduct an equity impact analysis. Such an assessment aims at quantifying the distribution of costs and effects by equity-relevant variables (Cookson et al., 2017).

Equity trade-off and equity impact analysis are illustrated in sub-sub-sections 3.3.3.1 and 3.3.3.2, respectively.

3.3.3.1. Equity trade-off analysis

One way of introducing vertical equity into the evaluation is to use 'equity weights' to value health benefits that apply to people with characteristics that reflect considerations of fairness (Cookson et al., 2017). Attaching greater weights to some populations over others implies ensuring greater access to scarce resources. A more equitable distribution of health can therefore be achieved relative to the dominant health maximisation approach implied in the acceptance of the assumption that "a QALY is a QALY" (Round & Paulden, 2018).

If equity weights are introduced, it becomes necessary to also identify who bears the opportunity cost of the additional weight given to the beneficiaries' health. In fact, if equity weights are applied only to the (identifiable) beneficiaries of a treatment, and not to the (similar but unidentifiable) bearers of the opportunity cost, the principle of horizontal equity would be violated (Paulden et al., 2014).

Equity weighting

On the basis of vertical equity considerations, the general public might assign a different value to health impacts falling on a specific subgroup of the population. For this reason, different weights could be associated with specific health variations. However, if vertical equity concerns are introduced in economic evaluations, horizontal equity issues should not be neglected.

The method proposed by Round and Paulden (2018) can be used to conduct an analysis that incorporates equity weights, but also complies with both horizontal and vertical equity principles. Their formulation allows analysts to incorporate differential weights associated with specific health impacts, and also reflects the weights in the opportunity costs. Required variables are defined as described in Table 3.3, where both effects on health and health care costs are expressed in QALYs.

Variable	Definition
Q_{gs}	QALYs gained by individuals with special characteristics
Q_{go}	QALYs gained by other individuals
Q_g	Total weighted QALYs gained
Q_{ds}	QALYs displaced in individuals with special characteristics
Q_{do}	QALYs displaced in other individuals
Q_d	Total weighted QALYs displaced
ω	Weight applied to QALYs of individuals with special characteristics

Table 3.3 Variables and definitions

As illustrated in Section 3.2.3.1, from the perspective of the health care system, the INHB of an intervention can be calculated as the difference between the direct impact on health and the health opportunity costs (i.e. what health could be generated if those resources were used elsewhere). Opportunity costs depend on the amount of direct costs falling on the HC budget, and the marginal productivity of the HC sector. In other words, opportunity costs are the QALYs that would be lost as a result of the resources displaced in order to accommodate the additional intervention's costs (Drummond, M. et al., 2015).

The formula to calculate the INHB was reported in Equation 3.1. Based on this equation, total QALYs gained and displaced described in Table 3.3 can be computed as shown in Equation 3.9.

$$Q_g = \Delta H \qquad\qquad Q_d = \frac{\Delta C_H}{k_H} \tag{3.9}$$

If equity weights are introduced, because the group from whom QALYs are disinvested also includes individuals with special characteristics, the total weighted QALYs gained and displaced can be computed as illustrated in Equation 3.10.

$$Q_g = \left(\omega * Q_{gs}\right) + Q_{go} \qquad\qquad Q_d = (\omega * Q_{ds}) + Q_{do} \tag{3.10}$$

To implement Equation 3.10, the components of QALYs gained and displaced by individuals with special characteristics and other individuals must be distinguished.

If total QALYs gained are greater than those displaced, then the corresponding Incremental Net weighted Health Benefit (INwHB) is positive, and the intervention is cost-effective. This can be expressed as shown in Equation 3.11.

$$INwHB = \omega * \left(Q_{gs} - Q_{ds}\right) + \left(Q_{go} - Q_{do}\right) > 0 \tag{3.11}$$

It must be stressed that the traditional INHB formula illustrated in Equation 3.1 and consequently the INwHB formula of Equation 3.11 do not actually reflect the full impact of the intervention on health care. These formulae do not include the spill-over effects on health (described in Section

3.2.3.4). The operationalisation of this method with the inclusion of the full impacts on health and its practical issues will be illustrated in detail in Chapter 7, using the case study.

3.3.3.2. Equity impact analysis: Extended and Distributional CEA

A potential alternative to the equity trade-off analysis is the equity impact analysis, which quantifies the distribution of costs and effects by equity-relevant variables. The aim of such an analysis is to use CEA to examine the distribution of DE and OC, broken down by one or more variables of concern to policy makers from an equity perspective. The two main analytical techniques available in the literature are Extended CEA (ECEA) and Distributional CEA (DCEA) (Cookson et al., 2017).

The ECEA approach permits production of breakdowns of costs, health benefits and financial risk protection benefits (i.e. prevention of illness-related impoverishment) by socioeconomic quintile group. Because of its focus on risk protection benefits, ECEA has mostly been implemented in low- and middle-income countries. (Verguet, Kim & Jamison, 2016). On the other hand, the focus of the proposed framework is on (potentially unfair) health inequality impacts and quantification of trade-offs between increasing total health and reducing health inequity, in a context of high-income countries with universal health systems and fixed health budgets (Cookson, 2016). Therefore, the instrument chosen to perform the equity impact analysis is DCEA, which is a framework for incorporating health equity concerns into the economic evaluation of health sector interventions (Asaria, Griffin & Cookson, 2016, Asaria et al., 2015).

DCEA methods

DCEA provides supplementary information about the health inequality impacts of the interventions evaluated. Moreover, it can provide decision makers with evidence on the nature and size of any trade-offs between improving total population health and reducing unfair health inequality (Asaria, Griffin & Cookson, 2016).

The analysis requires the estimation of the baseline health distribution (usually in terms of QALYs), and the modelling of changes to this baseline distribution due to DEs and OCs of the interventions being compared. Resulting modelled health distributions can then be adjusted for alternative social value judgments about fair and unfair sources of health variation (Asaria, Griffin & Cookson, 2016). As illustrated in the illustrative example proposed in Figure 3.2, in a DCEA the baseline health distribution and the new distribution after the intervention are compared. This allows a 'starting point' to be established, given by the baseline health distribution, and assess the impact on health inequalities compared to this.

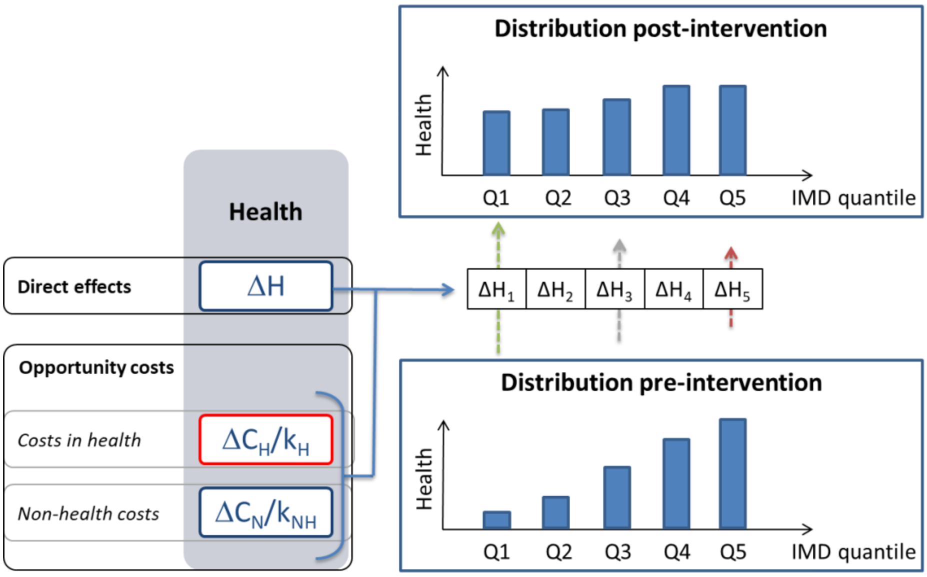


Figure 3.2 DCEA from the health care perspective

Figure 3.2 is an adaptation of the cross-sectoral scheme presented in Figure 3.1. Here the focus is only on the health care sector, and the impacts in terms of DEs and OCs are disaggregated according to the IMD quantiles. As a result, the distributions of health across the population before and after the intervention can be compared. In this illustrative example, greater health benefits are experienced by more disadvantaged individuals (green dashed arrow) compared to a smaller impact falling on richer individuals (grey dashed arrow) and negative health benefits experienced by the individuals in the highest IMD quantile (red dashed arrow).

The methodological advantage of the DCEA is that it does not prescribe in advance any particular set of social value judgments about health inequality. As a consequence, it is suitable for an analysis based on a social decision making approach. Of course, in the implementation phase it is necessary to make some social value judgments (e.g. the definition of unfair dimensions of inequality, and the nature and strength of inequality aversion). Nevertheless, these social value judgments are explicit and transparent in the analysis, and the sensitivity of conclusions based on alternative plausible social value judgments can be investigated (Asaria, Griffin & Cookson, 2016).

Presenting the results of the DCEA

In order to make trade-offs between cost-effectiveness and the alternative health equity objective (i.e. reducing inequality in lifetime health) explicit, results of the DCEA can be illustrated using the Health Equity Impact Plane (Figure 3.3). The vertical axis of the plane shows the net health impact (i.e. the cost-effectiveness) of the intervention, whereas the horizontal axis summarises its net health equity impact (Cookson et al., 2017, Asaria, Griffin & Cookson, 2016).

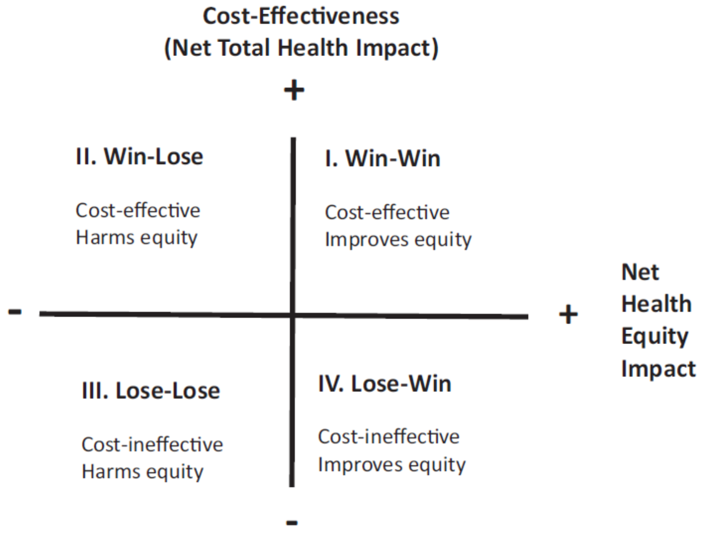


Figure 3.3 Health Equity Impact Plane, from Cookson et al. (2017)

The net equity impact can be assessed using a formal health equity metric that combines disaggregated information in a summary index: the Atkinson index of inequality *A(ε)* (Atkinson, A. B., 1970). When applied to health, Atkinson's measure of inequality measures the fraction of population health that society would be willing to sacrifice to achieve full equality in health, given the current health distribution and the degree of inequality aversion in the society (Norheim, 2012). The Atkinson index of inequality depends on the value of the inequality aversion parameter ε, and is scaled from 0 to 1 (where 0 represents no inequality and 1 represents full inequality).

To calculate A(ε), it is necessary first to compute the Equally Distributed Equivalent Health (EDEH), which combines the chosen measure of inequality with average health in the distribution. The EDEH summarises the health distribution as one number, representing the amount of health that each person in a hypothetically perfectly equal health distribution would need to have for the decision maker to be indifferent between the actual health distribution analysed and the perfectly equal health distribution (Asaria, Griffin & Cookson, 2016).

EDEH and A(ε) can be calculated as shown in Equation 3.12.

$$A(\varepsilon) = 1 - \frac{EDEH}{\bar{h}} \qquad EDEH = \left(\frac{1}{N}\sum h_i^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}} \qquad (3.12)$$

Where: $\bar{h}$ is the mean health in the general population; *N* is the total population size; h_i the individual health in QALE; and ε is the Atkinson inequality aversion parameter.

A decrease in A(ε) represents an improvement in equality. In order to resemble the cost-effectiveness plane, the sign of net equity impact is typically reversed. In the health equity plane, increases in -A(ε) represent an improvement in health equality (Love-Koh et al., 2018).

3.3.4. Incorporating health equity concern in the cross-sectoral analysis

Analytical techniques illustrated in Section 3.3.3 focus on health benefits and costs falling on the health sector budget, and can be used to aggregate outcomes within one dimension (i.e. the health care sector), with the incorporation of equity concerns. Basically, these methods would allow for revising the CEA from the health care sector perspective (Section 3.2.3.1), by incorporating alternative value judgments and relaxing the assumption that "a QALY is a QALY".

The aggregation within the health dimension, with the potential incorporation of considerations about health inequalities, is therefore achievable. Nevertheless, the objective of the proposed framework is to incorporate health equity concerns into the cross-sectoral analyses, in order to inform decision making from an integrated and broader perspective. This requires also, an aggregation across dimensions.

To aggregate across dimensions, the same cross-sectoral methods (illustrated in Section 3.2.3) can be used if either equity weighting or DCEA are conducted in the analysis from the health care perspective. The same formula shown in Equation 3.8 can be used, but it will be based on different criteria for aggregation within the health care sector.

If equity weighting is used, the adaptation of the cross-sectoral formula is straightforward. As shown in Equation 3.13 (which is an adaptation of Equation 3.7), Total Incremental Net weighted Monetary Benefit (INwMB) will be the result of the aggregation of INwHB, INNB and INCB.

$$INwMB = V_H * INwHB + V_N * INNB + INCB \qquad (3.13)$$

If a DCEA from the health care perspective is conducted, it would be necessary to use the EDEH to adjust the results of the analysis from the health care sector perspective. Practical implementation of the methods will be conducted in Chapter 7.

3.4. Implementation of the framework: comments and case study

The previous sections of this chapter illustrated the main stages of the theoretical analytical framework proposed for the evaluation of public health interventions. From the fourth chapter onwards, a case study will be used as a vehicle to illustrate the practical implementation of the framework.

This last section summarises the proposed framework and introduces the case study that will be used as a vehicle to illustrate the proposed methods.

3.4.1. A pragmatic approach to the cross-sectoral analysis

As mentioned in Sub-section 3.2.2, the proposed approach is adapted from Walker et al. (2019)[11]. However, the main objective of the present work was to develop a pragmatic framework, following and applying the principles and guideline by Walker et al. (2019) in the context of a real decision making process. Consequently, the framework proposed in this thesis is more prescriptive, less abstract, and less compatible with the use of alternative analytical techniques.

Table 3.4 summarises the analytical techniques that make up the proposed framework and can be employed to address the cross-sectoral, health equity and cross-temporal stages of the economic evaluation. The case study will be then used to illustrate how to operationalise the cross-sectoral and health equity analyses. The operationalisation of the cross-temporal extension of the analysis is discussed in Chapter 8.

[11] As pointed out previously, the framework was originally formulated by Claxton et al. (2007), and subsequently developed by Claxton et al. (2010), Sculpher et al. (2014) and Walker et al. (2019).

Stage	Analytical technique
Cross-sectoral analysis	***Impact inventory*** Quantification of DEs and OCs for each dimension ***Cost-effectiveness analyses*** Independent CEAs for each sector, including spill-over effects ***Aggregation*** Within-dimension approach. WTP used to aggregate across dimension ***Resources allocation scheme*** Compensation scheme with transfers across sectors to address resource allocation problems
Health equity analysis	***DCEA or equity weighting*** Incorporation of health equity concerns in the analysis
Cross-temporal analysis	***Extrapolation*** Setting a more appropriate horizon (probably beyond the available clinical evidence) for the estimation of costs and consequences of the intervention ***Impact of budgetary policies*** Assessing the impacts of 'soft' and 'hard' budgets (i.e. you can or cannot run deficits)

Table 3.4 Summary of the analytical techniques included in the framework

3.4.2. Identification of the case study

In order to show the full potential of the proposed framework, the objective of the selection process was to identify a public health intervention with a range of cross-sectoral costs and outcomes, with potential long-term impacts, and preferably affecting a disadvantaged subgroup of the population. The York trials Unit register[12] was checked in order to identify suitable trials which were already completed. Moreover, other potential studies were assessed through collaboration with a Public Health Research Consortium (PHRC) project (2011-2019), which aimed to develop a context for economic evaluations of public health programmes with costs and effects falling outside the NHS and local authority public health budgets.

Once the selection process was concluded, the Screening and Intervention Programme for Sensible drinking (SIPS) (Drummond, C. et al., 2014, Newbury-Birch et al., 2014, Kaner et al., 2013) appeared to be the most suitable study for the planned analysis and was therefore selected as the case study.

[12] Available at: https://www.york.ac.uk/healthsciences/research/trials.

3.4.2.1. Screening and Intervention Programme for Sensible drinking

SIPS had the objective of tackling the problem of alcohol misuse, which is a major public health issue in Britain, a country with one of the highest binge drinking rates in the world (WHO, 2014).

The impact of alcohol misuse may also extend beyond health consequences. First of all, problematic alcohol consumption is associated with criminal activity and reoffending; for example it was found that prisoners who had drunk alcohol every day shortly prior to custody were more likely to be reconvicted within a year following release than those who had not (Ministry of Justice, 2013). Moreover, interventions aimed at reducing alcohol misuse might be associated with a wide range of other non-health consequences, such as workplace and productivity losses, road traffic accidents, with the impacts often far in the future (Bhattacharya, 2017, Barbosa, Godfrey & Parrott, 2010).

Furthermore, alcohol programmes might have a significant role in addressing health inequalities. Such interventions target groups on the edges of society, which are deemed to be 'marginalised' through addiction and alcohol misuse (Bramley et al., 2015), and it has been found that the greatest effects are experienced by groups with lower SES (Burton et al., 2017).

3.4.2.2. SIPS trials

The SIPS programme comprised of three cluster randomised controlled trials (RCTs) across three settings: Emergency Departments (ED), Primary Care (PC) and Probation Services (or Criminal Justice Setting, CJS). The three trials had a similar design and aimed at evaluating a brief alcohol intervention.

All SIPS trials could have been eligible as case study for this analysis. Nevertheless, the first two trials were primarily based within the NHS, while the latter was based in the criminal justice system and provides an example of an evaluation of interventions with costs and consequences in different sectors. Alcohol misuse among offenders on probation has been found to be an important issue, and such an intervention is likely to have a significant impact on criminality. Among offenders with a community or suspended sentence order, just over one-third (35%) were identified as having an alcohol misuse need. Among offenders under post-release supervision, 16% were assessed as having an alcohol misuse need (Ministry of Justice, 2013).

Given the specific focus of this thesis on cross-sectoral and long-term impacts and health inequalities, the alcohol intervention implemented in the criminal justice setting appeared to be the most suitable to be used as a case study. Details of the SIPS trial conducted in the CJS are provided in the next sub-section.

3.4.3. SIPS trial in Probation Setting

3.4.3.1. Study design

The SIPS trial in CJS is a prospective factorial pragmatic cluster RCT with randomization at the level of offender manager (OM) (Newbury-Birch et al., 2014).

Table 3.5 summarises the characteristics of the trial.

Trial characteristic	Definition	Description
Prospective	The trial follows the individuals over time	Outcomes measured at baseline, 6 and 12 months
Factorial	The study allows investigators to assess more than one intervention in a single experiment (Montgomery, Peters & Little, 2003)	3 interventions are assessed: CIL, BA, and BLC
Pragmatic	Trials is designed to evaluate the effectiveness of interventions in real-life routine practice conditions (Patsopoulos, 2011)	Follow-up counselling and return visits were voluntary
Cluster	Trials where research subjects are not allocated to treatments independently, but as a group (Bland, 2004)	The target population was composed of offenders recruited by OMs

Table 3.5 SIPS trial characteristics

3.4.3.2. Intervention arms

OMs were randomised to three intervention arms. Each intervention arm was an intensified version of the previous one.

1. Client Information Leaflet (CIL), also called Minimal intervention, used as a benchmark.
   In the CIL control condition, an information booklet was given to the participant. The leaflet contained information about alcohol and included the Drinkline telephone number and contact details of local alcohol treatment agencies.
2. Brief advice (BA) and CIL.
   OMs randomised to deliver BA received training and gave up to five minutes of simple structured advice to the eligible offender. Offenders in this condition also received a CIL from the OM at the end of the advice.
3. Brief Lifestyle Counselling (BLC), BA and CIL.
   OMs randomised to deliver BLC, in addition to providing the BA and CIL, were trained to refer offenders to an Alcohol Health Worker (AHW). The AHW delivered a 20 minute brief

lifestyle counselling intervention to offenders who attend the appointment at the probation office.

3.4.3.3. Population and clusters

The target population were 525 offenders recruited by 131 OMs from three English regions (North East, London and South East). Characteristics of the population were presented in the clinical paper, and are reproduced in Appendix A.

Table 3.6 summarises the number of OMs (i.e. level 2 clusters) and individuals (i.e. level 1 units) in each intervention allocation according to both the Intention to Treat (ITT) and the per-protocol criteria. The ITT criterion identifies the number of offenders in each group according to their allocation to each intervention; the per-protocol criterion is based on the actual compliance of the offender, which is if he or she actually received the intervention.

Unit	CIL	BA	BLC	Total
OM	47	35	49	**131**
Offender (ITT)	184	178	163	**525**
Offender (per-protocol)	184	173	67	**424**

Table 3.6 SIPS population: level 1 and level 2 units

The number of OMs randomised to CIL, BA and BLC were 47, 35 and 49, respectively. The number of offenders allocated to CIL, BA and BLC were 184, 178 and 163, respectively. Nevertheless, checking the actual compliance, in the BA group, 173 of 178 received the BA; in the BLC group, 162 received BA but only 67 subsequently received BLC.

3.4.3.4. Outcomes

The primary clinical outcome collected was the self-reported hazardous or harmful drinking status, using the Alcohol Use Disorders Identification Test (AUDIT) at 6 months. Secondary outcomes were AUDIT at 12 months, alcohol problems questionnaire, readiness to change, EQ-5D-3L, visits to NHS and other resources use.

3.4.3.5. Main results from the previously published trial analysis

Newbury-Birch et al. (2014) did not report any significant differences in the collected primary and secondary outcomes[13]. The authors however highlighted the relevant role of re-offending and planned a subsidiary CEA using CJS data, in order to assess whether the choice of principal outcome measure altered the recommended decisions. However, no evidence was found that this analysis had been conducted or published (Newbury-Birch et al., 2014).

With regard to the CJS data, the authors found that over the next 12 months, offenders treated with BA and BLC were significantly less likely to be reconvicted. Odds of receiving a conviction were significantly lower for BA (36%) and BLC (38%), compared to CIL (50%). Nevertheless, according to the authors, the main limitation of this result was that reduction in criminal recidivism was found in the absence of significant differences in drinking consumption between the groups, which was the primary outcome of the study (OR of AUDIT negative status in BA group compared to CIL was 0.80; 95% CI: 0.39, 1.62. OR of AUDIT negative status in BLC group compared CIL was 0.73; 95% CI: 0.34, 1.53) (Newbury-Birch et al., 2014).

Nevertheless, the authors pointed out that even in the absence of differences in alcohol consumption across interventions, increased awareness might have resulted in a change in offending behaviour rather than consumption *per se*, or offending might be linked to particular patterns of drinking rather than overall consumption.

3.4.4. Rationale for conducting an economic evaluation of SIPS

Clinical trials are generally assessed on the basis of statistical significance, whereas economic evaluations assess both costs and effects and estimate when an intervention is cost-effective and how uncertain that estimate is. Statistical methods employed therefore differ between the two types of analysis (Raftery et al., 2015).

A CEA from the health care perspective should still be performed even if clinical studies fail to demonstrate a statistically significant difference in clinical end points, and demonstrable clinical benefit is not a prerequisite for economic evaluation (Ramsey et al., 2015, Whitehurst & Bryan, 2013). An economic evaluation can provide useful information regarding the budget impact

---

[13] Similar results were found in other settings. The effectiveness of the intervention was assessed also in ED and PC settings, and CIL turned out to be the better option (Drummond, C. et al., 2014, Newbury-Birch et al., 2014, Kaner et al., 2013).

implications of the trial. It may happen that large clinical benefits turn out to be too expensive, and on the contrary small clinical benefits may be cost-effective (Whitehurst & Bryan, 2013).

For these reasons, in the next chapter, a 'traditional' CEA of SIPS based on health-related costs and outcomes will be conducted. Clinical results of the trial previously reported will be revisited, and used in the economic evaluation of SIPS.

Chapter 4: CEA of SIPS from the health care perspective

4.1. Introduction to the case study

The second chapter of this dissertation explored the theoretical foundations and analytical techniques for the economic evaluation of public health interventions. The third chapter covered the description of the proposed comprehensive and theoretical framework of analysis. Moreover, the case study was presented. The aim of this and subsequent chapters is to operationalise the theoretical framework using the case study as a vehicle to illustrate the proposed methods.

In this chapter, the cost-effectiveness of SIPS is investigated from the health care perspective. The objective of a CEA is to inform decisions by assessing and comparing incremental costs and incremental effects of an intervention. Because the analysis is conducted from the health care perspective, only health care costs and effects of the intervention are assessed. In this analysis, only data observed within the trial period are used.

It is well recognised this 'traditional' CEA shows several limitations in informing decisions, especially when dealing with public health interventions (Sculpher et al., 2014). For this reason, in Chapter 5 and Chapter 6, the perspective of the economic evaluation will be extended beyond the narrow health care viewpoint. In the fifth chapter, another CEA from the CJS perspective will be conducted, using reductions in reconviction frequency as a measure of effectiveness. Afterwards, in Chapter 6, the CEA from the health care perspective will be adjusted to also include the spill-over effects from criminal justice on victims' health. In the same chapter, a CEA will be conducted from a broader and integrated perspective, including both impacts on HC and CJS. In Chapter 7, the effects of incorporating health equity concerns in the cross-sectoral analysis will be investigated. Lastly, the potential cross-temporal extension of the evaluation will be discussed in Chapter 8.

4.1.1. Structure of the chapter

Section 4.2 describes the methods employed for the analysis. Sections 4.3 and 4.4 cover the analysis of health and health care costs, respectively. A CEA under the complete case approach is conducted in Section 4.5. The impacts of missing data and outliers are assessed in Sections 4.6 and 4.7, respectively. A summary of results of the economic evaluation and policy implications are reported in Section 4.8.

4.2. Methods for the CEA from the health care perspective

4.2.1. Effectiveness assessment

The impact of the intervention on health was measured using QALYs. QALYs combine quantity and health-related quality of life (HRQoL) and are calculated by weighting each year of life lived using a quality adjustment weight. The quality weights represent the HRQoL of the health states under consideration and are based on individuals' preferences for the health states. On the conventional scale for QALY weights, zero corresponds to death and one to perfect health. The bottom of the scale is not well defined and states worse than death are allowed (Karimi & Brazier, 2016, Drummond, M. et al., 2015).

In the UK, the preferred generic measure of HRQoL impacts of health interventions is the EQ-5D (NICE, 2013). EQ-5D is a pre-scored multi-attribute health status classification system, with a questionnaire that categorizes individuals' health on 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. In the SIPS trial, the EQ-5D-3L questionnaire was used, where each dimension has three levels of severity, yielding up to 245 mutually exclusive health states (3^5 plus dead and unconsciousness). The questionnaire is presented in Appendix B. A more recent five-level version was developed in 2011 (when the trial was already concluded), but the three-level questionnaire is still the most commonly used (Oemar & Oppe, 2013).

HRQoL scores collected during the trial were converted into single-dimensional weights by applying a scoring function that reflects the general population preferences for the health states assessed by the EQ-5D instrument. In the UK, the preferences for the scoring function were measured in 1994 with the time trade-off (TTO) valuation technique on 3395 members of the UK public (Dolan, 1997, MVH Group, 1995). The resulting EQ-5D tariff is a set of numbers ranging from -0.594 to 1, where 1 is full health, 0 is death, and negative values are valued as worse than death.

The single-dimensional generic HRQoL weights obtained were used to calculate QALYs. QALYs can be obtained by drawing a graph with time on the x axis and HRQoL weights on the y axis, and calculating the area under the curve (AUC). Equation 4.1 describes the general expression for calculating QALYs from individual patient data, using individuals' scores at different time points (Hunter et al., 2015, Manca, Hawkins & Sculpher, 2005).

$$QALYs = \sum_{t=0}^{n} \left[\frac{(Q_1 + Q_{t+1})}{2} * \frac{(T_{t+1} - T_t)}{T} \right] \qquad (4.1)$$

In Equation 4.1, n is the number of measurements over the study period; Q_t is the individual EQ-5D score value obtained in the t^th^ measurement; T is the total duration of the study period expressed as the total number of time units in a year; T_t is the time period in which the t^{th} measurement takes place, again expressed as number of time units in a year (Manca, Hawkins & Sculpher, 2005). Equation 4.2 illustrates the application of the generic formula to this specific case.

$$QALYs = \frac{EQ5D_{Baseline} + EQ5D_{Month\,6}}{2} * \frac{6-0}{12} + \frac{EQ5D_{Month\,6} + EQ5D_{Month\,12}}{2} * \frac{12-6}{12} \qquad (4.2)$$

QALYs were discounted according to appropriate measures of time preferences (Drummond, M. et al., 2015), and regression methods were used to adjust differences in QALYs for potential differences in baseline health states between individuals or groups (Manca, Hawkins & Sculpher, 2005). Because of the clustered design of the study, the Huber–White sandwich estimation approach was employed when assessing the differences in QALYs in order to derive robust standard errors (SEs). Clustered SEs account for heteroscedasticity across clusters of observations (Cameron & Miller, 2015).

4.2.2. Costs assessment

Resource use data were collected alongside clinical data during the SIPS trial. Questionnaires were administered at baseline, and six and twelve month follow-ups. Unit costs were then associated with resource use data. On the basis of surveys' date of completion, 2010 was used as reference year for unit costs and all costs are presented in 2010 pounds.

Intervention costs were also calculated. These included staff costs and cost of leaflets. Appropriate intervention costs were linked to each individual, according to the actual delivery of the programme.

4.2.3. Assessment of cost-effectiveness

A complete case analysis was performed. ICERs, which represent the ratio of the difference in costs and the difference in outcomes between the trial arms, were presented and compared to alternative threshold estimates. Threshold estimates were based on values of £13,000 (Claxton, Martin, et al., 2015), £20,000 and £30,000 per QALY (NICE, 2012). INBs were presented as well. Uncertainty around the cost-effectiveness estimate was assessed using cost-effectiveness (CE) planes and cost-effectiveness acceptability curves (CEACs). CEACs were constructed to present the probability that each of the interventions (BA, BLC and CIL) is cost-effective according to different monetary values being attached to QALYs.

Lastly, even though the following analyses were not the main focus of the thesis, the impact of missing data and the appropriateness of building a multilevel model were also assessed. To overcome the missing data issue, Multiple Imputation (MI) was undertaken. Data after MI were bootstrapped to account for the expected skewness evident in economic cost data, and another CEA was conducted using the new dataset with imputed values. Analysis was performed using STATA 14 (StataCorp, 2015) and Microsoft Excel 2010 (Microsoft, 2010).

4.3. Effectiveness of SIPS: impacts on health

4.3.1. Analysis of EQ-5D scores

Participants were asked to complete the EQ-5D questionnaires before starting the trial, and twice over the duration of the trial, at 6 and 12 months. First, EQ-5D scores were calculated. A graphical description using box plots[14] is reported in Figure 4.1. Table 4.1 summarises the EQ-5D scores at baseline, 6 and 12 months for the three intervention allocations.

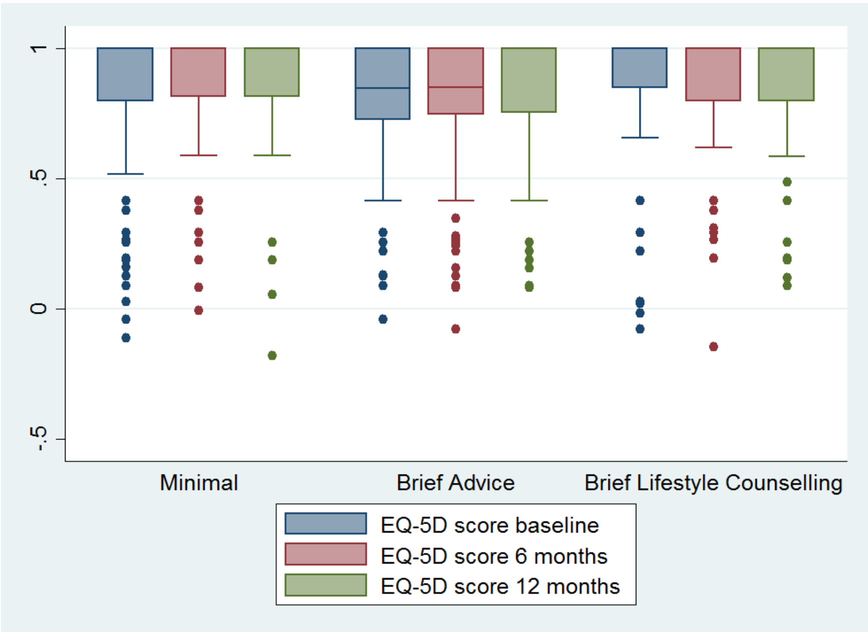


Figure 4.1 EQ-5D scores at baseline, 6 and 12 months

[14] Boxes capture the lower (first) and upper (third) quartiles, and the line subdividing the box represents the median; whiskers describe the distribution of all data within 1.5 interquartile ranges (i.e. the difference between the third and first quartile); dots capture outliers.

Intervention		Baseline	Month 6	Month 12
Minimal	Mean	0.836	0.873	0.881
	SD	0.243	0.208	0.199
	Count	178	116	101
Brief Advice	Mean	0.824	0.823	0.836
	SD	0.222	0.250	0.234
	Count	171	127	112
Brief Lifestyle Counselling	Mean	0.880	0.868	0.855
	SD	0.205	0.206	0.220
	Count	157	105	97

Table 4.1 EQ-5D scores at baseline, 6 and 12 months

Overall, it seems that the Minimal intervention performs slightly better compared to more intense interventions. The BA group is characterised by a slightly higher variability in the scores compared to other treatment allocations. Box plots suggest that EQ-5D scores in the Minimal and BA groups tend to remain stable over time, whereas a slight decrease in the scores can be observed in the BLC group.

Baseline EQ-5D scores appear to be different in the three treatment allocations. Such differences in initial health levels must be taken into consideration when calculating the differences in health scores. In order to estimate the differences in EQ-5D scores adjusted for the baseline scores, two regressions were performed. Month 6 and 12 EQ-5D scores were regressed on treatment allocation (using Minimal intervention as the benchmark) and EQ-5D baseline scores. Offenders were recruited by OMs and similarities between individuals recruited and treated by the same OM could be observed. For this reason, SEs of the regressions are robust to clustering. Results are reported in Table 4.2.

	EQ-5D scores at 6 months			EQ-5D scores at 12 months		
	Coefficient	95% CI *		Coefficient	95% CI **	
Intervention allocation:						
Brief Advice	-0.0396	-0.0889	0.0098	-0.0351	-0.0863	0.0160
Brief Lifestyle Counselling	-0.0315	-0.0755	0.0124	-0.0440	-0.0976	0.0096
Baseline EQ-5D	0.5507	0.4260	0.6755	0.5380	0.3894	0.6866
Constant	0.4108	0.2858	0.5359	0.4257	0.2815	0.5699

* Note: SE adjusted for 114 clusters

** Note: SE adjusted for 109 clusters

Table 4.2 Regression coefficients for EQ-5D at 6 and 12 months

Even though differences among treatment allocations are not statistically significant at 5% level, results show a negative correlation between levels of health and more intense interventions (BA and BLC compared to CIL).

4.3.2. Analysis of QALYs

Total QALYs resulting from the programme were computed using the AUC method (illustrated in Sub-section 4.2.1). QALYs were plotted using box plots for each intervention group, and differences in QALYs were assessed. Box plots are presented in Figure 4.2.

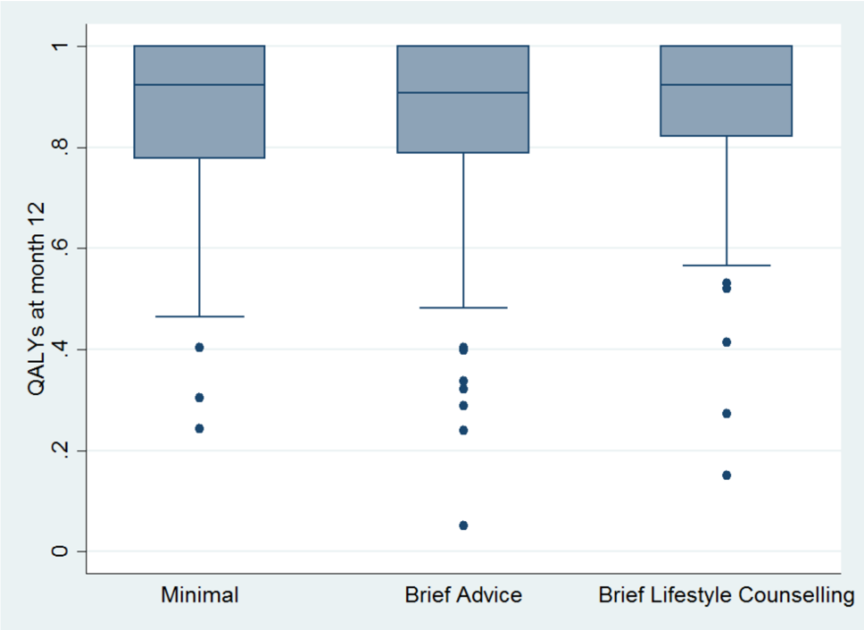


Figure 4.2 Distribution of QALYs at month 12

From the box plots it appears that QALYs distributions in Minimal and BA groups are very similar, with BA having a slightly lower median value. BLC group shows less variation in QALYs and slightly higher QALYs. Nevertheless, differences in QALYs might be biased by differences in health at baseline. The recommended method to investigate differences in QALYs between interventions and the role of baseline EQ-5D scores is to regress total QALYs on intervention allocation and baseline EQ-5D scores (Hunter et al., 2015).

Equation 4.3 illustrates the regression equation used, where total QALYs are the dependent variable, b_1 is the coefficient for the treatment effect and b_2 is the coefficient for baseline EQ-5D scores (Manca, Hawkins & Sculpher, 2005).

$$\text{QALYs} = b_0 + b_1 \text{ Intervention} + b_2 \text{ EQ5D}_{Baseline} \tag{4.3}$$

However, randomisation is not a perfect process and differences in QALYs might be biased also by differences in baseline characteristics between groups. For this reason the role of other baseline variables was investigated by including them in the regression as well[15].

4.3.2.1. Two potential regression models for QALYs

Two regression models were tested[16]. In regression Model 1, QALYs were regressed on intervention allocation (BA and BLC compared to Minimal) and adjusted for baseline EQ-5D scores differences. In regression Model 2 it was tested whether other baseline characteristics had an impact on QALYs. Age, gender and ethnicity were thus included in the regression, as well as the baseline measure of the drinking status (AUDIT score) and the baseline readiness to change. Regressions coefficients and their 95% confidence intervals are reported in Table 4.3.

	Model 1			Model 2		
	Coefficient	95% CI		Coefficient	95% CI	
Intervention allocation:						
Brief Advice	-0.0230	-0.0600	0.0141	-0.0237	-0.0608	0.0133
Brief Lifestyle Counselling	-0.0277	-0.0564	0.0011	-0.0291	-0.0589	0.0008
Baseline EQ-5D	0.6752	0.5942	0.7561	0.6536	0.5619	0.7453
Age				-0.0009	-0.0022	0.0005
Sex (1=male)				-0.0197	-0.0597	0.0204
Ethnicity:						
Black				-0.0011	-0.0510	0.0488
Asian				0.0413	-0.0127	0.0952
Mixed				-0.1364	-0.2816	0.0089
Other				0.0380	-0.0312	0.1071
Baseline AUDIT score				-0.0005	-0.0029	0.0019
Baseline RCR*:						
Sometimes think about drinking less				-0.0281	-0.0663	0.0101
I have decided to drink less				-0.0144	-0.0628	0.0340
Already trying to cut down				-0.0247	-0.0584	0.0090
Constant	0.2960	0.2171	0.3750	0.3602	0.2314	0.4890
R-squared	0.6299			0.6505		
AIC	-388.975			-378.382		
BIC	-374.810			-328.915		
Note: SE adjusted for 104 clusters * Readiness to Change Ruler						

Table 4.3 Regression estimates for QALYs

[15] On the other hand, adding covariates might generate issues due to small numbers and potential collinearities.
[16] Generalized linear models were also considered and Park test was performed to assess the appropriate model form. Test results indicated that Gaussian distribution provides the most appropriate fit to the data.

In both models baseline EQ-5D scores have a positive and statistical significant impact on the outcome of interest. As regards the other additional controls, it seems that none of them has a relevant impact on total QALYs, and including or excluding them from the regression does not affect coefficients and CIs of the intervention allocation and baseline EQ-5D. R squared values are comparable between the two models (though slightly higher in model 2), whereas AIC and BIC suggest that Model 1 performs better than Model 2. For this reason, Model 1 was chosen.

4.3.2.2. Average QALYs per intervention allocation

From Model 1, the coefficients associated with BA and BLC are -0.0230 (95% CI: -0.0600; 0.0141) and -0.0277 (95% CI: -0.0564; 0.0011) respectively. As discussed in Sub-section 4.2.1, QALYs are built on a scale going from zero to one, where zero corresponds to death and one to perfect health. Consequently, while these differences are small, they still equate to some meaningful drop in health status. In fact, these coefficients mean that on average BA is associated with a reduction of approximately 8 days in perfect health in a year, compared to the Minimal intervention; BLC is associated with a reduction of approximately 10 days in perfect health, compared to the Minimal intervention.

Using the results from the regression, it was then possible to estimate the predicted QALYs means adjusted for the differences in baseline health scores. As the discount rate is generally only applicable in cases where the duration of the trial is greater than 12 months, QALYs were not discounted in this analysis (Hunter et al., 2015). As illustrated in Table 4.4, a hypothetical individual having a 0.846 EQ-5D baseline score (which is the average EQ-5D score at baseline across all individuals) would experience 0.8670, 0.8441 and 0.8394 QALYs (i.e. 316, 308, and 306 days in perfect health) if treated for one year with CIL, BA and BLC, respectively. Differential QALYs calculated from the adjusted mean QALYs correspond to the regression coefficients.

Intervention	Observations	Unadjusted mean	Adjusted mean	Adjusted estimates 95% CI	
Minimal	83	0.8661	0.8670	0.8435	0.8905
Brief Advice	91	0.8351	0.8441	0.8158	0.8724
Brief Lifestyle Counselling	81	0.8672	0.8394	0.8214	0.8574

Table 4.4 QALYs adjusted for baseline EQ-5D scores

The number of total observations (255) is lower than the number of total patients (525). This is due to the considerable presence of missing observations. The issue of missing data will be explored via multiple imputation using the costs data in Section 4.6.

4.3.2.3. The impact of clusters

An additional analysis was conducted in order to assess the impact of clusters and investigate whether building a multilevel model was worth it. Alternative regression models were compared and the actual impact of clusters on QALYs turned out to be negligible. Multilevel regressions provided almost the same results as OLS regressions, and the Intracluster Correlation Coefficients (ICCs) were estimated to be zero. Details of this analysis can be found in Appendix C.

Given the negligible role of the clusters, it was concluded that performing a multilevel analysis would have only added layers of complexity (especially in performing missing data analysis) that would have not improved the quality of this case study.

4.4. Health care resource use analysis

4.4.1. Health care units

Table 4.5 summarises the health-related categories of resource use and associated unit costs, as they were recorded in the SIPS questionnaire. In the table, 'PSSRU' refers to the publication by Curtis (2013) on Unit Costs of Health and Social Care, and NHS reference costs are taken from the costs report published by the Department of Health and Social Care (2010). Costs reported in years different from the reference year (2010) were adjusted using the Bank of England Inflation Calculator[17].

Unit	Details	Cost (£)	Source
Treatment for drinking problems			
Drinkline calls	No information available in the literature. Here it is assumed these are 10-minute calls with an alcohol related worker (£33 per hour)	5.5	PSSRU
Counselling[18]	Cost per consultation by Alcohol Related Workers (i.e. experienced nurses who have undertaken specific training in counselling people who misuse alcohol), including qualifications	40	PSSRU
Day care[19]	No information available in the literature. Same costs as outpatient treatment for drinking problems were assumed	87	
Detoxification[20]	Cost per patient day of inpatient detoxification for people who misuse drugs/alcohol	142	PSSRU

[17] Available at: https://www.bankofengland.co.uk/monetary-policy/inflation/inflation-calculator
[18] Aggregated counselling in Hospital, Primary care, Day care programme, Alcohol Treatment facility, Voluntary.
[19] Aggregated day care in Hospital, Day care programme, Alcohol Treatment facility, Voluntary.
[20] Aggregated detoxification in Hospital, Primary care, Day care programme, Alcohol Treatment facility, Voluntary.

Unit	Details	Cost (£)	Source
Outpatient treatment for drinking problems[21]	Outpatient attendances, drug and alcohol consultant services (follow-up face to face)	87	PSSRU
Overnight hospital detoxification	No information available in the literature. Same costs as detoxification were assumed	142	
Overnight after care hostel	Hostel: £107.45 per week = 15.35 per day in 2012. Adjusted to 2010	14.14	Shelter (2012)
Overnight alcohol treatment facility	No information available in the literature. Same costs as detoxification were assumed	142	
Overnight residential programme	Temporary accommodation in housing association stock: £87 per week (£12.43 per day). 2012 costs adjusted to 2010	11.45	Shelter (2012)
Use of hospital services (excluding treatment for drinking problems)			
A&E visits leading to not admitted	Weighted average of attendance of A&E treatments leading to not admitted	97	PSSRU
A&E visits leading to admitted	Weighted average of attendance of A&E treatments leading to admitted	131	PSSRU
Inpatient nights	Cost per day calculated using the average cost per episode for elective/non-elective inpatient and the average length of stay	477	NHS Reference Costs
Outpatient visits	Weighted average of all outpatient procedures	152	PSSRU
Day case visits	Weighted average of all stays	637	PSSRU
Emergency ambulance travels	NHS Trusts Paramedic Services: Emergency Transfers / Urgents (PSETU). National Average Unit Cost	245.7	NHS Reference Costs
Ambulance travels	NHS Trusts Patient Transport Services (PTS): average of Admitted Patient Care (£40), Outpatient (£31) and Other (£26). National Average Unit Cost	32.3	NHS Reference Costs
Private travels	Average transportation cost per visit used for GP home visit	5	PSSRU
Primary care and social services			
GP visits	Sum of surgery consultation costs (including qualification costs, excluding direct care staff costs) and actual prescription costs per consultation (£39)	71	PSSRU
Nurse visits	Nurse (GP). Cost per consultation, including qualifications	12	PSSRU
Social worker visits	Costs per hour, including qualifications. From PSSRU it was not possible to estimate a cost per visit. Assumed 60 min per visit	53	PSSRU
Home Visit: GP	Cost per home visit, excluding direct care staff costs, including qualification costs, including actual prescription costs per consultation (£39)	147	PSSRU
Home Visit: Nurse	Nurse (GP). Cost per home visit, including qualifications	20	PSSRU
Home Visit: Community psychiatric nurse	Community nurse. Cost per home visit, including qualifications	27	PSSRU
Home Visit: Others (e.g. dietitian)	Costs per hour of home visiting, including qualifications. From PSSRU it was not possible to estimate a cost per visit. Assumed 60 min per visit	57	PSSRU
Home Visits: Social worker	No information available in the literature. Same costs as social worker visits were assumed	53	PSSRU

[21] Aggregated outpatient treatment in Hospital, Alcohol Treatment facility.

Unit	Details	Cost (£)	Source
NHS Direct	Average cost of a call to NHS Direct line estimated in 2007 and adjusted to 2010	27.63	(Hansard, 2009)
NHS walk in	Walk in services	37	PSSRU

Table 4.5 Unit costs of health care resource use

Accident and emergency (A&E) resource units were slightly adapted with respect to the way they were recorded in the SIPS questionnaire. Originally one of the resource use measures was the number of nights spent in Admission and Emergency Department (AED). Nevertheless, the values recorded were frequently not consistent with the information about inpatient nights spent as an inpatient from the same individuals, and there were some cases of double counting. A new rule was therefore implemented: A&E events were distinguished into A&E visits leading to not admitted and A&E visits leading to admitted. For those individuals reporting A&E visits leading to admission, all the additional nights after the first spent in AED were moved to inpatient nights (if not already recorded there). This categorisation is also more consistent with PSSRU estimates categories. Average prescription costs were included in visit tariffs.

4.4.2. Costs per resource use category

Unit costs were assigned to the health care resource consumption data in order to estimate the costs of the health care resource use. Health care resource use data are reported in Appendix D.

Average costs per group for each cost category are described in Table 4.6. Estimates are based on the total resource use over the whole follow-up, which is the sum of resource use of the first six months (recorded at month 6) and the second six months (recorded at month 12). By aggregating resource use data collected at the two time points, only complete cases are kept; if the resource use at either month 6 or 12 was missing, the aggregate resource use was recorded as missing as well.

Cost category	CIL	BA	BLC
Drinkline calls	0.9 (8.3) 86	0 (0) 100	0 (0) 87
Counselling	46.1 (232.7) 86	79.2 (491.7) 100	71.5 (259.2) 87
Day care	5 (27.8) 86	0 (0) 100	10 (76.7) 87
Detoxification	1.6 (15.3) 86	4.2 (31.6) 100	14.6 (137) 87
Outpatient treatment for drinking problems	2 (13.1) 86	0.8 (8.7) 100	10 (75.5) 87
Overnight hospital detoxification	0 (0) 87	0 (0) 100	27.7 (258.8) 87
Overnight after care hostel	0 (0) 87	0 (0) 100	0 (0) 87
Overnight alcohol treatment facility	1.6 (15.2) 87	0 (0) 100	0 (0) 87
Overnight residential programme	0.1 (1.2) 87	0 (0) 100	0 (0) 87
A&E visits leading to not admitted	26.4 (70.1) 88	49.9 (97.5) 101	36 (84.2) 86
A&E visits leading to admitted	8.9 (33.2) 88	12.9 (77.5) 101	12.1 (55.4) 86
Inpatient nights	37.5 (193.2) 89	318 (1985.9) 102	126.1 (501.1) 87
Outpatient visits	162.2 (665.5) 89	162.4 (490.6) 102	109.8 (290.4) 86
Day case visits	21.4 (115.6) 89	99.9 (769.7) 102	29.6 (134.9) 86
Emergency ambulance travels	93.8 (556.2) 89	21.6 (109.8) 102	45.1 (203.6) 87
Ambulance travels	0.7 (6.8) 89	0.3 (3.1) 102	0 (0) 86
Private travels	0.8 (8.4) 89	0.1 (0.8) 102	1.7 (16.1) 86
GP visits	197.1 (298.5) 89	314.9 (599.7) 100	211.4 (328.3) 86
Nurse visits	12.6 (30.2) 89	8.8 (24.2) 100	6.4 (16.8) 86
Social worker visits	25 (124.5) 89	42.7 (224) 101	11 (50.1) 86
Home Visit: GP	1.6 (15.5) 89	10.1 (63.2) 101	10.2 (66.8) 86
Home Visit: Nurse	2 (19) 89	0.9 (8.1) 101	0.6 (6.4) 86
Home Visit: Community psychiatric nurse	0 (0) 89	6.8 (40.1) 101	0.3 (2.9) 86
Home Visit: Others	1.9 (18.1) 89	1.6 (17) 101	0 (0) 86
Home Visits: Social worker	8.4 (45.3) 88	49.2 (206) 100	12.4 (42.7) 84
NHS Direct events	3.1 (11.4) 88	3.2 (12.5) 101	2.9 (13.5) 84
NHS walk in events	4.7 (13.7) 85	15.6 (40.7) 95	6.1 (21.3) 86

The table reports: mean cost in 2010 £ (SD[22]) number of observations, including imputed values

Table 4.6 Health care costs under the mean imputation scenario

Looking across the three intervention groups, the main factors contributing to total costs are outpatient visits, emergency ambulance travels, GP visits and A&E visits. Moreover, counselling and inpatient nights play a fundamental role in determining the additional costs for the BA group.

4.4.3. Intervention costs

The three interventions are described in Section 3.4.3. All OMs had to provide the information leaflet, which is a common cost for CIL, BA and BLC groups. The 5-minute brief advice from the

[22] Standard deviations (SD) reported in the tables are based on Normality assumption (which is not realistic in this context) and may lead to negative resource use. SDs here should be interpreted only as measure of uncertainty and are not used in models or simulations.

trained OM was planned only for OMs and offenders allocated to BA and BLC. The 20-minute brief lifestyle counselling from AHW was undertaken only if the offender was randomised to the BLC group. Resource use categories and their unit costs are reported in Table 4.7.

Resource use	Cost (£)	CIL	BA	BLC
Information leaflet	0.2	X	X	X
5-minute brief advice from trained OM	2		X	X
20-minute brief lifestyle counselling from AHW	11			X
Total intervention cost per individual (£)		**0.2**	**2.2**	**13.2**

Table 4.7 Intervention costs

The cost of the information leaflet was estimated as the price of printing a 16-page stapled booklet (from an online order[23] of 1,000 booklets). The cost of the 5-minute session was calculated from the cost per hour (24£) of a generic one-to-one session of an intervention for substance misuse (including qualifications costs). The PSSRU estimates of unit costs in criminal justice were used (Brookes et al., 2013). The cost of the 20-minute session with an AHW was calculated using the information reported in Table 4.5.

Costs were assigned according to the actual compliance to the intervention on the basis of the information reported in Table 3.6 about who actually received the intervention.

4.4.4. Costs and decision makers' budgets

For the sake of simplicity, frequently in CEAs all costs are aggregated. Here, on the contrary, particular attention was devoted to distinguishing costs falling on different decision makers' budgets, in order to assess cross-sectoral costs correctly. Health-related cost categories were therefore aggregated according to the different budgets on which costs fall.

SIPS costs categories were re-arranged according to areas of expenditure: primary care, secondary care, non-residential alcohol treatment, residential alcohol treatment, social care, other home visits, ambulances, private consumption and CJS. SIPS cost categories and decision makers' budgets are reported in Appendix E. Afterwards, areas of expenditures were associated with the appropriate decision maker's budget.

[23] From: https://www.helloprint.co.uk/

4.4.4.1. Social care and health care costs

In the theoretical framework proposed in Chapter 3, a broader perspective for the evaluation was described. As summarised in the examples in Table 3.1, a list of potential sectors was proposed, based on the 2nd Cost-Effectiveness Panel's impact inventory (Sanders et al., 2016). In this list, Health Care and Social Services were distinguished. From a purely theoretical point of view, it can be argued that health care and social care impacts fall on different budgets.

Nonetheless, from a more pragmatic point of view, the division between health and social care is more blurred. As mentioned in Sub-section 3.2.1.1, in the UK, NICE guidelines recommend the health and social care perspective for the evaluation (NICE, 2013). Consequently, when the analysis is carried out from the health care perspective, impacts of health and social care are typically measured in QALYs and decision makers in health care base their decision on evaluations where health care costs and social care costs are aggregated.

For this reason, this ‘traditional’ CEA will be based on NICE guidelines (NICE, 2013), and a health and social care perspective will be employed for the evaluation. It will be assumed that social care costs (i.e. social worker visits and home visits, community psychiatric nurse home visits, and visits and home visits by ‘others’) fall on the same health care decision maker’s budget, and effects of health and social care are both measured in QALYs.

The main objective of this case study is to illustrate the methods and the policy impacts of conducting a cross-sectoral analysis where budgets, effects and decision makers are clearly distinguishable (e.g. health and criminal justice). Nevertheless, the potential separation of health care and social care budget and its impacts on decision making are worth to be discussed more in detail. Policy impacts, and theoretical and technical issues of implementing the framework using ‘Social Services’ as a separate dimension will be discussed in the discussion Chapter 8, Sub-section 8.3.

4.4.4.2. Decision makers affected by the impacts of SIPS

Four decision makers (with their own budget) were identified: health care (including social care), LA, private consumption and CJS (that paid for the intervention). As illustrated in Table 4.8, it was assumed that CJS paid for the intervention.

Cost category	**Decision maker's budget**
Primary care Secondary care Social care and home visits Ambulances Non-residential alcohol treatment Overnight hospital detoxification	Health care
Overnight after care hostel Overnight alcohol treatment facility Overnight residential programme	Local authority
Private travels	Private consumption
Intervention costs	Criminal justice system

Table 4.8 SIPS costs and associated budgets

Table 4.9 summarises the average total costs of complete case observations (where all the resource use values in all categories at both 6 and 12 months were not missing) falling on each budget, grouped by treatment allocation.

	CIL	BA	BLC
Health care	573.6 (992.6) 75	1316.4 (3032.9) 83	768.7 (1469.5) 77
Local authority	1.8 (16.3) 75	0 (0) 83	0 (0) 77
Private consumption	1 (9.2) 75	0.1 (0.9) 83	1.9 (17) 77
Criminal justice system	0.2 (0) 75	2.1 (0.3) 83	6.7 (5.4) 77
Total costs	576.8 (999) 75	1318.7 (3033) 83	777.4 (1481.5) 77

Table reports: mean cost in 2010 £ (SD) number of observations

Table 4.9 Total costs of SIPS (complete case)

Total costs recorded using the health-related SIPS questionnaire included costs falling on HC, LA, consumption and CJS. Nevertheless, most of the recorded costs do actually fall on the HC budget, and only a very small proportion fall on other sectors. However, true impacts on CJS are much bigger than those reported in Table 4.9. Here only intervention costs are considered. The analysis of the cost-effectiveness of SIPS from the CJS perspective, with the inclusion of the full impacts on CJS, will be conducted in the next chapter.

4.4.5. Health care costs at baseline

In the SIPS questionnaire, health care resource use in the six months before the intervention was also collected. Mean costs at baseline are described in Table 4.10. Only costs falling on health care budget were recorded.

Cost category	Mean (£)	SD
ED patient	64.36	140.70
Nights as inpatients	387.68	2427.06
Admitted but not kept in	26.97	93.27
Outpatient appointments	86.10	299.23
GP visits	173.45	333.66
GP home visits (HV)	7.66	53.93
Practice nurse visits	5.29	34.28
Nurse HV	1.04	11.59
Prescriptions	107.00	273.41
Social worker HV	12.71	74.53
Social worker visit	27.27	215.53
Other HV	19.85	138.29
Other visits	52.71	212.18
Total baseline costs	974.47	2824.09

Table 4.10 Baseline health care costs

Nights as inpatients, GP visits, prescriptions and outpatient appointments are the main health-related cost categories that contribute to the total amount of baseline costs. This information could be potentially used to adjust cost differences of the intervention for differences at baseline. In fact, baseline health care costs could be interpreted as differences among individuals in the use of the resources due to causes not attributable to the intervention. The role of health care baseline costs will be investigated in detail in the next section.

4.5. CEA under the complete case analysis approach

This section illustrates the CEA of the SIPS trial under the complete case analysis approach. The economic evaluation was performed using the original dataset fixed for resource use intensity as illustrated in Appendix D.

Recalling the costs aggregation summarised in Table 4.9, only costs falling on health care and social care budget were considered in this CEA. It would not be correct to include also costs falling on other budgets, even though in this particular case their impact on final results would be likely to be marginal. Costs falling on CJS will be analysed in the next chapter, and a cross-sectoral CEA will be conducted in Chapter 6.

The distribution of health care costs were highly skewed to the right. This can be observed in Figure 4.3 and by comparing mean and median costs in all intervention groups. Median costs reported in Table 4.11 are in fact much lower than mean costs.

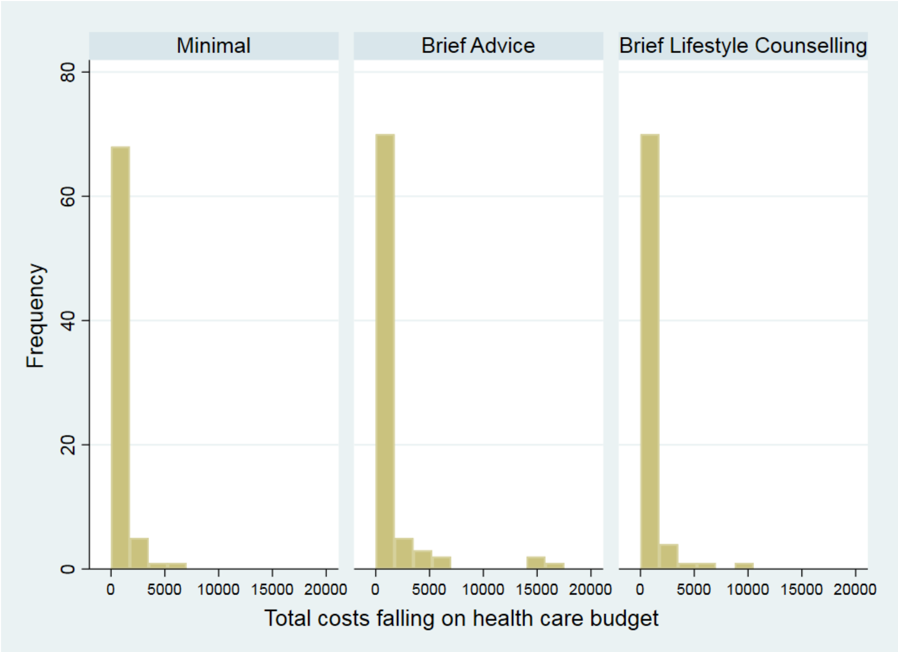


Figure 4.3 Distribution of costs falling on health care budget, by intervention allocation

	Median cost (£)	Mean cost (£)
CIL	173.5	573.7
BA	284	1316.4
BLC	249	768.7

Table 4.11 Median and mean health care costs per group

For this reason, two alternative analyses were explored when performing the CEA: costs and QALYs were analysed both via seemingly unrelated regressions (SURs) (Zellner, 1962) and using two separate regressions. SUR allows estimating not only adjusted mean differences in costs and QALYs, but also the correlation between the errors in costs and QALYs regression equations. Nevertheless, SUR assumes normality of cost data and there is no option to allow errors to be heteroskedastic (Cameron & Trivedi, 2010). Consequently, it might be worth trading off the simultaneous estimation of correlated costs and QALYs in favour of a better fit of the model, using two separate regressions, one for costs and one for QALYs.

4.5.1. Seemingly unrelated regressions for costs and QALYs

First, adjusted mean differences in costs and QALYs were estimated via SURs. Two alternative SUR models were compared. In Model 1, one regression had QALYs as dependent variable, regressed on intervention allocation and baseline EQ-5D scores; the other regression had total costs as dependent variable, regressed on intervention allocation. In an alternative SUR model (Model 2), the impact of baseline health care costs was investigated by including them in the regression having total health care costs as the dependent variable. Regression results are illustrated in Table 4.12.

Model 1: no adjustment for baseline costs						
	QALYs			Costs		
	Coefficient	95% CI		Coefficient	95% CI	
Intervention allocation:						
BA	-0.019	-0.054	0.016	742.714	101.382	1384.046
BLC	-0.020	-0.055	0.016	195.034	-458.051	848.119
Baseline EQ-5D	0.643	0.578	0.707			
Constant	0.321	0.260	0.381	573.689	108.860	1038.518
Model 2: adjusting for baseline costs						
	QALYs			Costs		
	Coefficient	95% CI		Coefficient	95% CI	
Intervention allocation:						
BA	-0.010	-0.046	0.025	373.646	-180.510	927.801
BLC	-0.018	-0.054	0.018	-226.624	-793.303	340.055
Baseline EQ-5D	0.658	0.593	0.723			
Baseline costs				0.613	0.491	0.736
Constant	0.305	0.244	0.366	344.457	-57.371	746.285

Table 4.12 SUR results for QALYs and costs

Test for error independence shows a statistically significant negative correlation between the errors in the two equations. At the same time, the correlation is not particularly strong (r_{12}= -0.197 in Model 1, r_{12}= -0.178 in Model 2, both statistically significant at 1% level[24]), and efficiency gains of using SUR might not be great in comparison to modelling costs and QALYs separately.

Compared to the Minimal intervention, both BA and BLC appear to be less effective. With regard to the impact on costs, BA is also more expensive than CIL. As a consequence BA results to be not a cost-effective alternative to CIL, because it is both more costly and less effective. On the contrary, BLC appears to be potentially less expensive than CIL. This is particularly the case when including baseline costs in the regression model. Therefore, BLC might be a cost-effective alternative to CIL, depending on the value assigned to health.

4.5.1.1. Results of the CEA based on SURs

In the absence of additional controls for potential baseline differences in resource uses in the costs regression, it was decided to include baseline costs in the regression in the primary analysis. However, an alternative analysis without adjusting for baseline cost in the SURs was conducted.

[24] P-values of the Breusch-Pagan test of independence were 0.0025 and 0.0072, respectively.

Results of this alternative analysis are discussed briefly at the end of this section, and full results are reported in detail in Appendix F.

ICERs and INBs

Once differential QALYs and costs (in 2010 £, the reference year for the case study) of BA and BLC compared to CIL were estimated, ICERs could be computed. INBs were calculated as well.

As illustrated in Chapter 3, Section 3.2.3.1, from the health care perspective, the INHB of an intervention is given by the difference between the direct impact on health (ΔH) and the health opportunity costs (i.e. what health could be generated if those resources were used elsewhere). Opportunity costs depend on the amount of direct costs falling on HC budget (ΔC_H), and the marginal productivity of the HC sector (k_H). In other words, opportunity costs are the QALYs that would be lost as a result of the resources displaced in order to accommodate the additional intervention costs. Equation 3.1 showed how to calculate the INHB. By re-arranging Equation 3.1, in Equation 4.4 the same INB is expressed in monetary terms, using the INMB (Drummond, M. et al., 2015).

$$INMB = \Delta H * k_H - \Delta C_H \quad (4.4)$$

QALYs, costs and NMBs associated with BA and BLC, and differential QALYs, costs and NMBs of BA and BLC compared to CIL are presented in Table 4.13, where k corresponds to a marginal productivity of the HC sector of £ 13,000 per QALY.

Treatment allocation	**QALYs (95% CI)**	**Costs (95% CI)**	**NMBs (k = £13,000)**	**Average differential with CIL**			**ICER** *(£/QALY)*
				QALYs	**Costs**	**NMBs**	
CIL	0.864 (0.839; 0.890)	849.0 (446.8; 1251.2)	**10387.0**	-	-	-	-
BA	0.854 (0.830; 0.878)	1222.6 (843.6; 1601.6)	**9877.6**	-0.010	373.6	-509.4	Dominated
BLC	0.846 (0.821; 0.871)	622.4 (227.5; 1017.2)	**10378.5**	-0.018	-226.6	-8.5	12,527

Table 4.13 CEA results based on SURs (adjusting for baseline costs)

In Table 4.13 treatment options are presented in rank order of effectiveness. The highest NBs are associated with CIL, while BA is dominated by CIL, achieving poorer outcomes at higher costs. With regard to BLC, ICER is 12,527 £/QALY, but INMBs are negative. This is because BLC is on average cheaper than CIL, but also less effective[25]. INMBs of BLC are negative also if assuming a *k* that

[25] BLC is therefore cost-effective if its ICER is higher than the cost-effectiveness threshold (Drummond, M. et al., 2015)

reflects more traditional recommendations from NICE (2013) of either £20,000 or £30,000 per QALY. In fact, for these two alternative values of k, INMBs associated with BLC are £ -135 and £ -316, respectively.

CE plane, CEACs and uncertainty around INB estimates

Results can be illustrated also using the cost-effectiveness (CE) plane, presented in Figure 4.4. The CE plane was obtained via bootstrapping and reflects the uncertainty around the estimates. It shows that in most of the cases differential costs and QALYs associated with BA fall in the north-west quadrant, where the incremental cost of BA compared to CIL is positive and the incremental effect is negative. With regard to the comparison of BLC and CIL, most of the estimates fall in south-west quadrant (i.e. comparator is cheaper and less effective). BLC is therefore cost-effective if its incremental cost-effectiveness ratio is higher than the threshold for ICERs (Drummond, M. et al., 2015).

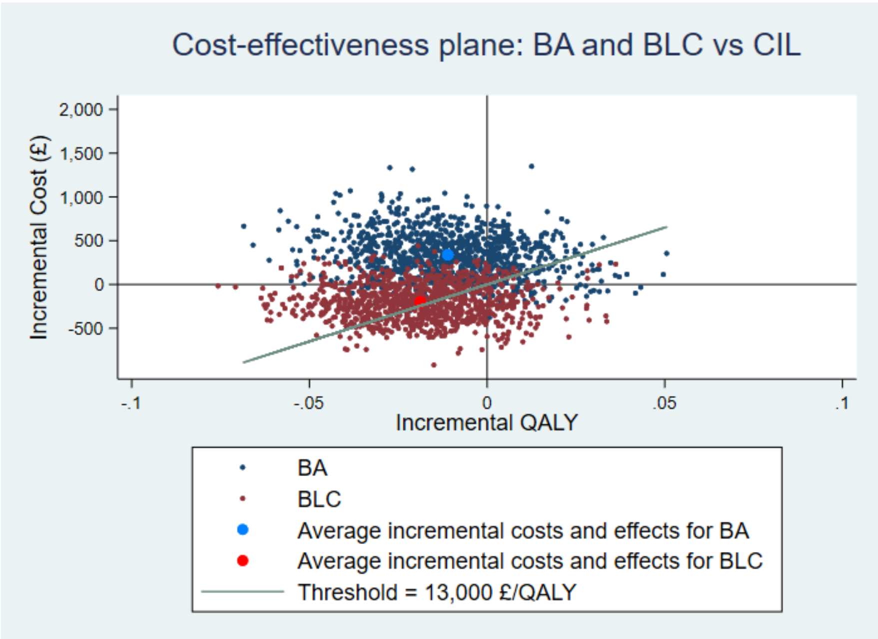


Figure 4.4 CE plane

To reflect the uncertainty around the estimates, the confidence interval around the INBs was estimated, and CEACs were plotted.

The 95% CI around the INMBs was estimated parametrically. As shown in Figure 4.5, even though the expected INMBs associated with both BA and BLC are always negative for thresholds higher than £13,000 per QALY, INMBs associated with BLC are more uncertain and also include positive values within the 95% CI. Such difference is particularly clear when looking at the blue line representing the 95% CI for k equal to 13,000 £/QALY: the interval associated with BLC is much wider.

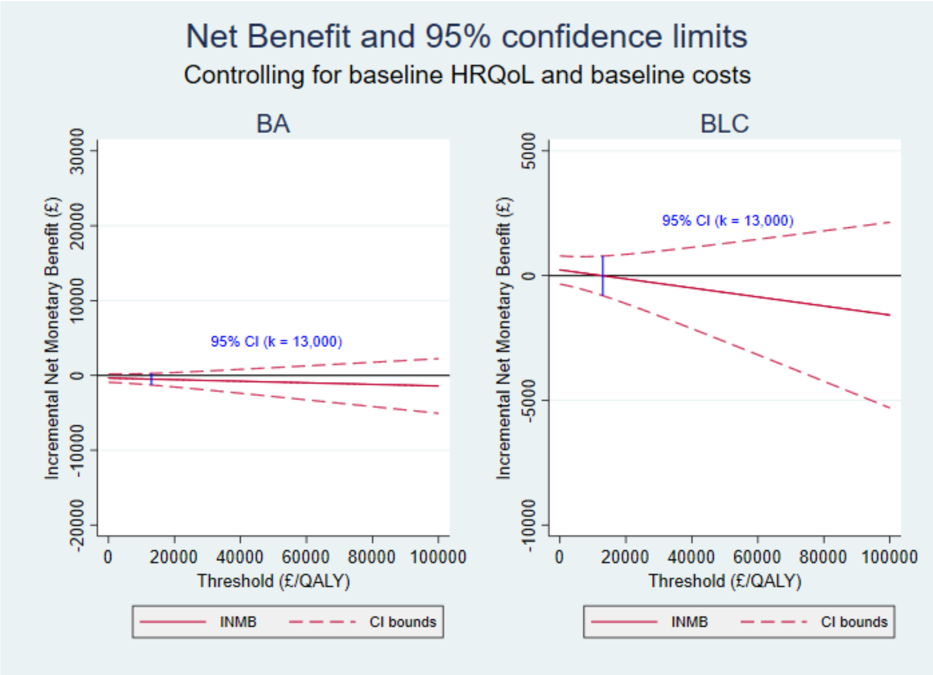


Figure 4.5 INMBs and 95% CIs

CEACs were obtained both parametrically and via bootstrapping. Results were very similar and CEACs based on bootstrapping are presented here; parametric CEACs can be found in Appendix G. As shown in Figure 4.6, the probability that BA is cost-effective is always close to zero. For very low values of the threshold, BLC has the highest probability of being cost-effective. The higher the threshold (i.e. the higher the value assigned to health) the lower is the probability that BLC is cost-effective. Three grey dashed lines represent three potential threshold values of 13,000, 20,000 and 30,000 £/QALYs.

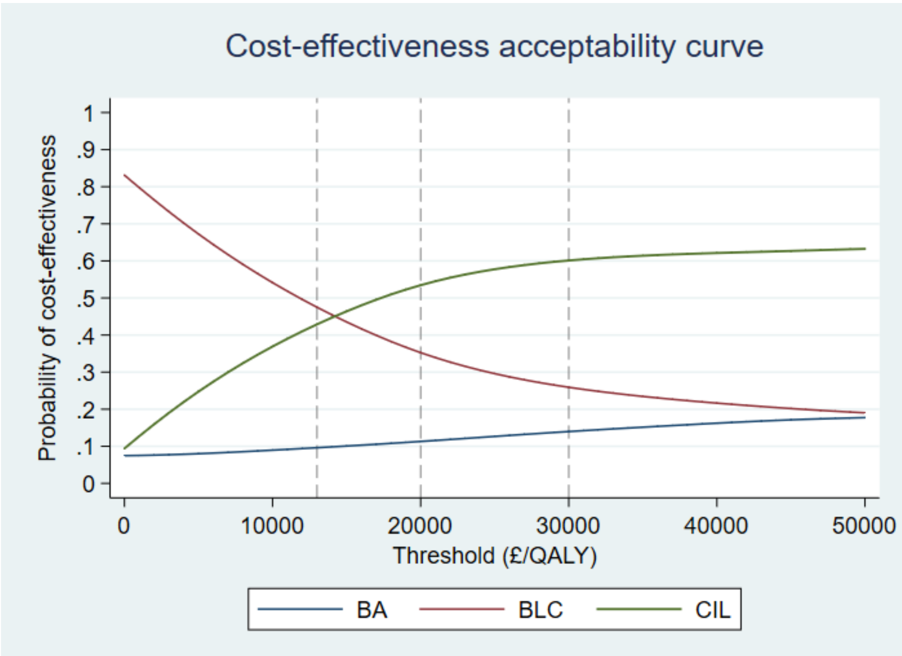


Figure 4.6 CEACs

Figure 4.6 shows that with higher threshold values, CIL has the highest probability of being cost-effective. However, with a threshold of 13,000 £/QALY, BLC is associated with the highest probability of cost-effectiveness (even though really similar to CIL). However, at the same time, negative INBs are associated with BLC even for a threshold of 13,000 £/QALYs. This is because BLC has indeed the highest probability of being cost-effective, but still the (less likely) negative impacts more than offset the (more likely) positive impacts.

Results of the CEA based on an alternative SUR model

As highlighted at the beginning of this section, these results are based on SURs that included baseline health care costs in the costs analysis. If baseline costs are not included in the regression models, it is even more likely for CIL to be cost-effective when compared to both BA and BLC. Results of this alternative analysis can be found in Appendix F.

In brief, without adjusting for baseline costs, CEACs show that both BA and BLC have a very low probability of being cost-effective when compared to CIL, independently of the threshold value. Moreover, looking at the ICERs, both BA and BLC are dominated. In the CE plane, most of estimates of differential costs and QALYs fall in the north-west quadrant, where the comparator (BA or BLC) is dominated by CIL.

4.5.2. Separate regressions for QALYs and costs

As an alternative to SUR, costs and QALYs can be modelled separately. A regression model for QALYs was already illustrated in Section 4.3.2, and regression estimates are already reported in Table 4.3. With regard to costs, alternative generalized linear model (GLM) models were explored. In order to choose the most appropriate family function, a Park test was performed (Jones et al., 2013) and results of the test indicated that Gamma distribution provides the best fit to the data. Tests for the link function also confirmed the appropriateness of using a log-link function. An alternative analysis was therefore conducted, where a log-gamma regression (with robust SE) was employed for costs. Again, the impact of baseline costs was also assessed (excluded in Model 1, included in Model 2). Costs regression outcomes and average marginal effects are summarised in Table 4.14.

Log-gamma regression models for costs						
	Model 1			Model 2		
	Coefficient	95% CI		Coefficient	95% CI	
Intervention allocation:						
BA	0.831	0.251	1.410	0.263	-0.301	0.826
BLC	0.293	-0.223	0.808	0.020	-0.495	0.536
Baseline costs				0.000	0.000	0.001
Constant	6.352	6.036	6.668	6.026	5.592	6.460
AIC	15.508			15.092		
BIC	-934.415			-963.507		
Average marginal effects						
	Model 1			Model 2		
	dy/dx	95% CI		dy/dx	95% CI	
Intervention allocation:						
BA	720.575	32.901	1408.249	3780.621	-12697.420	20258.660
BLC	263.361	-222.606	749.329	289.369	-7189.627	7768.365
Baseline costs				7.053	-28.904	43.010

Table 4.14 Log-gamma regression outputs

The role of baseline costs appears to be less crucial.

CEACs for Model 1 and Model 2 are reported in Figure 4.7 and Figure 4.8, respectively. CEACs show that including baseline costs in the regression (Model 2) only increases the uncertainty around the estimates, but the probability of cost-effectiveness for BA and BLC is always much lower than CIL.

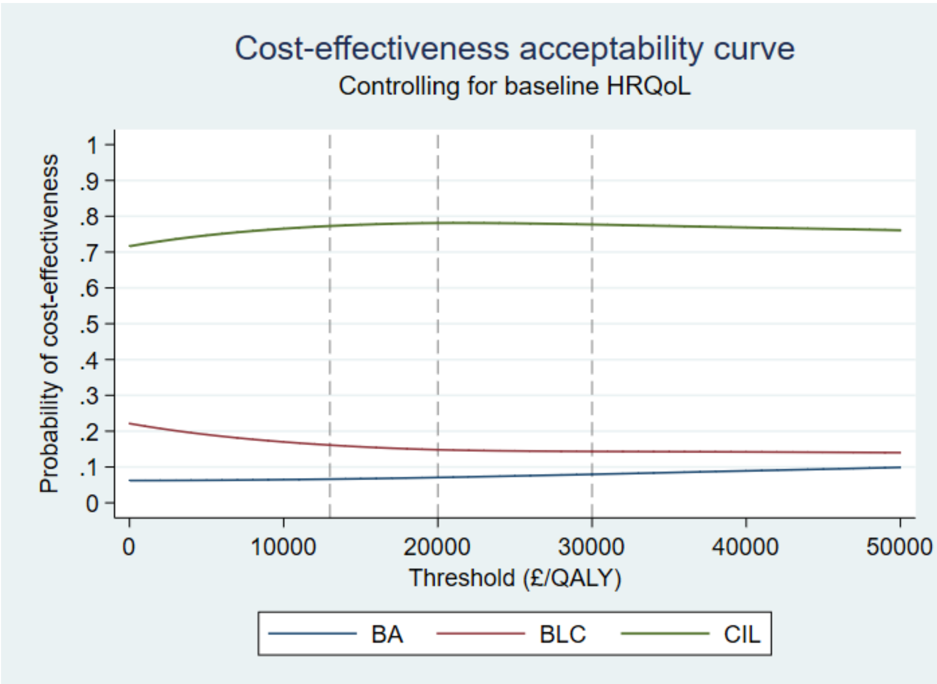


Figure 4.7 CEACs based on separate regressions for costs and QALYs (Model 1)

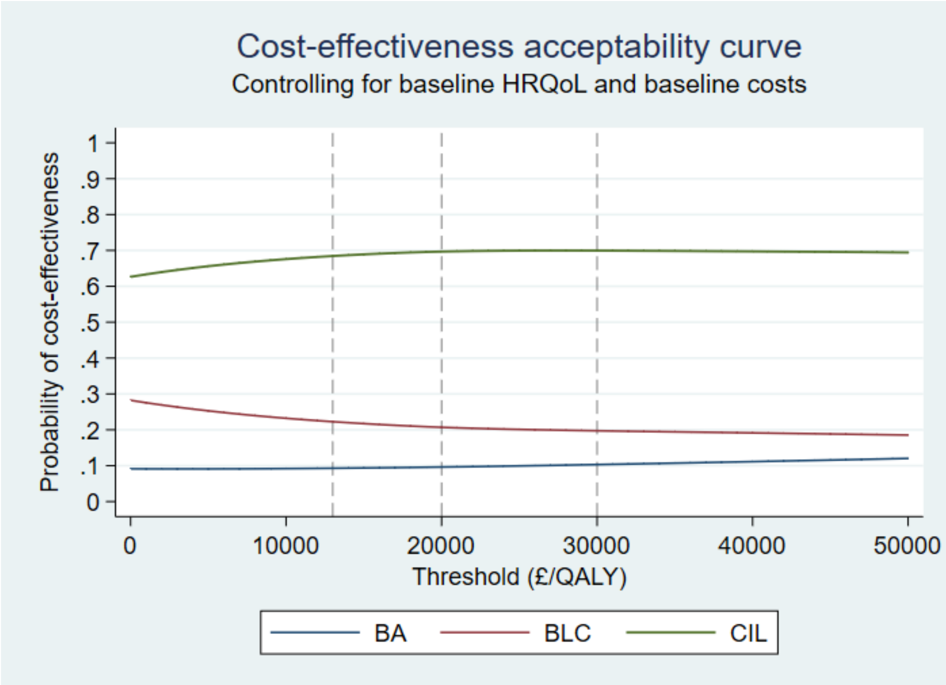


Figure 4.8 CEACs based on separate regressions for costs and QALYs (Model 2)

4.5.3. Conclusions

From the analysis based on the complete case approach, it could be concluded that (from the health care prospective) BA and BLC are not cost-effective compared to CIL.

4.6. Missing data and Multiple Imputation

4.6.1. The role of missing data

Missing data are a frequent problem in CEA conducted alongside RCTs (Faria et al., 2014). In particular, as described in Table 3.5, SIPS trial was a pragmatic study. Attending follow-up visits was entirely voluntary and it is therefore reasonable that only a minority of offenders returned questionnaire replies. For this reason, the amount of missing data is considerable, and its impact must be investigated.

In the presence of missing data, the complete case approach (i.e. ignoring missing data) might produce biased results. This is because costs or health outcomes in individuals with missing data may be systematically different from those with fully observed information (Faria et al., 2014). Inadequate handling of the missing data in a statistical analysis can lead to biased estimates of parameters as well and ultimately can affect the decision of whether an intervention is good value for money (Faria et al., 2014, White, Royston & Wood, 2011).

4.6.1.1. Assumptions about missing data

According to the framework proposed by Little and Rubin (2014), missing data can be categorised as:

- Missing completely at random (MCAR): the probability of data being missing does not depend on the observed or unobserved data;
- MCAR conditional on baseline missing data (CD-MCAR): the probability of data being missing does not depend on the observed or unobserved data, after controlling for baseline variables;
- Missing at random (MAR): the probability of data being missing does not depend on the unobserved data, conditional on the observed data;
- Missing not at random (MNAR): the probability of data being missing depend on the unobserved data, conditional on the observed data.

4.6.1.2. Multiple imputation

Multiple imputation (MI) is a popular statistical technique for handling missing data which uses the distribution of the observed data in order to predict and replace missing observations with a set of plausible imputed values. MI can handle missing data under MAR and can be modified to handle MNAR (Faria et al., 2014, White, Royston & Wood, 2011).

When missing values occur in several variables, MI by chained equations (MICE) can be used. MICE generates imputations based on a set of imputation models, one for each variable with missing values. Imputed values in one variable are used to predict missing values in other variables in an iterative way until the model converges to a stable solution (Faria et al., 2014, White, Royston & Wood, 2011).

In the following section, SIPS missing data are analysed and a MI procedure based on MICE is conducted.

4.6.2. MI procedure

4.6.2.1. Descriptive analysis of SIPS missing data

Only a few values were missing in the baseline variables. On the contrary, approximately one third of the effectiveness outcomes were missing at month six, and roughly 40% were missing at month twelve. The amount of missing data for cost variables was considerable, and ranged from more than 30% at month 6, to around 40% at month 12. Roughly half of the total costs were missing.

Details of the descriptive analysis of missing data are reported in Appendix H.

Aggregation of cost categories

There were too many cost variables with too few observations to run the MI model. For the MI model to converge, it was necessary to reduce the number of cost categories drastically. Costs aggregation aimed at merging similar cost categories at 6 and 12 months, separately.

It must be stressed again that only costs falling on the health and social care budget were considered here, and costs related to overnight after care hostels, overnight alcohol treatment facilities, overnight residential programmes and private travels were excluded from this analysis.

Three groups were created: treatment for drinking problems, hospital and AED, visits and home visits. Components of each grouping are illustrated in Table 4.15.

Treatment for drinking problems	Hospital and AED	Visits and home visits
Drinkline calls	A&E visits (not admitted)	GP visits
Counselling	A&E visits leading to admitted	Nurse visits
Day care	Inpatient nights	Social worker visits
Detoxification	Outpatient visits	Home Visit: GP
Outpatient treatment for drinking problems	Day case visits	Home Visit: Nurse
Overnight hospital detoxification	Emergency ambulance travels	Home Visit: Community psychiatric nurse
	Ambulance travels	Home Visit: Others
	NHS Direct events	Home Visits: Social worker
	NHS walk in events	

Table 4.15 Grouping of cost categories before MI

Missing data patterns

Figure 4.9 shows that over time EQ-5D scores exhibit an intermittent missing pattern: there are offenders who did not report the outcome at time *t*, but they did report it at time *t+1*. Information about QALYs is missing when EQ-5D score is missing at baseline or at 6 or 12 months. Missing data patterns for the grouped cost at both month 6 and 12 are reported in Figure 4.10 and exhibit the same intermittent pattern.

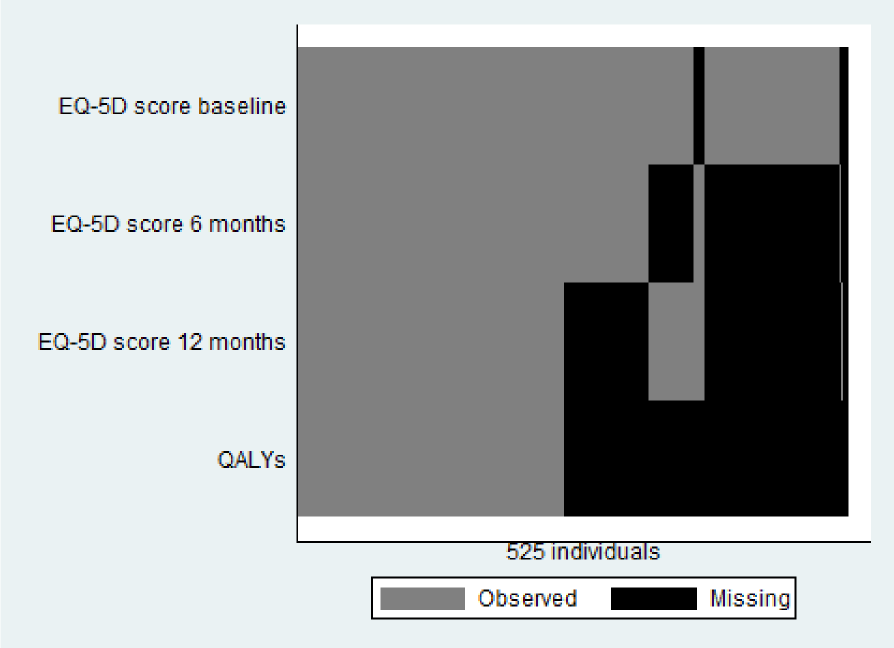


Figure 4.9 Missing data patterns for health outcomes

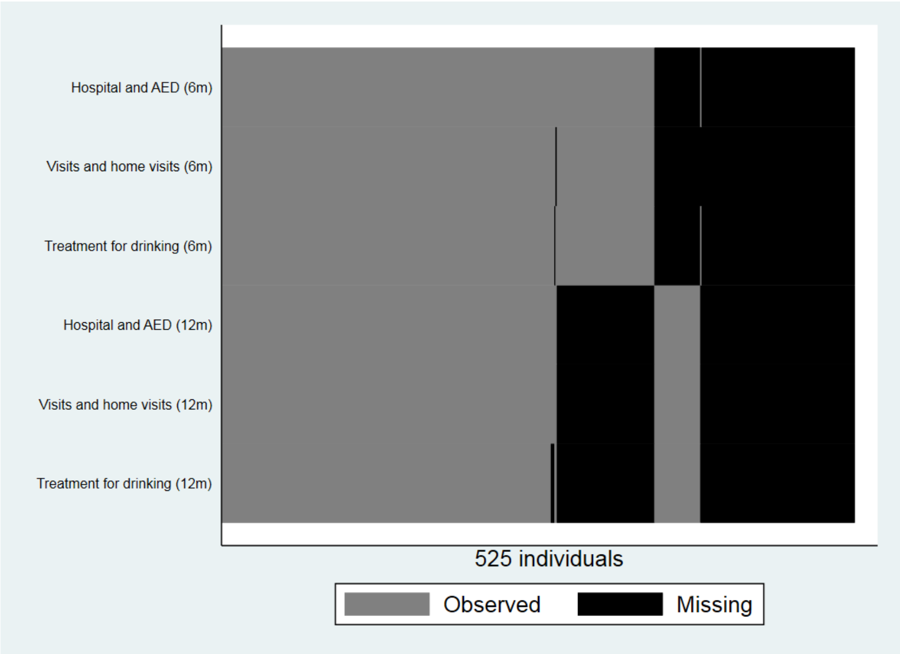


Figure 4.10 Missing data patterns for costs at month 6 and 12

4.6.2.2. Associations with missing outcomes

Logistic regressions were used to investigate the association of baseline and outcome variables on the probability of missingness. Indicators of missingness were regressed first on baseline characteristics. Afterwards other observed covariates were included in the regressions. Full regression results are reported in Appendix H.

Testing for MCAR

Firstly, the dependence of health outcomes missingness on baseline variables was tested. Age was found to have a positive statistically significant (at 5% level) impact on missingness of health outcomes at month 6 or 12. Moreover, AUDIT scores at baseline were significantly and negatively associated with missingness of outcomes at month 6, and high values of baseline readiness to change were significantly positively associated with missing data at 12 months.

While it appears reasonable that older offenders are less responsive to such an intervention, the impact of drinking status and readiness to change is less clear. Regressions show a positive correlation between missingness and readiness to change (more likely to be missing when more willing to change behaviour), and a negative correlation of missingness with AUDIT score (more likely to be missing if the individual had lower scores at baseline in the questionnaire about alcohol use disorders). These findings might suggest a lower engagement with the programme for individuals with lower levels of alcohol dependence and more inclined to change their behaviour.

With regard to the missingness of cost variables (i.e. hospital and AED costs, visits and home visits costs, and treatment for drinking problems costs) at 12 months, it was found that they were all significantly associated with age, sex and having high levels of readiness to change.

These results highlight that data are unlikely to be MCAR. This information would support both CD-MCAR and MAR assumptions.

Testing for CD-MCAR

The subsequent step was to check whether the probability that data are missing depends on observed baseline covariates, but is independent of the missing and observed outcome (i.e. data are CD-MCAR).

Firstly, the association between missing health endpoints and other outcomes was tested. An indicator of missing EQ-5D at 12 months was regressed on previously observed EQ-5D scores and previously observed AUDIT, alcohol related problems – measured using the Alcohol Problems Questionnaire (APQ) – and readiness to change scores. Regression results showed that other than age, readiness to change at 6 months could have a role in predicting missing health outcomes data.

Secondly, indicators of missing health care costs at 12 months were regressed on costs at 6 months, and it was found that previously observed costs related to the treatment of drinking problems were significantly and negatively associated with missingness of cost data. This finding is very reasonable and highlights that the more it was spent in treating the individuals, the less likely it was to having missing values at 12 months.

Lastly, it was investigated whether costs could have been predictors of health outcomes missingness. A statistically significant negative association between missingness of health outcomes at 12 months and visits and home visits costs at 6 months was found. In other words, as it could be expected, the more resources were spent in visits and home visits, the less likely it was to have missing values in health outcomes at 12 months.

These findings show that data are not CD-MCAR.

Conclusions

Given the previous results, data were assumed to be MAR. It is thus assumed that the probability of data to be missing does not depend on the unobserved data, conditional on baseline and observed data (Faria et al., 2014).

Nevertheless, such an assumption cannot be demonstrated using the data collected within the study (Mason et al., 2018, Faria et al., 2014). For example, it has been claimed that if data are based on self-reported outcomes, patients in relatively poor health may be less likely to complete the requisite questionnaires. Outcome data may be therefore MNAR (Mason et al., 2017).

The assumption that data are MAR is therefore arguable and the true data generating process is unknown. However, the potential MNAR nature of the data is not demonstrable. Moreover, from a pragmatic point of view, modelling MNAR data is not the main focus of this analysis, and would require more advanced methods (e.g. a Bayesian framework incorporating expert opinion) that go beyond the purpose of this analysis (Mason et al., 2018). It is however acknowledged that if data were actually MNAR, not only would both the MI and complete case approach be biased, but also MI might yield results that are more biased than those under complete case approach. This is the case even with very strong correlations between fully observed variables and variables with missing values, such that the data are very nearly MAR (Pepinsky, 2018).

4.6.2.3. Imputing missing values in the baseline variables

Some missing values were found in the baseline variables. This could affect the analysis because baseline variables must be used as predictors of missingness in the MI process. Following Faria et al. (2014), mean values were imputed to replace missing values in baseline variables. Such operation ensures that imputed values are independent of the treatment allocation and prevents the creation of additional covariate imbalance (Faria et al., 2014). This was performed for age, degree, education, AUDIT score, readiness to change and baseline costs. As regards baseline EQ-5D, to ensure not to lose any information available, mean imputation was undertaken for each of the five missing individual EQ-5D scores at baseline, and not for the total score. Some replies to the health questionnaires were in fact partially complete, even though the overall EQ-5D scores were missing.

4.6.2.4. Specification of the MICE model

Given the previous results, missingness of health outcomes and health care costs was assumed to depend on baseline covariates (and in particular on: age, ethnicity, sex, baseline AUDIT score, and baseline RCR), readiness to change at 6 months and other observed health outcomes and costs at 6 and 12 months. The MI model was thus constructed and included all baseline variables as predictors, and the following incomplete variables to be imputed: EQ-5D scores at 6 and 12 months, health care costs at 6 and 12 months and RCR at 6 months. Observed values of these incomplete variables were also used as predictors for each other. Total health care costs and QALYs were re-calculated passively using imputed values.

As a rule of thumb, White et al. (2011) suggested to create *m* imputed datasets such that *m* is equal to the percentage of incomplete cases. In this analysis the proportion of missing values ranged from around 31% to 41%, and *m* was thus set to 40. Basically, 40 imputations by treatment allocation were conducted.

Ordered logistic regression and predictive mean matching (PMM) were used to predict ordinal and non-Normal continuous variables, respectively. PMM is a method of imputing missing values using observed values from other similar individuals; this ensures that imputed values are plausible (e.g. positive costs and HR-QOL always lower than one) (Faria et al., 2014, White, Royston & Wood, 2011).

4.6.2.5. MI results

Once appropriate methods for imputing each endpoint were defined, MI with chained equations was conducted.

Imputed and observed data were then compared and it was found that imputed values generally resembled the distribution of observed values. Imputation graphical diagnostics for the first four imputations for all dependent variables (EQ-5D scores and health care costs at 6 and 12 months) are reported in Appendix H. The imputation model appeared to predict the missing values failry.

4.6.3. Economic evaluation after MI

The economic evaluation was conducted again using the multiple imputed dataset.

Table 4.16 reports the results of the SURs adjusted for baseline health and costs. ICERs and INMBs were computed and are reported in Table 4.17. CEACs were re-estimated parametrically and are illustrated in Figure 4.11.

	QALYs			Costs		
	Coeff.	95% CI		Coeff.	95% CI	
Intervention allocation:						
BA	-0.031	-0.061	-0.002	395.538	-137.864	928.940
BLC	-0.025	-0.055	0.005	-258.613	-742.812	225.586
Baseline EQ-5D	0.590	0.524	0.657			
Baseline costs				0.242	0.144	0.340
Constant	0.371	0.310	0.433	720.410	382.823	1057.997

Table 4.16 SUR results after MI for QALYs and costs

Compared to CIL	ICER	INMB (for k=13,000)
BA	Dominated	-803 £
BLC	10315 £/QALY	-67 £

Table 4.17 ICERs and INMBs after MI

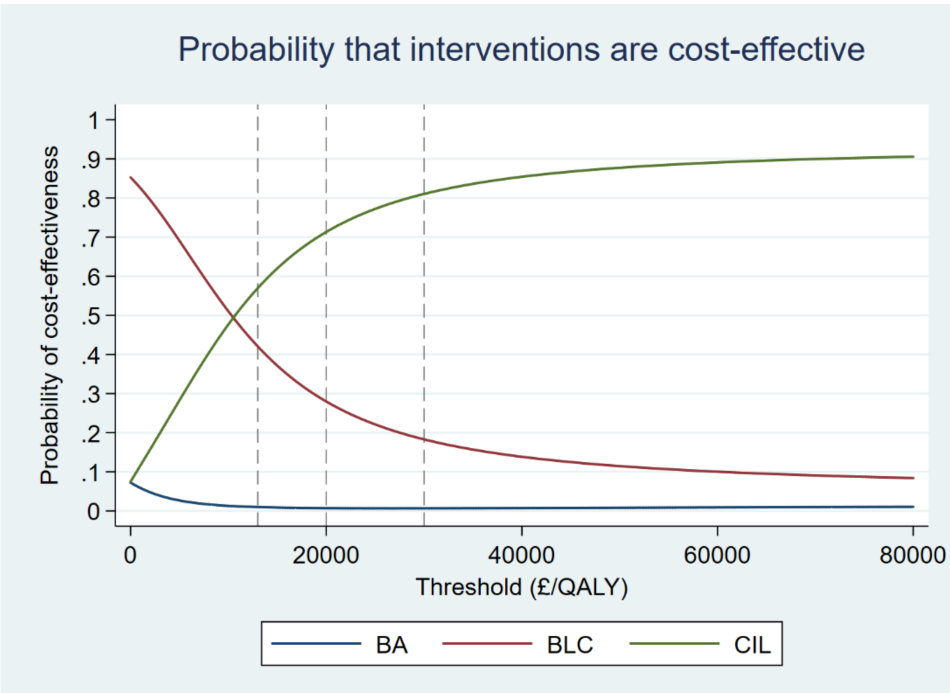


Figure 4.11 CEACs after MI

Results resemble the findings of the CEA under the complete case analysis approach, and confirm again that from the health care prospective it is probably not cost effective to choose BA or BLC, when compared to CIL.

Alternative results based on SURs without baseline adjustment of costs are reported in Appendix H.

4.7. Impact of outliers

Heath care cost categories were summarised in Table 4.6. The main cost components contributing to the amount of total costs were: counselling, emergency ambulance travels, outpatient and GP visits. In this section, the impact of potential outliers affecting the mean values of the four main cost components is investigated.

Figure 4.12 reports the box plots showing the distributions of costs for each component and the presence of outliers. With regard to the costs associated with emergency ambulance travels, there is one outlier that affected costs in the CIL group considerably. Differentials in emergency ambulance travel costs are mostly due to one individual allocated to CIL who used the emergency ambulance 20 times. Similarly, one individual allocated to the BA arm affects considerably the total costs for counselling in their own treatment group. The presence of outliers in GP and outpatients visits is less evident.

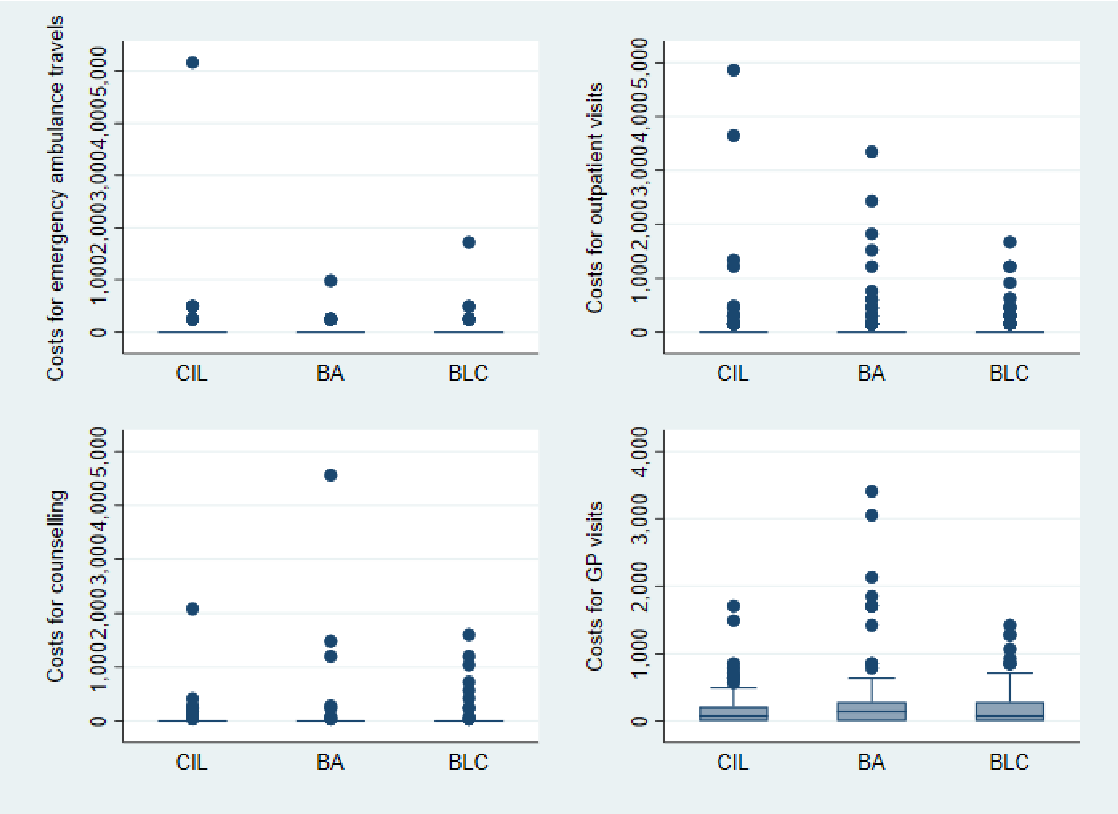


Figure 4.12 Distributions of the main cost components

Adjusting for baseline costs might be a way of smoothing the impacts of outliers if their particularly high consumption of resources is persistent in time. Nevertheless, it was found that baseline costs associated with these individuals were not particularly high.

If there are reasons to believe that such an intense use by only one individual is caused by something not related to the intervention, or if there was an error in the data collection procedure, the outlier might actually bias the analysis. However, it was found that even removing the individual in the BA group did not change the results. Removing outliers in the CIL group would only emphasize its cost-effectiveness by potentially reducing costs even further. Results of the analysis results appear to be robust.

4.8. Conclusions

In this chapter, a 'traditional' CEA from the health (and social) care perspective was conducted. This CEA was based exclusively on costs falling on the health and social care budget, and effects on trial participants measured in QALYs.

Alternative regression models under the complete case approach led to similar results and same judgments on cost-effectiveness. A substantial proportion of missing values was found. Nevertheless, results based on MI data did not affect the recommended adoption decision. In summary, from this first analysis, it could be concluded that (from the health care prospective) BA and BLC are not cost-effective compared to CIL.

Uncertainty appears to have a significant role. Value of Information (VOI) analysis could be performed to investigate the value of reducing the uncertainty in the estimates. Nevertheless, such an analysis would go beyond the purpose of this thesis. Instead of conducting further analyses from the HC perspective, the objective of this work is to investigate the impact of the intervention first from an alternative perspective, and then from a broader and integrated point of view.

The next chapter will investigate the impact of SIPS from the perspective of the CJS decision maker.

Chapter 5: CEA from the criminal justice system perspective

5.1. Introduction

5.1.1. Objective of this chapter

Chapter 4 described the first part of the case study, and assessed the cost-effectiveness of SIPS from the health care system perspective. However, the impact of alcohol misuse may also extend beyond health consequences. In fact, in the UK, alcohol is a major factor in crime and anti-social behaviour. One third of people in police custody have committed an alcohol related offence (Man, 2002) and 40% of prisoners claimed alcohol played a role in the crime they committed (Parkes et al., 2010). Further, 47% of violent offences in England and Wales were alcohol related (Flatley, 2016). Alcohol imposes significant costs on the NHS and the criminal justice system with alcohol related crime and social disorder estimated to cost UK tax payers £11bn per year (Institute of Alcohol Studies, 2016).

The aim of this chapter is to carry out a CEA of SIPS from the CJS perspective. In the next chapter, results of the CEAs from the HC and CJS perspectives presented in Chapters 4 and 5 will be linked and a cross-sectoral analysis will be conducted.

5.1.2. Structure of the chapter

This chapter is divided into two parts.

The first part spans from Section 5.2 to Section 5.4 and describes the methods used to conduct CEA in the criminal justice (CJ) context. Specifically, Section 5.2 provides an overview of the use of economic evaluation in the CJS context and compares the methods to those employed in the HC context. Section 5.3 introduces the proposed methods for the CEA, which reflect the analysis usually conducted from the HC perspective. Section 5.4 explores alternative strategies for the estimation of the opportunity costs.

The second part (Sections 5.5 - 5.9) is devoted to the description of the case study. The analysis plan for the case study is illustrated in Section 5.5. In Sections 5.6 and 5.7, the effectiveness of the intervention is assessed using two alternative data sources. The analysis of costs in response to criminal activity is conducted in Section 5.8. Results of the CEA are described in Section 5.9; alternative scenarios are also presented here. Conclusions of the CEA from the CJS perspective are summarised in Section 5.10.

5.2. Methods for economic evaluation in the CJ context: overview of current practice and comparison to HC context

This section aims to provide an overview of the use of economic evaluation in the CJS context. First, Sub-section 5.2.1 describes the existing methods for CEA in CJS. Then, Sub-section 5.2.2 compares them to the methods typically employed in HC and discusses potential limitations.

5.2.1. An assessment of previous CEAs from the literature

Similarly to the HC sector, the CJS has limited resources. Assessing only the effectiveness of an intervention conducted in the CJ sector would not be sufficient to inform decision making. By contrast, as compared to an analysis solely of the effect of an intervention, an economic evaluation provides additional information to policy makers (Marsh, Chalfin & Roman, 2008).

Looking at the literature of economic evaluation of CJ interventions, various and different analytical techniques have been employed to assess the impacts of interventions in the CJS sector, such as cost analysis, CEA and CBA (French & Drummond, 2005, Swaray, Bowles & Pradiptyo, 2005, McDougall et al., 2003). Because the proposed framework is based on CEAs, the focus of this chapter is on this particular analytical technique. CEA has the objective of informing decisions by assessing and comparing incremental costs and effects of a CJ intervention. CEAs might show that expenditures on treatment programs for offenders might be not only justified in terms of effectiveness, but also produce substantial monetary and nonmonetary benefits (Daley et al., 2004).

This sub-section reviews some examples of CEAs taken from the literature of economic evaluations of CJ interventions. The aim is not to produce an exhaustive review of the economic evaluation methods in CJ context, but to develop an understanding of which methods are employed and how these compare to the healthcare sector. Studies were taken from published reviews of economic analyses of criminal justice interventions (Swaray, Bowles & Pradiptyo, 2005, McDougall et al., 2003).To identify more recent economic evaluations, a rapid review was also conducted. A summary of the main characteristics and potential issues (if any) of the selected CEAs is reported in Appendix I.

In the CEA performed by Daley et al. (2004), results were presented as a comparison between costs and cost savings. Opportunity costs were not taken into consideration. Moreover, cost savings were not included in the total costs of the intervention and were instead used as a measure of benefit (Daley et al., 2004). In many CEAs performed by McCollister et al. (2004, 2003, 2003), re-incarceration was recognised as a useful measuring tool of criminal recidivism, and ICERs resulting from CEAs were interpreted as the marginal cost of achieving one less incarceration day in the treatment study condition relative to control study condition (McCollister, French, Prendergast, et al., 2003). Nevertheless, decision criteria were vague and benefits (i.e. recidivism and cost savings) were not explicitly compared to opportunity costs. Analyses did not take into account the alternative results that could have been achieved with the same resources. Frequently ICERs were compared to the average daily cost of incarceration: if the ICER was higher than the daily cost of incarceration, the intervention was deemed not to be cost-effective (McCollister et al., 2004, McCollister, French, Inciardi, et al., 2003, McCollister, French, Prendergast, et al., 2003).

In the studies reviewed, the need of a cost-effectiveness threshold and its fundamental role in informing investment decisions was frequently not mentioned. One example is the evaluation from Griffith et al. (1999), which is to my knowledge the first CEA conducted from CJS perspective that can be found in the literature. The choice of conducting a CEA was based on the controversy surrounding the methodologies used for translating variables into monetary values. ICERs (representing the amount of money required per person to realise a 1% reduction in reimprisonment) were estimated, but the need of comparing ICERs to a cost-effectiveness threshold was not mentioned. Similarly, Shanahan et al. (2004) undertook an analysis considering both time to the first offense and offending frequency per unit time, and they included all relevant costs from the CJS perspective (i.e. treatment costs, cost of court appearance, costs of probation, and penalty per day). Nevertheless, estimated ICERs were not compared to any benchmark value and it is not clear how conclusions about cost-effectiveness were drawn.

By contrast, Barrett and Byford (2012), after conducting a CEA in order to determine the incremental cost of the programme per serious offence prevented, acknowledged the necessity of comparing results to a benchmark[26]. According to the authors, judgement on the programme could be made depending on a decision maker's willingness to pay for preventing a serious offence, but such an estimate is not available. The role of a WTP to prevent criminal events was also mentioned in the analysis by Muser et al. (2015). The authors estimated several ICERs per event avoided, and compared them to a benchmark value represented by the community WTP to prevent incarceration estimated previously by Cohen et al. (2004).

[26] The authors also compared monetary benefits and costs, but their analysis was explicitly described as a cost-offset analysis.

Focusing more specifically on the alcohol context, in the CEA conducted by Mansdotter et al. (2007) costs and cost savings of the intervention were simply compared. The authors did not consider where the intervention costs fall and aggregated the monetary consequences to the judicial system, production changes, HC and other damage. Moreover, to assess the cost-effectiveness of the intervention from the HC perspective, they compared the ratio of the cost of implementing the intervention and health gains to a cost-effectiveness threshold representing the incremental marginal costs per QALY. By doing this, they compared intervention costs that did not fall entirely on HC sector to an estimate of health care opportunity costs.

5.2.2. A comparison of economic evaluations in HC and CJS contexts

When comparing CEAs conducted in HC and CJ contexts, some inconsistencies in the definition and in the operationalisation of the analytical techniques can be identified. This sub-section explores the similarities and potential links between methods for the economic evaluation in HC and CJ.

First of all, in the CJ context, CEA is frequently viewed as an incomplete CBA exercise, and the main difference between the two is that CBA allows monetising program costs and outcomes (McIntosh & Li, 2012, Swaray, Bowles & Pradiptyo, 2005). By contrast, in the HC context, both CBA and CEA could in principle include a monetary valuation of health outcomes (Sculpher & Claxton, 2012), and the main difference is that CEA recommendations depend crucially on the explicit comparison of health impacts (e.g. expressed in terms of QALYs) with an estimate of the opportunity costs (Culyer & Chalkidou, 2019).

Theoretical and pragmatic inconsistencies in the approach to economic evaluation of interventions conducted in the HC and CJS contexts become apparent when looking at examples of guidance on the methods for the economic analysis in crime prevention. For example, in the Canadian Department of Public Safety guidance (McIntosh & Li, 2012), it is argued that the cost-effectiveness (CE) ratio can be obtained by comparing what the author defined as ‘net effects of programme’ and its ‘total costs’ (Equation 5.1).

$$CE\ ratio = \frac{Net\ Effects\ of\ Program\ (crimes\ averted)}{Total\ Program\ Cost} \qquad (5.1)$$

The CE ratio is claimed to be the result of the CEA, and an intermediate step before computing the benefit-cost ratio (BCR) of the programme. The formula to compute the BCR is reported in Equation 5.2.

$$BCR = Potential\ Societal\ Costs\ (per\ crime) * CE\ ratio \qquad (5.2)$$

As illustrated in the previous sub-section, several examples of CEAs of CJ interventions in the literature follow this approach to the analysis. When such an approach is compared to the techniques typically employed in the HC sector (see Chapter 3, Sub-section 3.2.3.1), inconsistencies in the methods for measuring costs, valuing benefits and opportunity cost become evident. Details of each are listed below.

5.2.2.1. Methods for measuring costs

In CJS, the total cost of a programme is defined as the direct cost of the intervention, i.e. the sum of administrative, capital and indirect costs. Other CJS cost components such as police and investigative costs, prosecution, courts, corrections, are not included (McIntosh & Li, 2012). Thus, if an intervention is associated with a reduction in crime, cost savings from the CJS perspective (e.g. reduced correction costs) are not included in the total cost of the programme; instead they are treated as societal benefits.

For the CE ratio to be consistent with the ICER typically employed in health care context, it would be necessary to estimate the full costs falling on the CJS budget and incorporate them in the total cost of the programme. These costs might be negative due to reduction in crime.

Moreover, the victims' medical costs, both tangible (i.e. medical and mental health care) and intangible (e.g. pain and suffering, decreased quality of life, fear of crime, emotional/psychological distress), should be included in the analysis as well. However, it should be remembered that these costs do not fall on CJS budget. The same applies to the direct property losses and productivity losses (i.e. lost workday and schooldays for victims and incarcerated offenders), which fall on private consumption.

5.2.2.2. Methods for valuing effects

The analysis conducted using Equation 5.2 represents a comparison between costs and cost savings; benefit is assumed to be captured in cost savings related to crime, and no value is placed on the reduction in crime itself. By contrast, in CEA in the HC context, the health itself is included in the outcomes. This is because the HC decision maker has the objective of maximising health, which has a value *per se* (Drummond, M. et al., 2015).

Similarly, it can be argued that the aim of the CJS decision maker is to reduce crime because a negative value is associated with crime *per se*[27] and that the socio-democratic process confers the authority to the decision maker to strive for a reduction in crime. If the assumption about the intrinsic value of crime is accepted, measuring the effectiveness of an intervention in terms of cost savings alone can underestimate the benefit of the intervention. Cost savings of course must be taken into account, but they should be included in the total costs (as highlighted in section 5.2.2.1). Afterwards, total incremental costs (obtained as the sum of costs and cost savings) must be compared to the variation in crime, which has a value itself.

5.2.2.3. Opportunity costs

As it is formulated in Equations 5.1 and 5.2, the analysis does not take into account the opportunity costs of the intervention and the existence of a CJS budget. To be consistent with the CEA approach in the HC context, total costs falling on CJS budget (including costs savings) should be then compared to CJS opportunity costs.

As explained in Chapter 3, the objective of the proposed framework is to conduct similar and compatible analyses from different perspectives, and potentially combine their results. Therefore, Section 5.3 proposes an approach to conduct CEA from the CJS perspective which is compatible with the methods used in the HC context.

5.3. Proposed approach for the CEA from the CJS perspective

The importance of developing a standardised methodology for calculating the relative costs and benefits of CJ programs has been pointed out in the literature (Swaray, Bowles & Pradiptyo, 2005, McDougall et al., 2003). This is also because the lack of a standardised outcome measures prevents the comparison of costs and benefits of different sentencing options (McDougall et al., 2003). Moreover, it has been argued that policy-relevant findings appear to be generally based on less rigorous[28] methods of economic analysis when compared to HC interventions (Swaray, Bowles & Pradiptyo, 2005).

The aim of this section is to borrow the CEA methods usually employed in the HC context, and adapt them to the CJ context. As explained in Chapter 3, Section 3.2.3.3, from the perspective of a 'non-

---

[27] As explained later in Section 5.4, individuals have in fact a positive WTP to reduce crime (Cohen et al., 2004).

[28] On the basis of an economic analysis rating scale that helps researchers to judge the quality of costs and benefits information in the literature developed in McDougall et al. (2003).

health care sector', direct non-health effects (ΔN) and direct non-health costs (ΔC_N) can be assessed. From the narrow perspective of the non-health sector decision maker, if the Incremental Net Non-health Benefit (INNB) – given by the difference of DEs and OCs – is positive, the recommended decision will be to implement the intervention.

The formula to compute a generic INNB was illustrated in Equation 3.3, and can be adapted for the specific case of the CJS. In the CJ context, the INNB depends on the two usual components: DEs (falling on offenders treated with the intervention), and OCs. Sections 5.3.1 and 5.3.2 describe each component in further detail, while section 5.3.3 illustrates how to assess the incremental net benefits.

5.3.1. Direct effects

To measure the DEs of an intervention, it is necessary to select an outcome measure that is deemed appropriate by the decision makers involved in the decision making process. For example, a CEA from the health care perspective aims to assess incremental cost per QALY, while criminal justice providers base their funding decisions on re-offending (Fox & Albertson, 2011). Recidivism (often measured using reconvictions as a proxy) has been deemed an appropriate outcome measure to assess the effectiveness of interventions from the CJS perspective (Pearson et al., 2016, Muser et al., 2015, McCollister, French, Inciardi, et al., 2003, Griffith et al., 1999).

In the approach proposed here, avoided reconvictions are employed as a proxy to measure criminal activity, and DEs are measured in terms of reconvictions avoided (ΔR) by the offenders treated with the programme.

5.3.2. Opportunity costs

Similarly to the HC context, OCs in the CJS context are generated by the displacement of CJS resources to accommodate the costs of the new intervention. Displacement of resources implies that other (unknown) offenders will commit crimes as the resources for their programmes are displaced. The magnitude of the OCs depends on the marginal productivity of CJS sector in avoiding reconvictions (k_R) and on the total costs (ΔC_{CJS}) falling on the CJS budget (i.e. the amount of resources displaced).

Total costs falling on the CJS budget (ΔC_{CJS}) are composed by the incremental costs of the intervention and the costs related to re-offending (e.g. prosecution, courts, custodial sentencing costs, prison costs). Costs associated with re-offending are those costs that occur in response to crime, and therefore become cost-savings if the intervention reduces crime. Total costs falling on CJS budget include thus also a cost saving component if crime is reduced. In other words, the novelty of the proposed approach is that a component of the total costs is a function of the direct effects of the intervention. The implications of such relationships are explored in the case study, in Sub-section 5.9.1.

5.3.3. Incremental Net Benefits

INBs are obtained by the DEs net of the OCs. If expressed in natural units (i.e. reconvictions), INBs can be defined as Incremental Net Reconviction Benefits (INRBs) and can be computed as illustrated in Equation 5.3. The same benefits can be expressed in monetary terms (INMBs) as well.

$$INRB = \Delta R - \frac{\Delta C_{CJS}}{k_R} \quad (5.3)$$

As opposed to the methods discussed in the previous section, such a formulation considers both the DEs and the potential cost savings associated with potential crime reduction. These costs components will be illustrated in detail using the SIPS case study and are examined again in Sub-section 5.9.1.

To operationalise the formula illustrated in Equation 5.3, a cost-effectiveness threshold representing the marginal productivity (*k)* of the CJS would be required. Once an estimate of *k* is available, INBs of the intervention from the CJS perspective can be calculated. Nevertheless, as explained in the next section, such an estimate is not available in the literature, and alternative proxies must be used.

5.4. Strategies for the estimation of the opportunity costs in CJS

As explained in Chapter 3, Section 3.2.3.1, a pragmatic approach typically followed in the HC context is to measure health in QALYs and use an estimate of the marginal productivity (*k)* of the HC sector to estimate the opportunity costs of the investment (i.e. what could be alternatively done with the same resources) (Claxton, Martin, et al., 2015). Similarly, to perform a CEA from the CJS perspective, and to operationalise Equation 5.3, it would be necessary to use an estimate of the marginal productivity of the CJS sector in avoiding re-offending.

Nonetheless, a well-recognised limitation in conducting CEAs outside the HC sector is the lack of evidence on the benefits displaced elsewhere by diverting resources to the intervention being evaluated (i.e. what benefits could be generated if the resources were used for another purpose)

(Hill et al., 2017). In fact, no estimates of the marginal productivity of the CJS in reducing re-offending currently exist.

In light of this, an alternative pragmatic approach might be to use an estimate of the social decision maker's WTP for a gain in a unit of outcome from the production of a service (i.e. for reducing re-offending) as a proxy[29]. Under the assumption that social decision maker's WTP is the best estimate of the OCs, it can be considered a proxy of *k* for that sector (Remme, Martinez-Alvarez & Vassall, 2017).

Two approaches to estimating the social decision maker's WTP are proposed in this section. The first approach uses examples of Payments by Result (PbR) agreements and Social Impact Bonds (SIB) already in place in the CJS as proxies for decision makers' WTP (Palumbo & Learmonth, 2014, Fox & Albertson, 2011). The second approach uses individuals' WTP (*v*) for reducing crime to estimate *k*. Such an approach assumes that budget allocation across sectors is consistent with the individual WTP for the selected outcome. Sections 5.4.1 and 5.4.2 describe and discuss each approach in detail.

5.4.1. Payment by result (PbR) and social impact bond (SIB)

PbRs are schemes that allow the government to pay a provider of services on the basis of the outcomes their service achieves rather than the inputs or outputs the provider delivers. More precisely, social impact bonds (SIBs) are one specific form of PbR which allow the financing of social outcomes via private investment (Fox & Albertson, 2011). Social decision maker's WTP estimates could be therefore estimated from examples of PbR agreements already in place in the CJS.

Examples of SIBs are already in place in the CJS (Palumbo & Learmonth, 2014, Fox & Albertson, 2011). For example, details of the Peterborough SIB – the first pilot model based on SIB in the English and Welsh CJS – can be found in Appendix J (Anders & Dorsett, 2017, Social Finance, 2017, Disley et al., 2015, Jolliffe & Hedderman, 2014, Cave et al., 2012).

These SIBs provide evidence that decision maker in the CJS does actually base investing decisions on reductions in reconviction frequency. However, available information was not sufficient to infer an estimate of the decision maker's WTP for reducing re-offending. Publicly available information

[29] Such an estimate should not be confused with the individuals' WTP for a unit of outcome, generally referred to as *v* (Woods et al., 2016).

about the actual payment and about the details on the agreement was vague. No clear decision criteria could be deduced from examples of actual SIBs.

5.4.2. Alternative pragmatic solution

The initial part of Section 5.4 highlighted the lack of an estimate of CJS's marginal productivity for a unit of outcome and the difficulties in determining a proxy of decision maker's WTP from SIBs. An alternative pragmatic solution is to estimate a proxy of CJS's marginal productivity starting from the individuals' WTP for reducing crime.

This alternative approach was proposed by Claxton et al. (2019) and is based on the assumption that the socio-democratic process allocates resources to each sector so that a pound spent in any sector at the margin generates the same benefit as a pound spent in another based on individual WTP. In other words, it is assumed that the ratio of societal willingness to pay (v) over marginal productivity (k) is constant across sectors, so that a pound spent in HC generates the same benefit as a pound spent in CJ, based on individual WTP. A similar explorative assumption was made by Woods et al. (2016) in order to obtain initial estimates of country-level cost-effectiveness thresholds. In their analysis, the authors assumed that the discrepancy between *v* and *k* for HC is constant in relative terms across countries. Here the assumption is made within the country, but across sectors. Limitations and assumptions of this approach are further discussed at the end of this section and in the discussion chapter, Sections 8.4 and 8.5.

As explained in the next sub-section, *k* and *v* in the health sector can be used to shadow price other forms of public expenditure where the equivalent estimates for that sector are absent (Claxton et al., 2019). Estimates of *k* and *v* are available for the HC sector. Consequently, *k* for CJS can be computed starting from societal WTP for reducing crime.

5.4.2.1. Individual WTP for improvements in health

With regard to the HC context, in 2009 the Department of Health and Social Care estimated that a QALY had a monetised value of £ 60,000 (Glover & Henderson, 2010). A more recent review of the evidence in the literature confirmed a mean WTP of around € 75,000 (in 2010 Euros), which means approximately £ 63,500 (in 2010 GBP) (Ryen & Svensson, 2014). It is thus assumed that the societal WTP for a QALY is £ 60,000 (in 2010 GBP)[30].

As illustrated in Chapter 3, Section 3.2.3.1, *k* for HC sector was found to be approximately £ 13,000 per QALY (Claxton, Martin, et al., 2015). The WTP for a QALY (v) is therefore about 4.6 times bigger than *k*.

5.4.2.2. Individual WTP for reducing crime

Moving to the CJ context, there are different WTPs that can be estimated. In the literature, several studies that investigate individuals' WTP to reduce their likelihood of being a victim of a violent crime can be found (Manning, Fleming & Ambrey, 2016, Ambrey, Fleming & Manning, 2013, Bishop & Murphy, 2011, Atkinson, G., Healey & Mourato, 2005). These analyses use contingent valuation (CV) methods, the life satisfaction approach or hedonic models, and their results can be interpreted as estimates of the intangible costs of crime.

However, the WTP for reducing the risk of being a victim of a violent crime is conceptually different from the WTP to reduce crime *per se*. In fact, the latter WTP might be driven by different reasons, not only the risk of being a victim (e.g. a natural aversion to criminality), and the inclusion of impacts on victims' health in the WTP estimate would lead to double counting of the effects of the intervention on health. The individual WTP for reducing crime *per se* is therefore needed for the calculation of the threshold representing the OCs of the CJS. Having this in mind, the study by Cohen et al. (2004) appears to be the best proxy available in the literature. Their findings have been already employed in another recent CEA (Muser et al., 2015).

Using CV methods, Cohen et al. (2004) estimated individuals' WTP for crime prevention programs, and found that the WTP amounts to approximately \$25,000 (in year 2000 \$) per burglary and \$70,000 per serious assault, with amounts increasing as the seriousness of the offense increased. With regard to White-Collar and Corporate Crime, Cohen (2015) estimated a WTP of \$1,200 per consumer fraud and \$12,000 for financial fraud.

Assuming the transferability of these specific WTPs to the UK context (limitations are discussed at the end of this section), a generic WTP for a generic crime averted can be calculated. A weighted average of WTPs per crime was computed using the numbers of criminal events in the UK as

---

[30] For the sake of simplicity, it could be assumed that this WTP is constant in time. However, the reference year of the SIPS case study is 2010. Therefore, the WTP value does not need to be discounted.

weights. The numbers of criminal events were taken from the Home Office Statistical Bulletin on crime in England and Wales (Chaplin, Flately & Smith, 2011).

WTPs and number of events for each crime category are reported in Table 5.1. Original WTP values for burglary, armed robbery, serious assaults, rape and sexual assaults, and murder were updated to the year 2010 (the reference year of the case study) and converted into GBP. WTP associated with fraud was obtained as an average of the two WTP values for consumer and financial fraud, and converted from 2015 USD in 2010 GBP[31].

Crime	Implied WTP per crime (2010 $)	Category of crime from Home Office Statistical Bulletin	No. of events in 2009/2010 in UK
Burglary	31,650	Total burglary offences	540,660
Armed Robbery	293,712	Robbery of personal property	66,923
Serious Assaults	88,620	Violence against the person with injury	401,629
Rape and sexual assaults	300,042	Total sexual offences	54,355
Murder	12,280,200	Homicide	618
Fraud	6,072	Fraud and forgery offences	152,241

Table 5.1 WTPs and number of events for each crime category

Resulting weighted average WTP (*v*) was approximately $ 80,000, or equivalently £ 51,000 when converted into 2010 GBP.

5.4.2.3. Resulting proxy for *k* in the CJ sector and its limitations

If the ratio of *v* and *k* is assumed to be constant across sectors, the proxy of *k* for the CJS would be approximately £ 11,000 per criminal event averted. In other words, with £ 11,000 CJS can prevent one episode of recidivism at the margin. As stressed previously, there are a number of limitations in the estimation of such a value.

First of all, the estimate is based on the assumption that resource allocation reflects the democratic process and is consistent across sectors when compared to the individual WTP values. This does not necessarily correspond to reality. Nevertheless, the pragmatic assumption that the ratio between *v* and *k* is constant can be modified if alternative evidence about the allocation across sectors is provided.

---

[31] Based on the 2010 spot exchange rates provided by the Bank of England.
Source: https://www.bankofengland.co.uk/boeapps/database/Rates.asp?

Secondly, it must be stressed that the study by Cohen et al. (2004) is quite old and it was conducted in the US. Individuals' preferences might change according to the year and the specific national context. WTP values reported might therefore not reflect the preferences of the UK population in 2010. Nevertheless, to my knowledge, no other more recent attempts of updating these estimates have been conducted. Moreover, the focus of this research is not to investigate societal WTP for reducing crime, but to illustrate how these estimates could be used in order to estimate the OCs of the CJS. If better evidence is provided, the proxy for *k* can be adjusted accordingly.

Thirdly, two issues regarding the value of the estimated WTP must be highlighted. The generic WTP per criminal event averted does not take into account other theft offences, criminal damage offences and drug offences because no individual WTP was found for these crime categories. Moreover, it is acknowledged that the estimated WTP might actually be an upper bound of the true intrinsic value of crime. In fact, it is probably unavoidable to somehow include also consequences on victims' health in the WTP to reduce crime *per se*, unless such a value is specifically investigated.

Lastly, it was assumed that the estimated WTP could be used to find a proxy of the marginal productivity of CJS in preventing crime. However, estimated WTP was referred to a reduction in criminal events, while marginal productivity of the CJS is measured in terms of reconvictions. It must be therefore assumed that the reduction in criminal events corresponds to a reduction in reconvictions.

That being said, it must be highlighted that this has the aim of being an exploratory work, with the objective of showing how to operationalise a theoretical framework. The adoption of some pragmatic assumptions is necessary to address the lack of data. Gaps in the literature and the need for cross-sectoral data pointed out in this analysis are further discussed in Chapter 8, Section 8.5.

5.5. Analysis plan for the case study

The Screening and Intervention Programme for Sensible drinking (SIPS) trial was selected as case study. Introduced extensively in Chapter 3, SIPS aims to tackle alcohol consumption in offenders and comprises three intervention arms: Client Information Leaflet, Brief advice, and Brief Lifestyle Counselling. Each intervention arm is an intensified version of the previous.

In the previous chapter, a CEA of the SIPS trial from the HC perspective was conducted. The aim of this second part of Chapter 5 is to operationalise the methods just illustrated and conduct a CEA of SIPS from the CJS perspective. This section summarises briefly the analysis plan.

The analysis of criminal outcomes (i.e. the effectiveness of the programme from the CJS perspective) is reported in Sections 5.6 and 5.7 using Police National Computer (PNC) data and self-reported data, respectively. In Section 5.7, the two measures of effectiveness are compared and a summary of the findings is presented. Section 5.8 covers the analysis of the budgetary impact of the intervention. Results of the CEA from the CJS perspective are reported in Section 5.9. CEA is based on the methods proposed in Section 5.3. INBs were computed using social decision maker's WTP to avoid one additional reconviction (based on individuals' WTP to reduce crime) as a proxy for *k* in the CJS context. To assess the robustness of results, alternative scenarios are also discussed. In Section 5.10, conclusions are presented. All analyses were performed using STATA 15 (StataCorp, 2015) and Microsoft Excel 2010 (Microsoft, 2010)

5.6. Assessment of recidivism using PNC data

For the analysis of the SIPS trial, two sources of information for criminal events were available: an extract from the PNC database and self-reported events from questionnaires. This section is devoted to the analysis of PNC data. Section 5.7 illustrates the analysis of self-reported outcomes.

5.6.1. Assumptions about available data

Due to the scarcity of information about the dataset extracted from the PNC database, crucial details about the collected variables (such as details about crime committed and dates of the offences) were unclear or unavailable. As a consequence, a few assumptions were made before starting the analysis process.

Firstly, in the available PNC dataset, names of the variables and labels were sometimes incoherent. For example, it was unclear whether PNC records measured conviction or caution episodes.

Nevertheless, in the analysis of SIPS previously published (Newbury-Birch et al., 2014), results were interpreted in terms of (re)-convictions[32]. The following analysis relied on this interpretation of the data.

Secondly, it was unclear whether conviction data were collected over the year following the screening, or the year after the intervention; previously published analyses provide contrasting descriptions (Newbury-Birch et al., 2014). Keeping in mind that this case study has mainly an illustrative purpose, in the following analysis it was assumed that conviction rates were collected in the year following the intervention.

Thirdly, convictions before the intervention were also reported. This variable was used in the analysis to adjust the number of convictions after the intervention for potential baseline differences in previous conviction events. However, it is important to note that the timing of the reported convictions was not stated, and it was unlikely to reflect the previous year as average values were higher than national annual frequencies.

5.6.2. Descriptive analysis of reconviction rates

The criminal behaviour in the year following the intervention was assessed using the PNC conviction data. Table 5.2 summarises the number of individuals for each intervention group who were convicted at least once in the year following the intervention.

Convicted	CIL	BA	BLC	Total
No	90 (50.28 %)	110 (63.95 %)	99 (61.88 %)	299 (58.51 %)
Yes	89 (49.72 %)	62 (36.05 %)	61 (38.13 %)	212 (41.49 %)
Total	179 (100 %)	172 (100 %)	160 (100 %)	511 (100 %)

Table 5.2 Convictions in the year after the intervention

As reported in the original analysis of SIPS (Newbury-Birch et al., 2014), over the next 12 months, offender treated with BA and BLC were less likely to be reconvicted. Half of the offenders treated with CIL were reconvicted at least once during the year after the interventions, whereas only 36% and 38% of the individuals in the BA and BLC groups were reconvicted.

[32] Only probation workers participated in the trial. Consequently, any conviction after the intervention was actually a re-conviction episode.

In Table 5.2, reconvictions were measured using a binary variable assessing the number of individuals reconvicted at least once. Although this can provide a good overview of the effectiveness of the intervention, as stressed in the introduction, CJ decision makers usually base their decisions on the frequency of reconviction events (Disley et al., 2015). Boxplots in Figure 5.1 show the effectiveness of the programme if frequency of reconviction events is selected as the outcome metric.

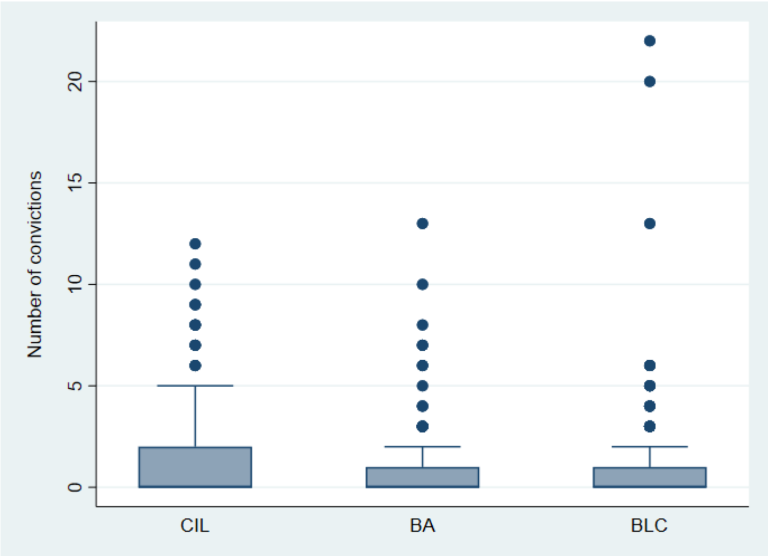


Figure 5.1 Reconviction events

Data appear to be right skewed. Median values for all groups are zero, and BLC has the largest interquartile range. The histogram in Figure 5.2 emphasises the presence of a high number of individuals who were not re-convicted after the intervention.

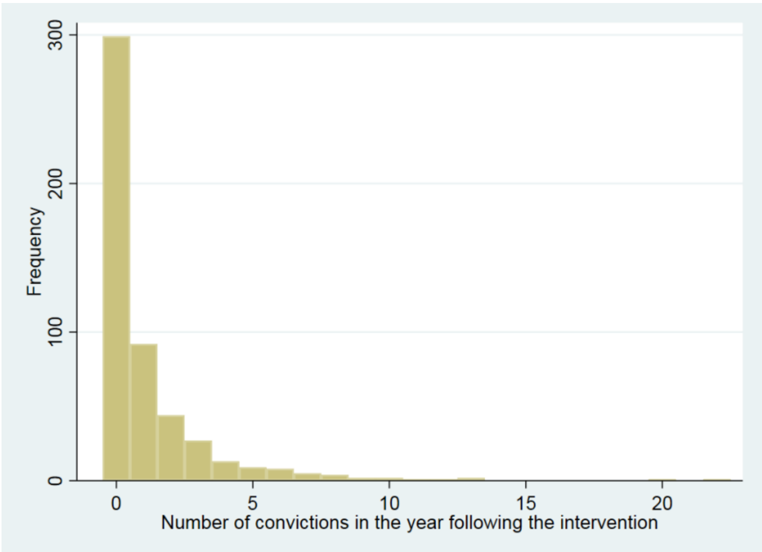


Figure 5.2 Distribution of convictions after the intervention

An excessive number of zeros could potentially introduce biases in the analysis. For this reason, appropriate regression models are investigated in the next sub-section.

5.6.3. Analysis of reconviction frequency

To assess differences in reconvictions, it might be desirable to adjust the average number of reconviction events after the intervention for differences at baseline. In fact, randomisation itself is not a perfect process, and CJ outcomes were not the main focus of the SIPS trial. Therefore, unadjusted differences in reconvictions might be biased also by differences in baseline characteristics between groups.

To adjust the differences in CJ outcomes it would be necessary to take into account not only the available baseline characteristics (e.g. age and gender), but also other factors such as: social economic status, violent behaviour, substance misuse. However, these factors are not available in the dataset. Moreover, it would not be feasible to build a complex regression model and adjust for all these covariates due to the limited sample size. For these reasons, baseline conviction events were used as a proxy to correct for potential imbalances in personal characteristics, criminal history, and prior sentences between groups at baseline. It was therefore assumed that convictions at baseline were the best proxy to summarise potential baseline differences in criminal behaviour. None of the other baseline variables were included to avoid potential collinearity issues, and two alternative baseline corrections were investigated: the before-intervention conviction data from PNC, and a self-reported reconviction count before the intervention.

The number of reconvictions events in each group was investigated. Table 5.3 summarises the two potential alternatives for baseline correction (i.e. self-reported arrests or cautions in the previous 6 months and reconvictions pre-intervention) and the variable of interest (i.e. reconvictions one year after intervention).

	Self-reported data		**PNC database**			
	N	**Number of arrests or cautions in the previous 6 months, at baseline (SD)**	**N**	**Number of reconvictions pre-intervention (SD)**	**N**	**Number of reconvictions one year after intervention**
CIL	180	1.116 (1.499)	179	2.508 (2.768)	179	1.458 (2.358)
BA	175	0.857 (1.211)	172	2.616 (2.064)	172	0.912 (1.853)
BLC	161	1.18 (2.402)	160	2.893 (3.01)	160	1.15 (2.824)

Table 5.3 Reconvictions before and after the intervention

Arrests and reconvictions at baseline appear to be similar across groups, and both BA and BLC appear to be associated with greater reductions in reconviction episodes. However, alternative regression models were employed in order to adjust group differences in reconvictions for potential baseline differences, and obtain unbiased estimates. Three regression models were explored: PNC data and self-reported data were used in Model 1 and Model 2, respectively; Model 3 includes both sources of information and adjusted for the presence of clusters, due to the randomization at the level of offender manager (OM). With regard to the functional form of the regression models, three alternatives were investigated: Poisson, negative binomial (NB) and zero inflated regression models. Appropriateness, strengths and limitations of each alternative are discussed in the next sub-sections.

5.6.3.1. Poisson regression model

Since the dependent variable was a count variable, a Poisson regression model was judged the most appropriate. This consideration was confirmed by a Park test; compared to a Gaussian, Gamma and Inverse Gaussian, results of the test indicated that a Poisson distribution provided the most appropriate fit to the data. The model passed also two tests (Pregibon's link test and Hosmer-Lemeshow Test) for the appropriate link function (Jones et al., 2013).

The three regression models using PNC and self-reported data to adjust for baseline differences provided similar results. The number of cautions before the screening were found to be positively and statistically significantly associated with the number of convictions after the intervention. Outputs of the Poisson regression models are reported in Appendix K.

Poisson regression models are based on the assumption that the dependent variable is not over-dispersed (i.e. mean and variance are the same) and does not have an excessive number of zeros (Jones et al., 2013). As reported in the two following sub-sections, both assumptions were tested, and regression models based on alternative functional forms were investigated.

5.6.3.2. Negative binomial regression models

Negative binomial regressions can take into account extra variation (i.e. greater than the mean) and do not assume the absence of overdispersion. Outputs of the three negative binomial regression models are reported in Table 5.4, to compare alternative corrections for baseline differences.

	Model 1		Model 2		Model 3	
	Coeff.	95% CI	Coeff.	95% CI	Coeff.	95% CI
Intervention allocation:						
Brief Advice	-0.388	[-0.755, -0.020]	-0.362	[-0.731, 0.006]	-0.356	[-0.727, 0.015]
Brief Lifestyle Counselling	-0.542	[-0.936, -0.147]	-0.209	[-0.689, 0.270]	-0.537	[-0.951, -0.123]
Number of cautions prior to screening	0.195	[0.148, 0.242]			0.177	[0.127, 0.227]
Number of times in last 6 months been arrested or cautioned (at baseline)			0.225	[0.100, 0.349]	0.108	[0.002, 0.214]
Constant	-0.263	[-0.545, 0.018]	0.032	[-0.265, 0.329]	-0.349	[-0.649, -0.049]
Alpha	1.861	[1.457, 2.378]	2.437	[1.918, 3.096]	1.850	[1.448, 2.364]
AIC	1401		1420		1374	
BIC	1422		1441		1400	

Table 5.4 Negative binomial regression models

Overdispersion (i.e. the *Alpha* parameter in Table 5.4) is estimated to be 1.86 and the null hypothesis that it equals zero is rejected. This means that the equidispersion property imposed by the Poisson model is rejected, and a negative binomial regression model would provide a better fit for the data.

Comparing the three models, Akaike information criterion (AIC) and Bayesian information criterion (BIC) values appear to be slightly in favour of Model 3. The second best option would be Model 1. Results across the three models are however quite similar. When compared to Model 3, Model 1 would be more consistent in terms of source of data, because both the explanatory and dependent variable are taken from the PNC database. Moreover, including in the regression two variables which should theoretically measure the same outcome (i.e. criminal activity before the intervention) might generate multicollinearity issues. For these reasons, Model 1 was chosen as the final model for the analysis.

By exponentiating the coefficients of a negative binomial regression model, incidence rate ratios (IRRs) can be obtained. According to Model 1, the IRR associated with BA is 0.678 (95% CI: 0.469, 0.980), and the IRR associated with BLC is 0.582 (95% CI: 0.392, 0.863). In other words, when treated with BA (BLC) the mean rate of conviction is expected to decrease by a factor 0.678 (0.582) compared to CIL. The corresponding adjusted numbers of reconvictions associated with each intervention allocation were also computed, and results are summarised in Table 5.5.

Intervention allocation	Average reconvictions	Compared to CIL
CIL	1.57	-
BA	1.06	-0.51
BLC	0.91	-0.66

Table 5.5 Average reconvictions

5.6.3.3. Zero inflated regression models

As highlighted in Figure 5.2, the distribution of reconviction events after the intervention was characterised by a high proportion of zeros. For this reason, zero inflated regression models were employed to investigate whether positive values and zeros might be generated by different mechanisms (Cameron & Trivedi, 2010).

Full regression results of the zero inflated models are reported in Appendix K. In summary, comparing the zero inflated negative binomial (ZINB) regression model to the zero-inflated Poisson (ZIP) regression model, the test on overdispersion confirmed the relevant role of the alpha parameter. Due to the presence of overdispesion, a ZINB regression was more appropriate than a ZIP regression. Nevertheless, AIC and BIC were almost identical in ZINB and NB models. More specifically, AIC and BIC for non-inflated NB Model 1 regression were slightly lower than ZINB, and this suggested using the non-inflated model. Tests comparing zero-inflated models to standard negative binomial models did not provide evidence for adopting two part models, especially for Model 1.

5.6.3.4. Summary

Alternative regression models were explored, and the negative binomial model ‘Model 1’ (described in Sub-section 5.6.3.2) appeared to be the most appropriate.

5.7. Self-reported crime outcomes

Results presented in Section 5.6 were based on PNC data on reconvictions. This section illustrates the analysis of self-reported data on criminal episodes committed in the last six months, at six and twelve months of follow-up.

SIPS questionnaires collected information about a list of the most common violent and criminal events. Individuals reported on these events either as a victim or a perpetrator. In this analysis only events committed as the perpetrator were considered. Episodes of violence included: generic violence, assault, and wounding. With regard to criminal offences, the list of episodes included: vehicle and non-vehicle theft, burglary, criminal damage, robbery, and 'other'. When choosing the 'other' option, respondents could add details about the criminal or violent event committed. Other events were therefore re-classified in the following additional categories: shop lifting, sexual offences, drug offences, driving offences and other fines.

The overall proportions of self-reported crimes committed for each category across all three intervention arms are reported in Table 5.6.

Violent offences		**Criminal offences**	
Violence	21.7%	Theft (non-vehicles)	12.7%
Assault	27.6%	Vehicles theft	1.3%
Wounding	25.4%	Burglary	1.8%
Sexual offences	0.3%	Criminal damage	3.8%
All violent offences	**75.0%**	Robbery	1.6%
		Shop lifting	0.3%
		Drug offences	0.8%
		Driving offences	2.4%
		Other fines	0.5%
		All criminal offences	**25.0%**

Table 5.6 Distribution of self-reported criminal events

Of all self-reported crimes, 75% were violent offences, while the rate of generic violence, assault and wounding were comparable. Most of the remaining criminal offences were thefts.

Before analysing the self-reported data, the correspondence between self-reported and PNC data was explored. Results are reported in Sub-section 5.7.1. It must be however remembered that the potential correspondence of self-reported and PNC data is based on the crucial assumptions enunciated in Section 5.6.1. It is possible that in reality the two variables simply did not measure the same outcome. In fact, from the available data, it was not clear whether convictions were measured one year after the screening or the intervention, and it was not possible to check whether PNC and self-reported data had different time horizons.

5.7.1. Exploring the correspondence between self-reported and PNC data

The aim of this sub-section is to check whether individual self-reported data corresponds to PNC records. A two-way table was created to assess the proportions of convicted individuals who did or did not self-report at least one criminal event. Results are reported in Table 5.7.

	PNC database		
	Not convicted	**Convicted**	**Total**
No self-reported criminal events	153 82%	47 55%	200 74%
At least one self-reported criminal event	33 18%	39 45%	72 26%
Total	186 100%	86 100%	272 100%

Table 5.7 Convictions and self-reported crimes

Out of 186 offenders who were not reconvicted after the intervention, 153 (82%) did not self-report any criminal offence; the remaining 33 (18%) declared that they committed at least one criminal event, but this did not result in a conviction. Out of 86 offenders who were reconvicted at least once after the intervention, only 39 (45%) self-reported at least one criminal event; the remaining 47 (55%) did not report the criminal event that they did actually commit.

Even though assumptions stated in Section 5.6.1 would theoretically imply the consistency of the information collected from the two sources of data, it appears from Table 5.7 that PNC and self-reported variables do not actually match. Moreover, looking at the frequencies of convictions and self-reported criminal offences, variables show even greater discrepancies. For this reason, no attempts of linking each reconviction episode to a specific criminal offence was made.

The two variables are therefore investigated as if they were two independent CJ outcomes. Similarly to what has been done with PNC data, in Sub-section 5.7.2 the analysis of self-reported criminal events is conducted to see if there are differences in terms of effectiveness. Afterwards, in Sub-section 5.7.3, the two measures of effectiveness are compared and discussed.

5.7.2. Analysis of self-reported criminal events

The purpose of this analysis is to assess the effectiveness of the intervention from the CJS perspective, using self-reported outcomes. In particular, the aim is to investigate the frequencies of violent and criminal offences.

5.7.2.1. Data cleaning and assumptions about missing criminal events

The structure of the CJ section of the SIPS questionnaire was similar to the HC resource consumption questionnaire illustrated in Appendix D. A generic 'yes or no' question was followed by more specific questions about the frequencies of criminal events[33].

As summarised in Table 5.8, the proportion of incomplete records (i.e. missing frequencies of criminal events corresponding to a 'yes' in the generic question) is really small when compared to the totality of correctly reported values (i.e. answer to the generic question is coherent with the stated frequency of a criminal event).

	Complete records	**Incomplete records**
6 months	4573	25
12 months	4006	22
Complete cases	**3531**	**36**

Table 5.8 Criminal events correctly reported and imputed

Only 36 out of 3567 (i.e. approximately 1%) records have missing frequencies of criminal events which corresponded to a 'yes' in the generic question. Similarly to the analysis of HC resource use (see Appendix D), keeping in mind the almost negligible proportion of incomplete records, two alternative scenarios were constructed to enlarge the complete cases data set on the basis of alternative assumptions on the missing information about crime frequencies when the generic question was instead completed: the 'mean imputation scenario' and the 'minimum imputation scenario'.

Tables summarising criminal events at six and twelve months in the original, mean and minimum imputation scenarios are presented in Appendix L. As expected, given the small proportion of missing values, the numbers of average criminal events are almost identical in minimum and mean imputation scenarios. Only the average frequency of assaults slightly differs depending on the

[33] E.g. have you committed this crime? If yes, how many times?

different assumptions regarding the missing values. Nevertheless, (as it will be illustrated later in the analysis of costs) the unit cost of assaults in one of the lowest, therefore the impact of their variation on total costs is still minimal. To sum up, the total number of imputed events is low and overall differentials between alternative scenarios are negligible. As previously done in the analysis of HC resources consumption, the mean imputation scenario was chosen as the reference case.

It must be stressed that this analysis is still conducted under the complete case criterion, because only individuals who answered to the generic questions about criminal activity at both six and twelve months were considered. The mean imputation only refers to a small proportion of missing frequencies that were not correctly reported in the questionnaires. However, these results are based on only 269 out of 525 offenders who self-reported the crimes they committed. Such a considerable presence of missing observations could be investigated via multiple imputation procedure (see Chapter 4, Section 4.6). However, such an analysis would go beyond the scope of this thesis.

5.7.2.2. Self-reported criminal events

Average numbers of criminal events committed in the year after the intervention for each group are presented in Table 5.9 and Table 5.10. In Table 5.9 violent crimes and other criminal offences were aggregated. In Table 5.10 all criminal events were distinguished, and table reports: average frequency (SD) $_{\text{group size}}$ by intervention allocation. Minimal differences in total frequencies between the two tables are due to slightly different sample sizes: some individuals did not provide information across all criminal categories and were therefore dropped when aggregating the group of offences.

	CIL			**BA**			**BLC**		
	N	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**
Violent offences	86	0.77	2.17	100	1.32	6.18	83	0.94	3.75
Criminal offences	87	0.37	1.45	100	0.12	0.48	83	0.61	2.9
Total offences	86	1.14	3.12	100	1.44	6.23	83	1.55	6.31

Table 5.9 Summary of violent and criminal offences across groups

	CIL	BA	BLC
Violent offences			
Violence	0.38 (1.33) $_{88}$	0.28 (0.89) $_{102}$	0.24 (0.88) $_{87}$
Assault	0.36 (1.06) $_{88}$	0.37 (2.13) $_{102}$	0.41 (1.71) $_{87}$
Wounding	0.14 (0.46) $_{88}$	0.63 (3.63) $_{102}$	0.23 (1.25) $_{87}$
Sexual offences	0.01 (0.1) $_{87}$	0 (0) $_{100}$	0 (0) $_{83}$
Criminal offences			
Theft (non-vehicles)	0.07 (0.31) $_{87}$	0.03 (0.3) $_{100}$	0.46 (2.55) $_{83}$
Vehicles theft	0.03 (0.23) $_{87}$	0 (0) $_{100}$	0.02 (0.15) $_{83}$
Burglary	0.06 (0.45) $_{87}$	0.01 (0.1) $_{100}$	0 (0) $_{83}$
Criminal damage	0.06 (0.45) $_{87}$	0.02 (0.16) $_{100}$	0.07 (0.3) $_{83}$
Robbery	0.02 (0.15) $_{87}$	0.01 (0.1) $_{100}$	0.03 (0.24) $_{83}$
Shop lifting	0.01 (0.1) $_{87}$	0 (0) $_{100}$	0 (0) $_{83}$
Drug offences	0.03 (0.18) $_{87}$	0 (0) $_{100}$	0 (0) $_{83}$
Driving offences	0.05 (0.27) $_{87}$	0.03 (0.17) $_{100}$	0.01 (0.1) $_{83}$
Other fines	0 (0) $_{87}$	0.02 (0.14) $_{100}$	0 (0) $_{83}$

Table 5.10 Complete cases under the mean imputation scenario

The average number of criminal events per individual in the three groups were: 1.14 (PIL), 1.44 (BA) and 1.55 (BLC). While the proportion of violent episodes and criminal offences committed by individuals treated with CIL and BLC are comparable (approximately 65% violent crimes and 35% criminal offences), offenders treated with BA committed almost exclusively violent crimes. Looking at the specific crimes committed, violence and assaults are the most common events across all groups. Wounding and non-vehicle thefts have higher frequencies in BA and BLC groups, respectively. Even if the distribution of criminal events do not follow a Normal distribution, standard deviations help highlighting the great degree of uncertainty around these estimates. Standard deviations are very high when compared to mean estimates, and this indicates the presence of high variability in the groups.

5.7.2.3. Results of the analysis of self-reported criminal events

For each group, all the self-reported criminal events were aggregated and their frequency was assessed. Similarly to the analysis of PNC data conducted in Sub-section 5.6.3, Poisson, negative binomial, and zero inflated regression models were investigated.

The overdispersion coefficient was found to be statistically significant, therefore the Poisson regression was deemed inappropriate. Values of AIC and BIC associated with the NB and ZINB models were almost identical, even though a slightly lower BIC was associated with the NB model. Moreover, the appropriate test did not support the creation of a zero inflated negative binomial model. Consequently, a negative binomial regression was performed using the total number of self-reported offences in year following the intervention (complete cases only) as the dependent variable. Estimates were adjusted for self-reported baseline criminal events (presented previously in the first column of Table 5.3) and for the presence of clusters. Complete output of the regressions can be found in Appendix M.

IRR associated with BA was 1.24 (95% CI: 0.46, 3.32); IRR associated with BLC was 1.35 (95% CI: 0.41, 4.49). Offender treated with BA (BLC), are therefore expected to have an offending rate 1.24 (1.35) times greater compared to those treated with CIL. The resulting average self-reported criminal offences are summarised in Table 5.11.

Intervention allocation	Average criminal offences	Compared to CIL
CIL	1.16	-
BA	1.44	0.28
BLC	1.57	0.41

Table 5.11 Average self-reported criminal offences

5.7.3. Summary of effectiveness estimates

The analysis of PNC and self-reported data provide different insights about the impact of SIPS on criminal activity. As reported in Table 5.5, according to the PNC data, BA and BLC are both associated with lower reconviction episodes when compared to CIL. On the contrary, results from the analysis of self-reported data indicate that the average number of self-reported crimes is higher in BA and BLC groups compared to CIL. The uncertainty around the estimates is quite high.

Nevertheless, if the focus of the analysis is shifted from the frequencies of criminal events to the proportion of offenders re-committing at least one crime, PNC and self-reported data show similar patterns across the groups. In fact, from PNC data, being allocated to BA or BLC is associated with a lower risk of recidivism. Similarly, according to self-reported data, being allocated to BA or BLC is associated with a (non-statistically significant) reduction in the risk of committing another crime at least once within the year after the intervention. Details of the analysis are reported in Appendix N.

Even though self-reported and PNC data exhibit some differences in the frequencies of criminal events, the proportions of individual re-committing at least one crime are similar. As summarised in Table 5.12, ORs of getting reconvicted and re-offending at least once after the intervention are in fact lower than one for both BA and BLC groups compared to CIL group.

Intervention Allocation	OR (95% CI) of committing at least one crime	
	PNC data	**Self-reported data**
BA	0.53 (0.34, 0.83)	0.58 (0.27, 1.25)
BLC	0.60 (0.38, 0.95)	0.71 (0.33, 1.56)

Table 5.12 Probability of committing at least one crime after the intervention

Results might suggest that even if the number of crimes self-reported by offenders treated with CIL is lower when compared to BA and BLC, the proportion of individuals recommitting at least one crime is higher in the CIL group. Nevertheless, self-reported data are probably less reliable than the more objective records from the PNC. In fact, the reliability of self-reports offending surveys is typically based on the finding that individuals who self-report offenses are more likely to have official records than those who do not report offenses, and individuals with officially recorded crimes usually admit these crimes[34] (Jolliffe & Farrington, 2014). However, such a correlation between self-reported and official records was not observed in this analysis (see Sub-section 5.7.1). For this reason, in this analysis it has been decided to use PNC data as the main measure of effectiveness of the intervention.

The CEA will be therefore based on PNC data. However, as explained in the next section, the analysis will borrow information on criminal events from self-reported data. In fact, due to the lack of details on PNC records, self-reported data will be used to integrate the information about the nature of the criminal events.

[34] This evidence does not actually cover self-report surveys of probation workers with alcohol misuse issues. Moreover, self-reported offending surveys appear to be more valid for some demographic categories than for others.

5.8. Costs in response to crime

In this analysis only costs falling on CJS budget are considered. Other budget impacts such as HC costs (e.g. impacts on victims of crime), lost wages, and any other impact on private consumption and damaged or stolen properties are not considered. The inclusion of these components would imply considering other sectors in the analysis (e.g. HC or private consumption), whereas the focus now is on the impact of SIPS from the CJS perspective. This CEA conducted from a 'narrow' CJS perspective will be extended in the next chapter, where an integrated cross-sectoral analysis considering both CJ and HC sectors will be conducted. The potential inclusion of other dimensions such as private consumption (in terms of lost wages, property stolen or damaged) will be discussed in Chapter 6, Sub-section 6.3.3.3.

The use of CJS's resources was not collected alongside the SIPS trial. For this reason, total cost of SIPS falling on the CJS budget were estimated indirectly. As mentioned in Sub-section 5.3.2, total costs are composed by the incremental costs of the intervention and the costs in response to crime. The first component is investigated in Sub-section 5.8.1. To estimate the second component, PNC re-conviction records were used to determine the variation in criminal activity, and therefore the variations in CJS resource use. Nevertheless, PNC records of re-conviction episodes were not associated with a description of the type of crime committed. For this reason, self-reported questionnaires were used to investigate the potential distribution of specific criminal events. PNC records of criminal activity based on re-conviction frequencies were therefore integrated with a more specific list of self-reported criminal episodes (Sub-section 5.8.2). Afterwards, appropriate unit costs (described in Sub-section 5.8.3) were associated with each criminal event (Sub-section 5.8.4).

5.8.1. Costs of the intervention

In Chapter 4, Table 4.8, all decision makers and associated budgets involved in the programme were summarised. As discussed in Sub-section 4.4.4, costs of implementing SIPS were assumed to fall entirely on the CJS budget.

Costs of the intervention were computed already in Chapter 4, Sub-section 4.4.3. Costs of implementing SIPS were very low and are reported in Table 4.7. The cheapest treatment is CIL, which costs £0.2 per individual (the cost of the leaflet). BA and BLC are slightly more expensive: £2.2 and £13.2, respectively. Such estimates are likely to be a lower bound of the actual costs of the intervention. For this reason a wide range of variation around intervention costs is explored in the scenario analysis presented in Section 5.9.

5.8.2. Using self-reported data to characterise the distribution of criminal events

The impact of SIPS on criminal activity can be measured by assessing variations in reconvictions from PNC data. As reported in Sub-section 5.6.3.2, compared to CIL, average reconvictions avoided by BA and BLC were 0.51 and 0.66, respectively. On the contrary, to estimate the costs in response to crime, a more detailed description of the types of crime committed is necessary. However, the available records extracted from the PNC database do not provide any information about the specific offence committed for each reconviction episode.

Given the lack of detail about reconviction episodes, PNC records were integrated with the information provided by self-reported questionnaires. Self-reported data offered in fact an overall representation of the types of crimes committed more frequently by offenders treated with SIPS, and were used as a reference to characterise the distribution of criminal events recorded in the PNC database. However, the assumption that the distribution of criminal events followed the self-reported data was necessary only because of the lack of data. The need of a more accurate procedure to record resource use consumption in the CJS is further discussed in Chapter 8, Sub-section 8.5.1.

Categories of self-reported criminal events across all three intervention arms were summarised in Table 5.6. It was found that 75% of self-reported crimes were violent offences, roughly equally distributed among generic violence, assault and wounding. Thefts constituted most of the remaining criminal offences. Self-reported frequencies were used to integrate the PNC records and estimate the number of criminal events avoided by the intervention. These are reported in Table 5.14.

Once estimates of the specific criminal offences committed are available, appropriate unit costs can be associated.

5.8.3. Unit costs of criminal events

Estimates of the unit costs for each specific criminal event were taken from: the Home Office Research Study 217 (Brand & Price, 2000), the Home Office Online report 30/05 (Dubourg, Hamed & Thorns, 2005), an analysis of alcohol misuse costs (Leontaridi, 2003), and a National Audit Office technical paper (Imran Akhtar, 2011). For each criminal event, unit costs included all the costs in response to crime, namely: police activity, prosecution, magistrates' court, crown court, jury service, legal aid, non-legal-aid defence, probation service, prison service, other CJS costs, CJS overhead, and Criminal Injuries Compensation Authority costs. With regard to driving offences, cost components were: drink driving arrests, magistrates' courts, and crown courts (including sentencing). Cost per violent episode was assumed to be the same as the cost of an assault.

Unit cost estimates were updated to 2010 prices (the reference year of the case study) using the Bank of England Inflation Calculator[35]. As summarised in Table 5.13, the highest unit costs are associated with sexual and driving offences, followed by robbery, wounding, burglary in a dwelling and drug offences.

Category	**Unit cost *(in 2010 £)***	**Source**
Crime against individuals and households		
Wounding	2,189	Home Office 30/05
Sexual offences	4,067	Home Office 30/05
Common assault	314	Home Office 30/05
Violence	314	Home Office 30/05
Robbery	3,207	Home Office 30/05
Burglary in a dwelling	1,402	Home Office 30/05
Theft - not vehicle	371	Home Office 30/05
Theft of vehicle	245	Home Office 30/05
Criminal damage	155	Home Office 30/05
Commercial and public sector victimisation		
Theft from a shop	27	Home Office 217
Drug offences	2,601	NAO 2011
Driving offences	4,231	Leontaridi 2003
Other fines	283	NAO 2011

Table 5.13: Unit costs per criminal event

It is important to note that the main reference used for the estimation of the unit costs is quite old and the accuracy of results might be disputable (Bhattacharya, 2017). However, to my knowledge, no other more recent attempts of updating these cost estimates have been conducted. Moreover, the focus of this research is not on the values of the unit costs of criminal events themselves, but on how to use them to inform a CEA from the CJS perspective. If better evidence is provided, unit costs can be adjusted accordingly.

[35] Available at: https://www.bankofengland.co.uk/monetary-policy/inflation/inflation-calculator

The need for better quality data about costs of crime is further discussed in Chapter 8, Sub-section 8.5.1. In this analysis, uncertainty around cost estimates is addressed in the scenario analysis described in Section 5.9.

5.8.4. Cost (savings) in response to (reductions in) crime associated with SIPS

Unit costs described in the previous sub-section were linked to the avoided criminal events associated with BA and BLC. Resulting costs in response of crime falling on the CJS budget were thus estimated. Table 5.14 summarises the estimated criminal events avoided and the corresponding cost savings associated with BA and BLC.

Offences	**Events avoided**		**CJS cost savings (£)**	
Violent offences	**BA**	**BLC**	**BA**	**BLC**
Violence	0.111	0.143	34.8	45.0
Assault	0.141	0.182	44.2	57.2
Wounding	0.130	0.168	283.9	367.4
Sexual offences	0.001	0.002	5.4	7.0
Criminal offences				
Theft (non-vehicles)	0.065	0.083	23.9	31.0
Theft vehicles	0.007	0.009	1.6	2.1
Burglary	0.009	0.012	13.1	17.0
Criminal damage	0.019	0.025	3.0	3.8
Robbery	0.008	0.010	25.7	33.2
Shop lifting	0.001	0.002	0.0	0.0
Drug offences	0.004	0.005	10.4	13.5
Driving offences	0.012	0.016	50.8	65.8
Other fines	0.003	0.003	0.8	1.0
Total	**0.510**	**0.660**	**497.7**	**644.1**

Table 5.14 SIPS costs in response to crime

As illustrated in Sub-section 5.6.3.2, it was found that both BA and BLC were associated with a reduction in criminal activity. In this analysis, cost savings triggered by the reduction in criminal activity were estimated to be approximately £498 and £644 for BA and BLC, respectively. Moreover, a 'generic' cost of £976 per SIPS reconviction episode was also estimated. This was obtained as a weighted average of the specific unit costs, using the frequency of events self-reported by SIPS offenders as weights.

5.9. Cost-effectiveness analysis of SIPS from the CJS perspective

5.9.1. A note on the relationship between effectiveness and total costs

As explained in Sub-section 5.3.2, total costs falling on the CJS budget (ΔC_{CJS}) are composed by the incremental costs of the intervention and the costs in response to crime. This second component is a function of the direct effects (i.e. reconvictions avoided) of the intervention. In particular, because of the way costs were estimated in this analysis, it is acknowledged that there is a linear function linking the effectiveness output and total costs. In fact, for reasons related to data availability, total costs falling on CJS budget were simply estimated as shown in Equation 5.4. In this formulation benefits are measured in terms of reconvictions avoided (ΔR). For this reason, costs in response to crime are treated as cost-savings associated with each reconviction episode avoided.

$$\Delta C_{CJS} = \Delta R * (-Reconviction\ cost) + Intervention\ costs \qquad (5.4)$$

As a consequence, both differential direct effects and differential costs depend on recidivism. However, the relationship between effectiveness and cost savings does not imply any double counting. Actually, as pointed out already in Section 5.2.2, costs (savings) due to (reductions in) recidivism must be added to the incremental intervention costs and contribute to the total impact on CJS budget. Avoided reconvictions (and not the related cost savings) are a proxy of criminal activity and measure the effectiveness of the intervention.

5.9.2. Adapting the methods to the analysis of SIPS

Incremental costs and effectiveness of the alternative intervention allocations can be compared using the CEA methods illustrated in Section 5.3. Similarly to the HC context, resulting Incremental Net Benefits (INBs) from the CJS perspective can be expressed in natural units (i.e. reconvictions) or in monetary terms. If expressed in monetary terms, direct effects (falling on offenders treated with SIPS) can be computed as shown in Equation 5.5, where k is the marginal productivity of the CJS.

$$DEs = \Delta R * k \qquad (5.5)$$

The two components of the opportunity costs (falling on other unknown offenders) are calculated as illustrated in Equations 5.6 and 5.7, where C_R is the 'generic' cost of reconviction for SIPS (see Sub-section 5.8.4).

$$Cost\ in\ response\ to\ crime = \Delta R * (-C_R) \qquad (5.6)$$

$$Costs\ of\ the\ intervention = \Delta C_{int} \quad (5.7)$$

Incremental Net Reconviction Benefits (INRB) and Incremental Net Monetary Benefits (INMB) can be therefore obtained as reported in Equations 5.8 and 5.9, respectively.

$$INRB = \Delta R - \left(\frac{\Delta R*(-C_R)}{k} + \frac{\Delta C_{int}}{k}\right) \quad (5.8)$$

$$INMB = \Delta R * k - (\Delta R * (-C_R) + \Delta C_{int}) \quad (5.9)$$

Once an estimate of *k* is available, the cost-effectiveness of BA and BLC compared to CIL can be assessed.

5.9.3. Results of the CEA from the CJS perspective

As illustrated in Section 5.4, the required cost-effectiveness threshold representing the marginal productivity *k* of the CJS is not available. Nevertheless, from the analysis conducted in Sub-section 5.4.2, a pragmatic solution could be to assume that, at the margin, with £ 11,000 CJS can prevent one episode of recidivism, measured in terms of avoided reconvictions. In other words, the marginal productivity of the CJS is £ 11,000 per criminal event averted.

Table 5.15 provides a summary of all variables involved in the analysis and estimated in the previous sections.

Description	**Variable**	**CIL**	**BA**	**BLC**
Intervention costs (falling on CJS)	C_{int}	£0.2	£2.2	£13.2
Incremental intervention costs	ΔC_{int}	-	£2	£13
Reconvictions (95% CI)	R	1.57 (1.15, 1.99)	1.06 (0.74, 1.38)	0.91 (0.62, 1.21)
Reconvictions avoided	ΔR	-	0.51	0.66
Cost per reconviction (specific for SIPS crimes)	C_R	£976		
Marginal productivity of the CJS	k	£11,000/R		

Table 5.15 Summary of the variables

Looking at the DEs of SIPS on offenders, when compared to CIL, both BA and BLC are more effective (i.e. ΔR is positive for both, due to a greater reduction in recidivism) suggesting that more intense treatments help reduce crime.

With regard to the OCs, two components can be distinguished. More intense intervention allocations (BA and BLC) are more costly to implement (i.e. ΔC_{int} is positive for both). Consequently, to accommodate the additional cost of the treatment, the CJS must displace resources that could have been invested otherwise to achieve potential reductions in recidivism. A second component of the OCs is represented by the cost savings due to the reduction in criminal events.

DEs (falling on SIPS offenders) and OCs (falling on unknown offenders) were calculated using Equations 5.5, 5.6 and 5.7. INRBs and INMBs were obtained as illustrated in Equations 5.8 and 5.9, respectively. Resulting INBs and their components are reported in Table 5.16, and are expressed both in natural (i.e. recidivism) and monetary units.

	Natural units (Reconvictions)			**Monetary units (£)**		
	DE	**OC**	**INRB**	**DE**	**OC**	**INMB**
BA vs CIL	0.51	- 0.045	0.555	5610	- 496	6106
BLC vs CIL	0.66	- 0.057	0.717	7260	- 631	7891

Table 5.16 DEs, OCs and INBs of SIPS

Compared to CIL, both interventions are associated with more reconvictions avoided, and therefore to positive monetary impacts. With regard to the impacts on unknown offenders, both interventions are associated with negative OCs. This is because cost savings due to reduction in recidivism outweigh the costs of implementing the intervention. Consequently, both BA and BLC are associated with cost-savings, and therefore to further reductions in re-offending (due to additional resources being used by the CJS in other programmes).

Results presented in Table 5.16 did not include the uncertainty around the estimates. To reflect the uncertainty around the effectiveness of the interventions (which therefore translates into uncertainty about DEs and costs in response to crime), the 95% CIs of the INMBs were also computed. It must be pointed out that the covariance between costs and outcomes is equal to one, because costs directly depend on the same outcome variable (i.e. reconvictions). Moreover, in Sub-section 5.4.2, a marginal productivity of £ 11,000 per criminal event averted was estimated, and this value was used in the CEA. Nevertheless, such a value is highly uncertain and the impact of the intervention under alternative assumptions about the marginal productivity of the CJS should be explored.

INMBs for BA and BLC and their 95% CIs are presented in Figure 5.3. INMBs are expressed as a function of a range of potential values for *k*, going from 100 to 20,000 £ per reconviction averted.

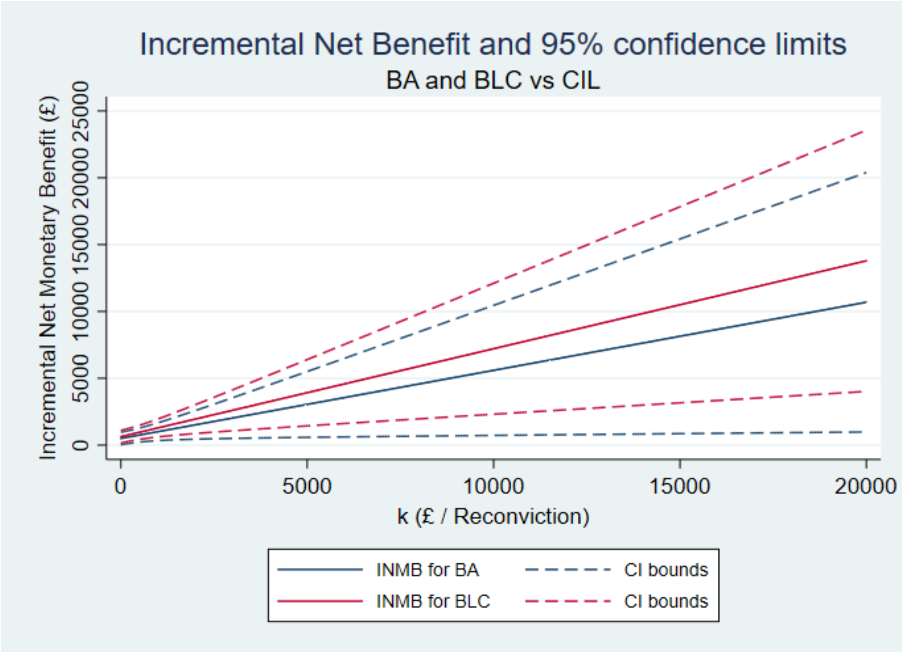


Figure 5.3 INMBs for BA and BLC

INMBs for both BA and BLC are positive and the 95% CIs do not include zero for all scenarios investigated. BLC appears to be the overall cost-effective option. For the same value of *k*, BLC is associated with greater INMBs compared to BA. Moreover, the higher the value of *k*, the greater the INMBs for both interventions compared to CIL. In fact, higher values of *k* mean that more money is required by the CJS to prevent one additional reconviction. Consequently, interventions that are more effective in reducing crime are associated with higher INMBs. A supplementary note about the interaction between *k* and cost components of SIPS is also provided in Appendix O. The notes explain how the magnitude of costs savings and incremental intervention costs affects the trends of INRB and INMB when *k* increases.

5.9.3.1. Scenario analyses

To explore the robustness of these results, two scenario analyses were carried out. Details can be found in Appendix O.

As pointed out in Sub-section 5.8.1, intervention costs might have been underestimated. A range of alternative costs of the intervention was therefore investigated in the first scenario analysis. It was found that INMBs were always positive, even when the costs of implementing the intervention were much higher. A second scenario analysis explored different costs in response to crime. In fact, as highlighted in Sub-section 5.8.3, estimates of the CJS costs associated with each re-conviction episode might be questioned. It was found that INMBs of BA and BLC compared to CIL were likely to be positive even if costs in response to crime (and therefore cost savings associated with BA and BLC) were lower.

5.10. Conclusions

While from the health care prospective both BA and BLC were found to be not cost-effective compared to CIL (see Chapter 4), from the CJS perspective, INBs for both BA and BLC are positive. More specifically, the INB associated with BLC is higher than the INB associated with BA. It is therefore worth investing both in BA and BLC when compared to CIL, but the overall cost-effective intervention is BLC.

From this analysis, it can be concluded that from the CJS perspective it would be probably cost-effective to invest in BLC. These results are based on several assumptions, mostly forced by the lack of data. Assumptions have been however explicitly stated, and sensitivity and scenario analyses have been conducted. Potential issues, limitations and policy implications of this analysis are further discussed in Chapter 8.

Further analyses from the CJS perspective could be performed (e.g. MI to address missing data issues, or VOI analysis), but are not strictly necessary to conduct the cross-sectoral economic evaluation proposed in this thesis. These analyses are therefore left as potential further research. In the next chapter, the cross-sectoral analytical framework proposed in Chapter 3 is operationalised using results from Chapter 4 and 5.

Chapter 6: Cross-sectoral analysis of SIPS

6.1. Cross-sectoral analysis of an intervention to reduce alcohol misuse

Chapter 2 firstly introduced methods and approaches to economic evaluation, and then illustrated from a purely theoretical point of view the issues in assessing public health interventions with impacts on multiple sectors. The focus of Chapter 3 was on the description of a theoretical framework for a cross-sectoral and cross-temporal cost-effectiveness analysis of public health interventions, with the inclusion of a health equity assessment.

The framework proposed in Chapter 3 would be particularly useful for the evaluation of alcohol interventions, where the need to consider wider societal perspectives, impacts on health inequity, and long-term outcomes has been highlighted (Hill et al., 2017). From a theoretical point of view, all three of these concerns (i.e. wider perspectives, impacts on inequity and cross-temporal) could be addressed using the proposed framework. However, the objective of this chapter is to focus on the cross-sectoral stage of the framework, and show how to pragmatically address the issue of conducting an evaluation from a wider perspective, using the SIPS case study. The other potential extensions of the cross-sectoral analysis are explored in the next two chapters. More precisely, the inclusion of health equity concerns in the economic analysis will be investigated in Chapter 7, whereas the cross-temporal aspects will be discussed in Chapter 8.

6.1.1. Structure of the chapter

From Chapter 4 onwards, a case study using SIPS trial was selected as a vehicle to demonstrate the framework proposed in Chapter 3. SIPS trial aimed at evaluating interventions to reduce alcohol consumption among offenders. Details about the SIPS trial can be found in Chapter 3, Section 3.4.

In Chapter 4, a 'traditional' CEA from the HC perspective was conducted. This perspective considered health-related costs and health effects on the offenders alone. In Chapter 5, a similar analysis was carried out, but from the CJS perspective. The CEA from the CJS perspective assessed costs falling on CJS budget and used recidivism (in terms of avoided reconvictions) as a measure of effectiveness. Figure 6.1 summarises the main variables assessed in the two independent CEAs conducted in the previous chapters.

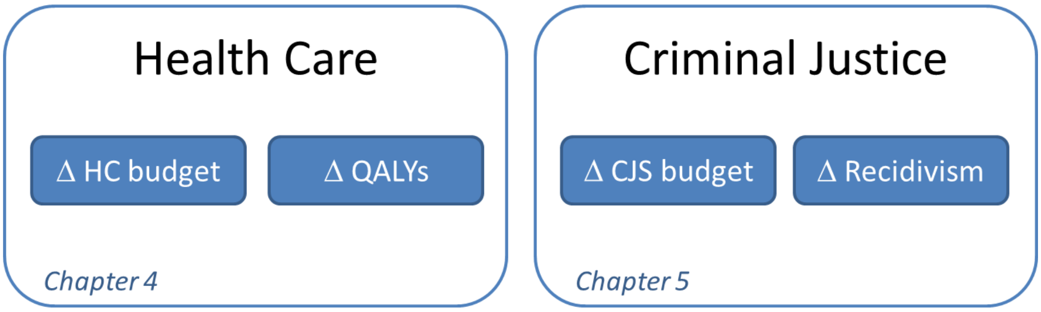


Figure 6.1 CEAs from HC and CJS perspectives

The analyses carried out in this chapter build on the evaluations conducted in Chapters 4 and 5. The chapter is composed of two parts: firstly, cross-sectoral data are used to revise the CEA from the narrow HC perspective; secondly, a cross-sectoral CEA from an integrated perspective is conducted.

6.1.1.1. Part 1: using cross-sectoral data to revise the CEA from the HC perspective

In Section 6.2, the 'traditional' CEA from the HC perspective (conducted in Chapter 4) is revised. The objective is still to maintain a 'narrow' HC perspective, but to expand the CEA to incorporate victims' health costs and effects as well. In fact, in the CEA conducted in Chapter 4, the focus was only on costs falling on HC budget and on QALYs impacts for trial participants (i.e. the offenders treated with SIPS). This approach failed to reflect the potential health and health care consequences on victims through impacts on the CJS. To estimate the true impacts on health and the HC budget, consequences on health of the victims and additional costs falling on HC budget should actually be considered, even if taking a narrow HC perspective.

In the revised CEA, impacts on crime victims in terms of QALYs lost and additional costs falling on HC budget are estimated and included in the evaluation, and revised incremental net benefits for each population subgroup from the HC perspective are computed. Such an adjustment of the CEA from the HC perspective requires the modelling of costs and consequences from the CJS perspective first. This is because health spill-over effects through the CJS can only be computed after modelling impacts on criminal offences and opportunity costs in CJS.

The analysis conducted in the first part of this chapter is therefore based on cross-sectoral data, but the perspective of the analysis is still the narrow viewpoint of the HC system. An analysis from an integrated perspective is conducted in the second part.

6.1.1.2. Part 2: cross-sectoral CEA conducted from an integrated perspective

In Section 6.3, an analysis across both sectors from an integrated perspective is conducted. Such an evaluation has the objective of showing the consequences (e.g. differences in the recommended decisions) of conducting a CEA from an integrated perspective, instead of conducting multiple fragmented and independent analyses, as is typically the case.

On the basis of the theoretical framework proposed in Chapter 3, cross-sectoral costs and benefits are made explicit and results of the CEAs from HC and CJS perspectives are aggregated. Potential compensation schemes that consider both HC and CJS costs and consequences are proposed to address resource allocation issues. In fact, when costs fall on one sector and benefits another, compensations across decision makers can lead to funding potentially cost-effective interventions.

6.2. Adjustment of the CEA from the HC perspective

The aim of this section is to revise the ‘traditional’ CEA of SIPS performed in Chapter 4, and include HC costs and consequences on victims, previously ignored. It must be stressed that this ‘revised’ CEA is still conducted from a narrow HC perspective, even though cross-sectoral data are used. The reason is that it is necessary to carry out an analysis from the CJS perspective first, in order to estimate the consequences of crime on victims and the CJ opportunity costs of funding an intervention.

From the HC perspective, three groups of individuals affected by the intervention can be identified: the offenders treated with the intervention, the general population bearing the opportunity costs and the victims of crimes. Victims are in turn composed of two subgroups: direct and unknown victims. The former are victims of crimes committed by offenders treated with SIPS; the latter bear the consequences of the opportunity costs generated in the CJS. In other words, when CJS displaces resources to accommodate the costs of a new intervention, unknown offenders will commit crimes (and hurt unknown victims) as the resources for their programmes are displaced. To model the impacts on unknown victims it is therefore necessary to assess the opportunity costs in CJS.

Figure 6.2 outlines the revised CEA scheme. Two CEAs were previously conducted, one from the HC perspective and another one from the CJS perspective. As illustrated in Chapter 5, the outcomes of interest in a CEA conducted from the CJS perspective are the impacts on CJS budget and recidivism. Nevertheless, consequences on (direct and unknown) victims can be computed as well, and these can be used to adjust the ‘traditional’ CEA from the HC perspective.

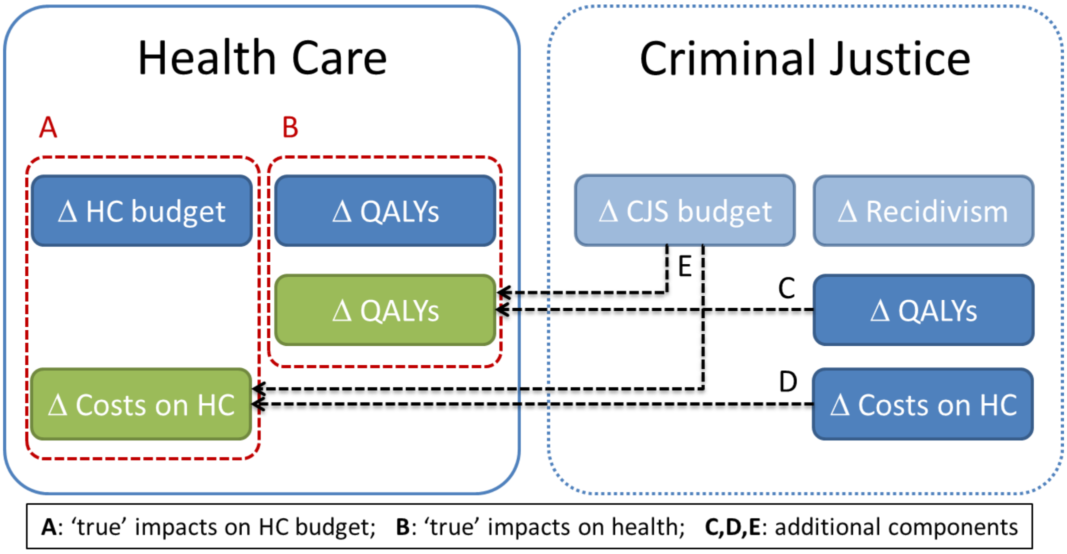


Figure 6.2 Scheme for the revised CEA from HC perspective

The focus of this analysis is on the impacts from the HC perspective. In Figure 6.2, the 'true' impacts on HC budget and health are described by the dashed rectangles A and B, respectively. These include the estimates from the original 'traditional' CEA, and the additional components (the green rectangles) that are usually (incorrectly) ignored. These additional components comprise: direct impacts on the victims' health (C); additional health opportunity costs due to the additional treatment necessary to assist the direct victims (D); health impacts and additional health opportunity costs related to crime involving unknown victims, due to the CJ opportunity costs of the intervention (E).

6.2.1. Methods for the estimation of the revised INHBs from the HC perspective

The general formula to calculate the INHB with the inclusion of health spill-over effects generated by a non-health sector was presented in Chapter 3, Equation 3.4. This formula can be adapted to the specific context of a CEA conducted from the HC perspective that includes health spill-over effects generated in the CJS.

Table 6.1 summarises DEs and OCs falling on the four groups previously identified. In addition to the DE on offenders treated with the intervention (ΔH_O), additional DEs on victims are included. ΔH_V represents the health effects (measured in QALYs) on direct victims. General population impacts depend on the total variations in the HC budget. Compared to the 'traditional' CEA, two additional components are introduced: the additional costs due to HC services for victims (ΔC_{Hv}) and unknown victims (ΔC_{Hvu}). Moreover, OCs in CJS also generate health effects (measured in QALYs) on unknown victims (ΔH_{Vu}). Table 6.1 reports the variables measured in natural units (i.e.

QALYs). Costs are therefore expressed in QALYs using the marginal productivity of the HC sector, k_H (Claxton, Martin, et al., 2015).

Group affected	DE	OC
Offenders	ΔH_O	
Direct victims	ΔH_V	
Unknown victims		ΔH_{Vu}
General population		$\frac{\Delta C_H + \Delta C_{Hv} + \Delta C_{Hvu}}{k_H}$

Table 6.1 Health impacts on population subgroups

As shown in Equation 6.1, the revised INHBs can be obtained as the DEs net of the OCs.

$$INHB = (\Delta H_O + \Delta H_V) - \left(\frac{\Delta C_H + \Delta C_{Hv} + \Delta C_{Hvu}}{k_H} + \Delta H_{Vu}\right) \quad (6.1)$$

In this analysis, all impacts are simply aggregated on the basis of the principle that "a QALY is a QALY", irrespectively of who is the beneficiary. Nevertheless, as discussed in Chapter 3, Section 3.3, the analysis can be underpinned by alternative value judgments, and potential equity concerns can be incorporated in the evaluation. In the next chapter, alternative solutions and considerations about how to aggregate health impacts falling on different population subgroups will be explored.

In the two following sub-sections, impacts on direct and unknown victims are estimated. Revised INBs are presented in Sub-section 6.2.4.

6.2.2. Impacts on direct victims of SIPS

As mentioned previously, impacts on direct victims consists of two components: direct impacts on the victims' health, and additional health OCs due to the additional treatments necessary to assist the direct victims.

QALY losses experienced by victims of various criminal offences are illustrated in Sub-section 6.2.2.1. Additional costs falling on HC budget due to the treatment of victims of violence are described in Sub-section 6.2.2.2. Resulting impacts of SIPS on direct victims are presented in Sub-section 6.2.2.3.

6.2.2.1. QALYs lost by victims of crime

Various methods can be used to estimate crime's impacts on health such as the WTP approach (see Sub-section 5.4.2.2) or the QALY approach (Wickramasekera et al., 2015). To be consistent with the

CEA methods in the HC context, the QALY approach has been chosen for this analysis (Dolan & Peasgood, 2007, Dolan, Loomes, et al., 2005).

Dolan et al. (2005) investigated the intangible costs of crime for the victims and provided estimates of the discounted QALY losses experienced by the victims of the main crime categories. Impacts on victims' health associated with each criminal event are summarised in Table 6.2.

Crime category	Discounted QALY loss
Homicide	17.791
Wounding	0.111
Sexual offences	0.160
Rape	0.561
Common assault	0.007
Violence	0.007
Robbery	0.028
Other criminal offences (e.g. burglaries in a dwelling, thefts, criminal damages…)	0.000

Table 6.2 QALY losses associated with violent crime, adapted from Dolan et al. (2005)

Due to the lack of data, in this analysis it was assumed that common assaults and generic violent episodes were associated with the same health impact on victims. Moreover, because Dolan et al. (2005) did not report any impacts quantifiable in QALYs associated with burglaries in a dwelling, thefts, criminal damages and other commercial and public sector victimisations, it was assumed that no QALY loss was associated with these events. Lastly, no direct impact on health was associated with driving offences because it was assumed that their impacts were already included in the estimate of generic wounding. The estimate for generic wounding was in fact obtained as an average of 'serious' (car accidents included) and 'other' wounding (Dolan, Loomes, et al., 2005).

In this analysis, it is assumed that all health consequences on crime victims are described by the QALY losses presented in Table 6.2. Nonetheless, theoretically, consequences on victims' health could be further explored and additional components could be incorporated in the analysis. Potential estimates of fear of crime can be found in the literature (Dolan & Peasgood, 2007) and impacts on mental distress could be investigated. This is further discussed in Chapter 8, Sub-section 8.4.1.3.

6.2.2.2. Additional HC costs to treat victims of crime

Criminal events are not only associated with direct health losses for the victims, but also to additional costs (falling on the HC budget) to treat victims of violence. These costs are usually neglected, but are actually additional HC opportunity costs and should be considered in the analysis. Additional HC costs fall on unknown patients that cannot be treated due to the resources displaced by the HC system to accommodate the additional treatments for the victims of crime.

Average costs falling on HC budget to treat victims of violent crimes were computed using the estimates from the Home Office Online Report 30/05 (Dubourg, Hamed & Thorns, 2005). All unit cost estimates were updated to 2010 prices (the reference year for the case study) using the appropriate inflator index, and are summarised in Table 6.3.

Crime category	Average costs for health services *(£, 2010)*
Homicide	950
Wounding	1,662
Sexual offences	1,129
Rape	2,567
Common assault	152
Violence	152
Robbery	596
Other criminal offences (e.g. burglaries in a dwelling, thefts, criminal damages…)	0

Table 6.3 Average HC costs associated with criminal events

From the HC perspective, the most expensive criminal event is rape, followed by wounding and generic sexual offences. These values should be interpreted with caution because the accuracy of the original estimates might be arguable and no confidence intervals were provided. However, to the best of my knowledge, no other more recent attempts of updating these cost estimates have been found in the literature.

6.2.2.3. Estimates of the impacts on direct victims

As illustrated in Chapter 5, Sub-section 5.8.2, PNC reconviction data were integrated with self-reported information in order to characterise the distribution of criminal events committed by the offenders, and estimate the number of criminal events avoided by SIPS. Compared to CIL, BA and BLC were associated with a reduction of 0.51 and 0.66 reconvictions, respectively (see Chapter 5, Sub-section 5.6.3.2). When PNC data were integrated with the frequencies of self-reported crimes (summarised in Table 5.6), the number of criminal events avoided were computed.

Criminal events avoided are summarised again in the first columns of Table 6.4. Appropriate discounted QALY losses (Dolan, Loomes, et al., 2005) and average health services costs (Dubourg, Hamed & Thorns, 2005) were linked to each criminal event avoided. Table 6.4 illustrates the resulting avoided victims' QALY losses and the HC cost savings corresponding to the criminal events avoided with BA and BLC compared to CIL.

Offences	Events avoided		Avoided victims' QALY losses		HC cost savings (£)	
Violent offences	**BA**	**BLC**	**BA**	**BLC**	**BA**	**BLC**
Violence	0.111	0.143	0.001	0.001	16.8	21.7
Assault	0.141	0.182	0.001	0.001	21.3	27.6
Wounding	0.130	0.168	0.014	0.019	215.6	279.0
Sexual offences	0.001	0.002	0.000	0.000	1.5	2.0
Criminal offences						
Theft (non-vehicles)	0.065	0.083	-	-	-	-
Theft vehicles	0.007	0.009	-	-	-	-
Burglary	0.009	0.012	-	-	-	-
Criminal damage	0.019	0.025	-	-	-	-
Robbery	0.008	0.010	0.000	0.000	4.8	6.2
Shop lifting	0.001	0.002	-	-	-	-
Drug offences	0.004	0.005	-	-	-	-
Driving offences	0.012	0.016	-	-	-	-
Other fines	0.003	0.003	-	-	-	-
Total	**0.510**	**0.660**	**0.017**	**0.021**	**260.0**	**336.5**

Table 6.4 Impacts of BA and BLC on victims, compared to CIL

Compared to CIL, BA and BLC were associated on average with cost savings of £ 260 and £ 336.5, and health gains of 0.017 and 0.021 QALYs, respectively. According to the distribution of avoided criminal events, the greatest benefits were associated with wounding episodes avoided.

6.2.3. Impacts on unknown victims of SIPS

A similar approach was followed to quantify the impact on unknown victims, generated by the opportunity costs in the CJS (i.e. what could have been obtained by using the CJS resources for an alternative purpose). Because both BA and BLC are associated with cost savings, more CJS resources would be available if BA or BLC were implemented (opportunity costs are negative). Additional reconvictions could be therefore avoided.

6.2.3.1. Frequencies of 'generic' criminal events

The distribution of criminal events reported in Table 5.6 refers specifically to the criminal activity associated with SIPS. These frequencies might be potentially appropriate for other alcohol interventions, but cannot be used to estimate the impacts on unknown victims resulting from OCs in CJS. In fact, when the resources for their programmes are displaced, unknown 'generic' offenders might commit all sorts of crimes. The frequencies of criminal events committed by generic offenders will be different from the criminal events committed by offenders treated with SIPS.

The 2011 Home Office Statistical Bulletin (Chaplin, Flately & Smith, 2011) was used to investigate the frequencies of criminal events committed in the year 2010 (the reference year for the case study). As summarised in Table 6.5, in 2010 in the UK more than 4 million criminal events were recorded. Of these, homicide accounted for 0.01%, wounding 1.03%, assault 8.15%, rape 0.35%, sexual assault 0.90%, robbery 1.72%, other offences without consequences on health (e.g. criminal damage, violence without injury and thefts) 87.83%.

Crime category	Events	Category from HORS	Proportion
Homicide	618	Homicide	0.01%
Wounding	45,043	Violence against the person with injury (excluding homicide and ABH)	1.03%
Assault	355,968	Actual bodily harm	8.15%
Rape	15,084	Rape of a female and male	0.35%
Sexual assault	39,271	Sexual offences (excluding rape)	0.90%
Robbery	75,105	Robbery	1.72%
Other offences without consequences on health	3,834,283	Burglary, violence against the person without injury, offences against vehicles, other theft , fraud and forgery, criminal damage, drug and other miscellaneous offences	87.83%
Total	4,365,372		100.00%

Table 6.5 Frequencies of criminal events committed in the UK in 2010

These frequencies provide an overall idea of the types of crimes committed more frequently by 'generic' offenders in the UK.

6.2.3.2. Estimates of the impacts on unknown victims of SIPS

OCs in CJS can be expressed either in monetary or natural units. When expressed in natural units, OCs represent the (avoided) reconvictions due to costs (savings) falling on the CJS budget. As reported previously in Table 5.16, compared to CIL, BA and BLC were associated on average to 0.045 and 0.057 avoided reconvictions, respectively.

Frequencies of criminal events for generic offenders summarised in Table 6.5 were used to estimate the numbers of specific criminal events avoided. Afterwards, appropriate discounted QALY losses and average health services costs (illustrated in Sub-sections 6.2.2.1 and 6.2.2.2, respectively) were linked to each criminal event. Resulting impacts on unknown victims in terms of QALY losses avoided and HC cost savings are shown in Table 6.6.

Crime category	Events avoided		Avoided QALY losses by victims		HC cost savings (£)	
	BA	**BLC**	**BA**	**BLC**	**BA**	**BLC**
Homicide	0.000	0.000	0.0001	0.0001	0.0	0.0
Wounding	0.000	0.001	0.0001	0.0001	0.8	1.0
Assault	0.004	0.005	0.0000	0.0000	0.6	0.7
Rape	0.000	0.000	0.0001	0.0001	0.4	0.5
Sexual assault	0.000	0.001	0.0001	0.0001	0.5	0.6
Robbery	0.001	0.001	0.0000	0.0000	0.5	0.6
Other offences without consequences on health	0.040	0.050	0.0000	0.0000	0.0	0.0
Total	**0.045**	**0.057**	**0.0004**	**0.0005**	**2.7**	**3.4**

Table 6.6 Impacts on unknown victims of BA and BLC compared to CIL

Average impacts on unknown victims' health and HC cost savings are close to zero. From a pragmatic point of view, health effects associated with the OCs in the CJS are probably negligible. Nonetheless, the focus of this analysis is more on the methodological approach, rather than on the numerical results of the analysis.

6.2.3.3. Marginal productivity of CJS in producing health

The magnitude of health and HC costs associated with the OCs generated in the CJS is linked to the inter-sectoral marginal productivity of CJS in producing health.

As explained in Sub-section 5.4, the marginal productivity of the CJS in avoiding reconvictions was estimated to be £ 11,000 per unit of effect. In other words, with £ 11,000 CJS can prevent one episode of recidivism (i.e. one reconviction episode) at the margin. Avoided recidivism is associated with impacts on victims that can be quantified with the methods illustrated in this section. Using the frequencies summarised in Table 6.5, it can be estimated that for each reconviction avoided, the CJS avoids also 0.0081 QALYs lost by victims, and generates cost savings for approximately £ 59. Knowing the marginal productivity of the HC system k_H (i.e. with £ 13,000 HC sector can generate 1 QALY), these cost savings correspond to additional 0.0045 QALYs gained.

A proxy for the inter-sectoral marginal productivity of CJS in producing health (k_{RH}) can be computed. On average, at the margin, with an investment of £ 11,000, the CJS generates 0.0126 QALYs through avoided reconvictions. Consequently, k_{RH} is approximately equal to 870,000 £/QALY.

As expected, k_{RH} is much higher than k_H. In fact, it is reasonable to assume that the most efficient way to generate QALYs is operating through the HC sector. In other words, if the main objective is to improve health, investments should be made in the HC sector. Nevertheless, the CJS can also generate QALYs, and k_{RH} can help to quantify them.

6.2.4. Revised estimate of incremental net benefits from HC perspective

6.2.4.1. Results of the 'traditional' CEA

Estimates of differential QALYs and costs of BA and BLC compared to CIL obtained from the 'traditional' analysis conducted in Chapter 4 were reported in Table 4.13.

In summary, compared to CIL, BA was associated with a negative impacts on offenders' health (on average: -0.010 QALYs), and it was also more expensive (on average £ 373.6 additional costs). BLC was also associated with negative impacts on health (on average: -0.018 QALYs), but it was cost saving (on average £ 226.6 cost savings). Overall, INBs associated with both BA and BLC were negative.

6.2.4.2. Revised results of the CEA from the HC perspective

Revised INBs of BA and BLC from the HC perspective were obtained as the difference between DEs and OCs. Details about methods used to compute the revised INBs were described in Sub-section 6.2.1.

INBs can be expressed either in natural (INHB) or monetary (INMB) units. Impact matrices and revised INHBs are presented in Table 6.7 for BA and BLC, compared to CIL. Revised INMBs from the HC perspective are reported in Table 6.8. In the first column of each table, the estimates from the 'traditional' CEA are reported [36]. 'Traditional' results are then adjusted for the additional components, previously ignored.

[36] If compared to the original results summarised in Table 4.13, small differences in the 'traditional' estimates of the INBs are due to rounding. For example, 'traditional' INMBs in Table 4.13 were -509.4 for BA and -8.5 for BLC, whereas in Table 6.8 these are -503.6 for BA and -7.4 for BLC.

Group affected	'Traditional' estimate		Additional component		Total	
	DE	OC	DE	OC	DE	OC
Brief advice (BA) vs CIL						
Offenders	-0.010				-0.010	
Victims			0.017		0.017	
Unknown victims				-0.0004		-0.0004
General population		0.029		-0.020		0.009
INHBs	**-0.039 QALYs**		**0.037 QALYs**		**-0.002 QALYs**	
Brief Lifestyle Counselling (BLC) vs CIL						
Offenders	-0.018				-0.018	
Victims			0.021		0.021	
Unknown victims				-0.0005		-0.0005
General population		-0.017		-0.026		-0.044
INHBs	**-0.001 QALYs**		**0.048 QALYs**		**0.048 QALYs**	

Table 6.7 Revised INHBs from the HC perspective

Group affected	'Traditional' estimate		Additional component		Total	
	DE	OC	DE	OC	DE	OC
Brief advice (BA) vs CIL						
Offenders	-130				-130	
Victims			215.7		215.7	
Unknown victims				-4.7		-4.7
General population		373.6		-262.6		111.0
INMBs	**-503.6 £**		**483.1 £**		**-20.5 £**	
Brief Lifestyle Counselling (BLC) vs CIL						
Offenders	-234.0				-234.0	
Victims			279.2		279.2	
Unknown victims				-6.0		-6.0
General population		-226.6		-339.8		-566.4
INMBs	**-7.4 £**		**625.0 £**		**617.6 £**	

Table 6.8 Revised INMBs from the HC perspective

In the proposed formulation, costs savings are expressed as negative OCs. For example, an intervention like BLC, which is on average cost saving from both HC and CJS perspectives, is associated with negative OCs (i.e. positive impacts) falling on both the general population and unknown victims (i.e. is associated with an increase in QALYs due to cost savings).

Looking at the additional components of the analysis, convictions avoided by BA and BLC engender QALY gains for both direct and unknown victims. These corresponds to the avoided QALY losses associated with avoided criminal events. Moreover, additional HC costs savings are generated, which fall on the general population. For example, impacts associated with BA falling on the general population were calculated as summarised in Table 6.9.

OCs falling on general population	Amount (£)	Source
Traditional estimate:		
HC costs from the 'traditional' CEA	373.6	Table 6.8
Additional components:		
HC cost to treat direct victims	-260.0	Table 6.4
HC cost to treat unknown victims	-2.7	Table 6.6
Total additional component	*-262.6*	Table 6.8 [37]
Total impacts of BA on general population	**111.0**	

Table 6.9 Impact on general population associated with BA

6.2.4.3. Conclusions of the revised CEA from the HC perspective

Looking at the total INHBs (last column in Table 6.7) and total INMBs (last column in Table 6.8), INBs are greater compared to their 'traditional' estimates. While average INBs of BA remain negative, the impact of the additional components is particularly relevant for BLC. In fact, in the traditional CEA, BLC was associated with negative INBs, while revised INBs are now positive. In other words, even from a narrow HC perspective, failing to account for the effects on health via criminal justice results in a failure to consider the full implications for the outcome of interest. If the full impacts on health are accounted for, BLC becomes on average cost-effective from the HC perspective. The conclusion of the revised CEAs would be to fund BLC, being the intervention that provides the highest INBs, compared to CIL and BA.

It must be stressed that these results are based on mean estimates. Further sensitivity analyses should be therefore conducted in order to investigate the uncertainty around these revised estimates and the robustness of the results. However, the focus of this analysis is more on the methods and less on the numerical results.

[37] The sum of 260.0 and 2.7 is 262.6 because, more precisely, cost savings components were 259.987 and 2.656. Resulting total cost savings amounted to 262.643.

6.3. Cross-sectoral CEA of SIPS from an integrated perspective

The analysis carried out in Section 6.2 was based on cross-sectoral data, but it was still conducted from the narrow HC perspective. Cross-sectoral data were only used to estimate health impacts on (direct and unknown) victims of crimes, and to adjust the amount of the HC opportunity costs.

The objective of this section is to carry out an evaluation from an integrated perspective, including costs and effects falling on both HC and CJ sectors. More specifically, Sub-section 6.3.1 aims to show how to operationalise the theoretical framework proposed in Chapter 3 and based on the work of Claxton et al. (2007), Sculpher et al. (2014) and Walker et al. (2019). Results of the cross-sectoral analysis are illustrated in Sub-section 6.3.2 and discussed in Sub-section 6.3.3.

6.3.1. Methods for conducting the analysis from an integrated perspective

In brief, the analytical framework proposed in Chapter 3 aims at informing decisions to be made by legitimate decision makers about interventions characterised by multiple objectives and impacts on multiple budget constraints. Cost-effectiveness is assessed from an integrated perspective, and results of the analysis can be used to suggest compensation schemes (i.e. transfers) across decision makers. Further details about the framework can be found in Chapter 3.

The formula to calculate the cross-sectoral Incremental Net Monetary Benefits (INMBs) was proposed in Chapter 3 (see Equations 3.7 and 3.8), and involved DEs and OCs falling on the HC sector, a generic non-HC sector, and private consumption. The original generic formulation can be adapted to the specific context of a cross-sectoral analysis involving the HC and CJ sectors, as shown in Equation 6.2.

$$INMB = V_H * INHB + V_R * INRB \tag{6.2}$$

Cross-sectoral INMB can be obtained by aggregating INHB and INRB when expressed in monetary consumption units. To measure health benefits in monetary consumption units, INHB is multiplied by the social valuation (estimated with WTP) of a QALY (V_H). Similarly, INRB is multiplied for the estimated WTP of a reconviction episode (V_R). As highlighted in Equation 6.3, INHB and INRB are given by the DEs net of the OCs. The resulting cross-sectoral INMB formula is illustrated in Equation 6.3 and can be used to aggregate the results of the CEAs conducted from the HC and CJS perspectives.

$$INMB = V_H \overbrace{\left(\Delta H - \frac{\Delta C_H}{k_H} - \frac{\Delta C_R}{k_{RH}}\right)}^{\text{INHB}} + V_R \overbrace{\left(\Delta R - \frac{\Delta C_R}{k_R} - \frac{\Delta C_H}{k_{HR}}\right)}^{\text{INRB}} \qquad (6.3)$$

A description of the variables included in Equation 6.3 is provided in Table 6.10. Further details about this formulation, the variables involved and the underlying assumptions can be found in Chapter 3.

Variable	Description
ΔH	Direct effect on health (including impacts on direct victims)
ΔR	Direct effect on recidivism
ΔC_H	Costs falling on HC budget (including impacts on direct victims)
ΔC_R	Costs falling on CJS budget
k_H	Marginal productivity of the HC sector (in generating health)
k_R	Marginal productivity of the CJS (in preventing recidivism)
k_{RH}	Marginal productivity of the CJS in generating health
k_{HR}	Marginal productivity of the HC sector in preventing recidivism
V_H	Social valuation (estimated with WTP) of a health outcome (QALY)
V_R	Social valuation (estimated with WTP) of a CJ outcome (recidivism)

Table 6.10 Description of the variables included in Equation 6.2

For the sake of simplicity, compared to the INB formulae used in Chapter 5 (e.g. Equation 5.9), incremental costs of the intervention (ΔC_{int}) and costs in response to crime[38] are aggregated here. Basically, all costs falling on CJS budget are aggregated, and resulting total costs are referred to as ΔC_R. Consequently, ΔC_R now does not include only reconviction costs, but all costs falling on CJS budget[39].

Moreover, for pragmatic reasons related to the lack of data, it was assumed that no impact on recidivism is generated from the HC sector. Put differently, it was assumed that k_{HR} is 'infinite' (i.e. an infinite amount of resources should be invested in the HC sector to reduce recidivism). With regard to the marginal productivity of the CJS in producing health (k_{RH}), an estimate was already implicitly calculated in Sub-section 6.2.3.3. In brief, k_{RH} was estimated to be approximately equal to 870,000 £/QALY.

[38] Given by the product of avoided reconvictions (ΔR) and cost savings per reconviction avoided ($-C_R$).
[39] Probably it would have been reasonable to label them C_{CJS} to distinguish them from the costs falling on HC budget (C_H). Nevertheless, using three letters in the subscript (CJS) might lead to confusion when dealing also with cross-sectoral marginal productivities (e.g. between HC and private consumption, k_{HC}). For this reason it has been decided to identify these costs only with the letter *R*, to also highlight their link to reconviction and recidivism (opposed to H for health).

As discussed previously in Chapter 5, Sub-section 5.4.2, it can be assumed that the socio democratic process allocates resources to each sector so that a pound spent in any sector at the margin generates the same benefit as a pound spent in another, based on individual WTP. Consequently, the ratio of *v* and *k* is constant across sectors and it is approximately equal to 4.6, being $V_H \approx$ £60,000 (Glover & Henderson, 2010) and $k_H \approx$ £13,000 (Claxton, Martin, et al., 2015). Equation 6.3 thus reduces to the formula described in Equation 6.4, where the two components falling on HC and CJ sectors are highlighted in red.

$$INMB = 4.6\overbrace{\left(k_H \Delta H - \Delta C_H - k_H \frac{\Delta C_R}{k_{RH}}\right)}^{\text{HC sector}} + 4.6\overbrace{(k_R \Delta R - \Delta C_R)}^{\text{CJ sector}} \qquad (6.4)$$

Looking at the impact on the HC sector, as pointed out in Table 6.10, ΔH includes the DEs falling on both offenders (ΔH_O) and direct victims (ΔH_V). Similarly, ΔC_H includes the HC costs to treat both offenders (ΔC_{Ho}) and direct victims (ΔC_{Hv}). The ratio of ΔC_R over k_{RH} describes the health consequences on unknown victims generated by the OCs in the CJS, namely: health effects on unknown victims (ΔH_{vu}) and additional HC costs to treat them (ΔC_{Hvu}). Equation 6.4 can be thus re-arranged as shown in Equation 6.5.

$$INMB = 4.6\,\overbrace{[k_H(\Delta H_O + \Delta H_V) - (\Delta C_{Ho} + \Delta C_{Hv} + \Delta C_{Hvu} + k_H \Delta H_{Vu})]}^{\text{HC sector}} + 4.6\,\overbrace{(k_R \Delta R - \Delta C_R)}^{\text{CJ sector}} \qquad (6.5)$$

From an integrated perspective, if the INMB is positive, then the new intervention should be introduced; if the IMNB is negative it should not. However, this will only be the case if HC and CJS decision makers accept this reflection of net benefits. In reality, HC (CJS) decision maker may only care about health (recidivism), as such only the net health (reconviction) benefit would be important (Walker et al., 2019).

6.3.2. Results of the cross-sectoral analysis of SIPS trial

Table 6.11 summarises the results of the cross-sectoral evaluation of BA and BLC, compared to CIL, based on methods illustrated in Sub-section 6.3.1.

Impacts on HC sector are taken from the revised results of the CEA from the HC perspective (presented in Section 6.2) and take into account also the spill-over effects on victims' health. Impacts on the CJ sector are taken from the CEA conducted in Chapter 5. DEs and OCs are presented both in monetary and in natural units (i.e. QALYs or Reconvictions). Total INMBs (expressed in monetary consumption units) are calculated using Equation 6.5 and are reported in the last column of Table 6.11.

Group affected	Health Care Sector		Criminal Justice System		Total	
	DE	OC	DE	OC	DE	OC
Brief advice (BA)						
Offenders	-130 £ (-0.010 QALYs)		5610 £ (0.51 R)		**25208 £**	
Unknown offenders				-496 £ (-0.05 R)		**-2280 £**
Victims	216 £ (0.017 QALYs)				**992 £**	
Unknown victims		-5 £ (-0.0004 QALYs)				**-22 £**
General population		111 £ (0.009 QALYs)				**510 £**
INMB	**-20 £ (-0.002 QALYs)**		**6106 £ (0.56 R)**		**27992 £**	
Brief Lifestyle Counselling (BLC)						
Offenders	-234 £ (-0.018 QALYs)		7260 £ (0.66 R)		**32320 £**	
Unknown offenders				-631 £ (-0.06 R)		**-2903 £**
Victims	279 £ (0.021 QALYs)				**1284 £**	
Unknown victims		-6 £ (-0.0005 QALYs)				**-28 £**
General population		-566 £ (-0.044 QALYs)				**-2606 £**
INMB	**618 £ (0.048 QALYs)**		**7891 £ (0.72 R)**		**39140 £**	

Table 6.11 Cross-sectoral INMBs of BA and BLC compared to CIL

From an integrated and broader perspective, both BA and BLC are associated with positive INMBs when compared to CIL. On average BLC provides the highest INMBs. Consequently, the recommended decision would be to implement BLC.

Looking at the impacts on HC and CJ sectors, the implementation of BLC would not require any compensation across sectors. This is because, after modelling the full impact of the programme, BLC is associated with positive INBs from both the HC and CJS perspective.

On the contrary, it would be more problematic for HC and CJS decision makers to find a consensus on whether BA should be introduced or not. In fact, even though BA appears to be cost-effective from an integrated perspective compared to CIL (i.e. cross-sectoral INMBs associated with BA are positive), on average BA is associated with negative INHBs. To fund the intervention it would be necessary to ensure that a compensation across sectors takes place. To put it differently, the CJS decision maker should renounce to a share of the benefits and compensate the negative impact of the intervention on HC.

6.3.3. Discussion

6.3.3.1. Uncertainty around the results of the CEA

Results presented in this chapter are based on mean estimates and should be therefore interpreted with caution. To assess the robustness of the results, sensitivity analyses reflecting the uncertainty about estimates should be conducted. However, the main objective of this analysis was to illustrate how to operationalise the theoretical cross-sectoral framework proposed in Chapter 3. The investigation of uncertainty around results is left as a suggestion for potential further research.

6.3.3.2. Investigating potential double counting of health impacts

Results of the cross-sectoral CEA of SIPS presented in Sub-section 6.3.2 showed that both BA and BLC provided positive INMBs compared to CIL. As highlighted in Chapter 3, Sub-section 3.2.3, cross-sectoral INMBs are based on the results of the CEAs conducted from the HC and CJS perspectives, under the assumption that outcomes are not overlapping or correlated, and the aggregation across dimensions is simply additive.

Nevertheless, as pointed out in Sub-section 5.4.2.3, it might be argued that the value associated with avoided recidivism already includes the impacts on victims' health. In other words, consequences on victims' health might be indirectly included in the WTP to reduce crime, which was instead assumed to be a reflection of the intrinsic value of reducing crime *per se*. If the WTP to reduce crime already includes the health benefits of avoiding being a victim, it would not be appropriate to aggregate the results by summing impacts across dimensions.

A pragmatic way to ensure potential double counting is avoided is to exclude the adjustment for health impacts on victims. For this reason, an alternative cross-sectoral analysis was conducted, which did not consider the adjustment for the health impacts on victims of crime. This alternative cross-sectoral analysis was based on estimates of DEs and OCs of the 'traditional' CEA from the HC perspective (see Chapter 4), and the CEA from CJS perspective (see Chapter 5). DEs and OCs are presented both in monetary and in natural units (i.e. QALYs or Reconvictions). INMBs for BA and BLC compared to CIL are reported in Table 6.12.

Group affected	Health Care Sector		Criminal Justice System		Total	
	DE	OC	DE	OC	DE	OC
Brief advice (BA)						
Offenders	-130 £ (-0.010 QALYs)		5610 £ (0.51 R)		**25208 £**	
Unknown offenders				-496 £ (-0.05 R)		**-2280 £**
General population		374 £ (0.029 QALYs)				**1719 £**
INMB	**-504 £ (-0.039 QALYs)**		**6106 £ (0.56 R)**		**25770 £**	
Brief Lifestyle Counselling (BLC)						
Offenders	-234 £ (-0.018 QALYs)		7260 £ (0.66 R)		**32320 £**	
Unknown offenders				-631 £ (-0.06 R)		**-2903 £**
General population		-227 £ (-0.017 QALYs)				**-1042 £**
INMB	**-7 £ (-0.001 QALYs)**		**7891 £ (0.72 R)**		**36265 £**	

Table 6.12 Revised cross-sectoral INBs of BLC compared to CIL

Total INMBs associated with BA and BLC are both positive, and the highest INMBs are associated with BLC. Therefore, from an integrated perspective, the recommended decision would not change. However, according to this alternative analysis, both interventions would require the CJS decision maker to compensate the HC decision maker. This is because both BA and BLC are not cost-effective from the HC perspective (i.e. they are associated with negative INHBs), whereas they are both associated with positive INRBs.

This analysis aimed to briefly highlight that cross-sectoral evaluation might become more problematic when outcomes are overlapping or correlated. This issue is further discussed in Chapter 8, Section 8.3.

6.3.3.3. Potential inclusion of impacts on private consumption

Time and resource constraints often prevent a comprehensive collection of all data required to expand the evaluation perspective. It has been argued that attempts should be made to include as many relevant costs and outcomes as possible within the resources available, prioritising those that are likely to have the greatest impact on the outcome of the evaluation (Hill et al., 2017). For this reason, in this analysis it was decided to focus the attention on impacts on health and criminal justice. A potential further extension of the cross-sectoral analysis could also include impacts on private consumption.

In the analysis carried out in Sub-section 6.3.3, impacts on private consumption were ignored. In other words, impacts on productivity, employment and properties stolen or damaged were not considered. Nevertheless, as pointed out in Chapter 3, Section 3.2.3.2, when INHB and INNB are aggregated using individual WTP values, Incremental Net Consumption Benefits (INCB) can be aggregated as well. To conduct a cross-sectoral analysis involving HC, CJ and private consumption, the formula used to aggregate the results of the CEAs conducted from the HC and CJS perspectives (see Equation 6.3) can be expanded as shown in Equation 6.6, where: ΔC_c is the DE on private consumption; k_{RC} is the marginal productivity of the CJS in generating private consumption; k_{HC} is the marginal productivity of the HC sector in generating private consumption.

$$INMB = V_H\left(\Delta H - \frac{\Delta C_H}{k_H} - \frac{\Delta C_R}{k_{RH}}\right) + V_R\left(\Delta R - \frac{\Delta C_R}{k_R} - \frac{\Delta C_H}{k_{HR}}\right) + \left(\Delta C_c - \frac{\Delta C_R}{k_{RC}} - \frac{\Delta C_H}{k_{HC}}\right) \quad (6.6)$$

The same procedure used to estimate HC impacts on victims can be employed to compute the additional impacts on individual consumption introduced in Equation 6.6. Lost outputs and stolen, damaged, destroyed and recovered properties associated with each criminal event are available in the literature (Dubourg, Hamed & Thorns, 2005). Numbers of criminal events avoided by the intervention can be computed, and can be linked to the appropriate impacts on private consumption. Lastly, it would be necessary to estimate the spill-over effects on consumption generated in the HC and CJ sector (see Sub-section 8.5.2.2).

6.3.3.4. Incorporation of health equity concerns and cross-temporal extension

In the theoretical framework proposed in Chapter 3, two additional stages of the evaluation were illustrated from a theoretical point of view: the incorporation of health equity concerns, and a cross-temporal extension.

The next chapter will cover the incorporation of equity concerns on health inequalities. As highlighted in Sub-section 6.2.1, this analysis was grounded on the principle that "a QALY is a QALY". Nevertheless, the evaluation can be underpinned by alternative value judgments, and alternative aggregation functions of health impacts can be used.

The operationalisation of the cross-temporal extension of the analysis (i.e. extrapolating effects and costs, and assessing the impacts on hard and soft budgets) will be discussed in the last chapter.

Chapter 7: Incorporating health equity concerns in the cross-sectoral CEA

7.1. Introduction

In Chapter 2, the methodological and technical issues of evaluating public health interventions were described. Similar considerations apply to the evaluation of interventions which specifically address alcohol issues. Three main challenges have been highlighted in the literature, namely: the consideration of wider societal perspectives, the incorporation of impacts on health inequity and the inclusion of long-term outcomes (Hill et al., 2017).

In Chapter 6, the 'wider perspectives' issue was addressed and a cross-sectoral analysis was carried out. The cross-sectoral analysis emphasised the importance of considering not only health and HC consequences of alcohol interventions, but also their impact on CJ. Focusing the analysis only on health costs and consequences potentially underestimates the full impact of the intervention. However, the analysis conducted in Chapter 6 overlooked that health impacts of alcohol interventions might fall on different groups of individuals (e.g. offenders, victims, individuals from various socio-economic groups), and this might raise potential equity issues.

As mentioned in Chapter 2, Sub-section 2.4.2, public health interventions often aim to tackle unfair inequalities in health. For this reason, the inclusion of equity concerns related to the distribution of health impacts between population sub-groups might become relevant (Weatherly, Cookson & Drummond, 2014, Weatherly et al., 2009, Chalkidou et al., 2008). With regard to the specific context of alcohol interventions, it has been argued that alcohol misuse has a wide impact across a range of issues related to (potentially unfair) health inequalities. For this reason, addressing alcohol issues is frequently linked to the attempt to reduce health inequalities in general (Smith & Foster, 2014).

7.1.1. Health equity concerns and alcohol interventions

Despite health equity being a recognised area of need, in the economic evaluations of alcohol prevention interventions there is little evidence of attempts to address this issue (Burton et al., 2017, Hill et al., 2017). For example, none of the studies identified in the systematic review conducted by Hill and et al. (2017) specifically addressed or even discussed equity in health. Similarly, the impact on health inequalities was frequently not identified in the effectiveness and cost-effectiveness studies of alcohol control policies in England reviewed by Burton et al. (2017). Nevertheless, it was reported that some interventions (e.g. regulating marketing and availability, providing information and education, managing the drinking environment) could be used to address health inequalities, if implemented in areas with greater deprivation or directed to specific groups (Burton et al., 2017).

Some evidence was identified on price regulation and taxation, which was found to provide greater health benefits to heavy drinkers who experience the greatest harm (Burton et al., 2017). Using the Sheffield Alcohol Policy Model[40] (SAPM), Holmes et al. (2014) investigated the equity implications of a minimum unit price (MUP) and how alcohol-related mortality and morbidity can vary by socioeconomic status (SES). Health benefits from the policy were found to be unequally distributed, with individuals in the lowest socioeconomic group benefitting more (Holmes et al., 2014). Similarly, using a newer version of the SAPM, Meier et al. (2016) found that different alcohol taxation and price control systems could contribute substantially to the reduction of health disparities between socioeconomic groups. Alcohol-content-based taxation and MUP policies can in fact reduce the harms associated with alcohol consumption, which are more likely to be suffered by less affluent groups (Burton et al., 2017, Meier et al., 2016).

With regard to Identification and Brief Advice (IBA) policies, the lowest socioeconomic groups were again estimated to experience the greatest absolute reduction in harms (Burton et al., 2017). Nevertheless, the impact of IBA policies is likely to be substantially lower than MUP policies in terms of reducing absolute socioeconomic inequalities in health, unless targeted at lower socioeconomic groups (Angus, Colin et al., 2015). For example, if implemented in the CJS context, IBA interventions can reduce alcohol consumption and harm in offenders and contribute to reducing health inequalities (Burton et al., 2017).

To summarise, alcohol-focused interventions have been rarely evaluated in a manner that enables an analysis of differential health impacts by population subgroups. More research is needed in order to explore how alcohol interventions can impact different social groups' health (Smith & Foster, 2014).

7.1.2. Objective and methods of this analysis

As summarised in Chapter 2, Sub-section 2.4.3, various analytical techniques could theoretically accommodate the inclusion of health equity concerns in the economic evaluation of a public health

[40] Also known the Sheffield model, it assesses effectiveness and cost-effectiveness of selected alcohol policies.

intervention. Similarly, the analytical framework proposed in Chapter 3 can be extended in order to assess also impacts on health inequities, based on alternative value judgments (see Section 3.3).

In Chapter 3, the 'health inequality assessment' extension of the framework was illustrated from a theoretical point of view. The aim of this chapter is to use the SIPS case study to show how to operationalise the inclusion of health equity concerns in the evaluation of a public health intervention with effects on HC and CJ.

7.1.2.1. Scheme of the analysis

In the previous chapters, the analytical framework was used to address issues related to the perspective of the evaluation. SIPS trial was selected as a vehicle to demonstrate the methods proposed in Chapter 3, and a cross-sectoral analysis of SIPS was carried out in Chapters 4, 5 and 6. The cross-sectoral analysis of SIPS consisted of four steps. A brief recap of the methods and results of each step is offered below.

The objective of this chapter is to revise the third and the fourth step of the SIPS case study, in order to incorporate equity concerns in the evaluation of the intervention.

Step 1

In Chapter 4, a 'traditional' CEA from the HC perspective was conducted. This perspective considered health-related costs and health effects only on the offenders. From this first analysis, it was concluded that from the HC prospective BA and BLC were probably not cost-effective compared to CIL.

Step 2

In Chapter 5, a CEA from the CJS perspective was conducted, using avoided reconvictions as a measure of effectiveness. From this second analysis, it was concluded that from the CJS perspective it was probably cost-effective to invest in BLC, when compared to CIL.

Step 3

After modelling the costs and consequences from the CJS perspective, in the first part of Chapter 6 the narrow CEA from the HC perspective was expanded to incorporate victims' health costs and effects.

In the 'traditional' CEA from the HC provider perspective (carried out in Chapter 4) the focus was on costs falling on HC budget and on QALY impacts for trial participants. This approach fails to reflect the potential health impacts on victims through the impacts on the CJS. In reality, reductions in criminal activities engender avoided QALY losses for the victims. Moreover, additional HC costs savings are generated.

From a 'narrow' HC perspective, the true impacts on health and HC budget can be described by the two dashed rectangles A and B in Figure 7.1. These include not only the impacts on offenders' health and the associated HC costs (the blue rectangles), but also the additional impacts on victims' health and the additional costs falling on HC budget to treat the victims of crime (the green rectangles).

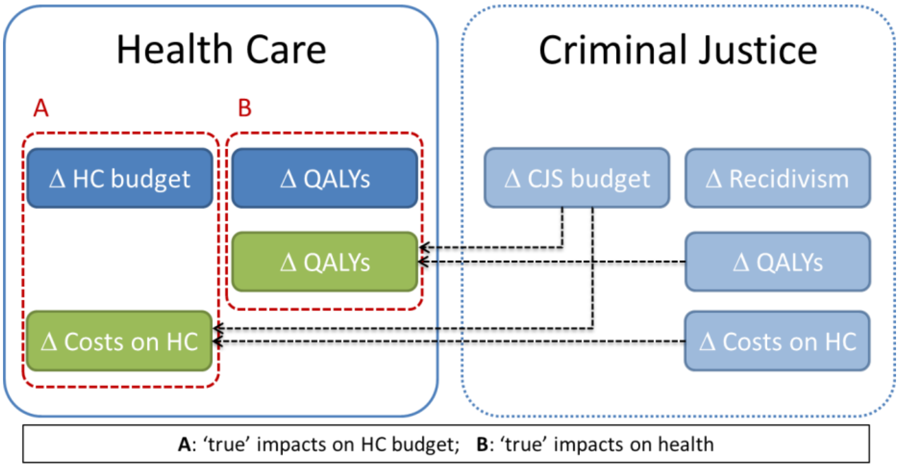


Figure 7.1 Scheme of the revised CEA from the HC perspective (Step 3)

When these additional components were included in the CEA, CIL provided on average positive Incremental Net Health Benefits (INHBs). The conclusion of the revised CEAs conducted from the HC perspective was to choose BLC, as that intervention provided the greatest INHBs.

Step 3 revised

The true impacts of SIPS on health fall on different groups of individuals, namely offenders, victims and general population. Potential equity issues might therefore arise when these health impacts are aggregated.

In the analysis previously conducted (Step 3), differential QALYs for offenders and victims were simply summed. In other words, the aggregation within the health dimension was simply based on the sum the unweighted changes in health across individuals. It was thus implicitly assumed that distribution of health effects was not relevant. However, the aggregation within dimension (see Sub-section 3.2.2.1) could potentially take into account the distribution of impacts across population subgroups, and alternative functional forms representing alternative normative judgements are possible (Walker et al., 2019).

In this chapter, the third step of the original analysis is revised. Using the methods illustrated in Chapter 3, Sub-Section 3.3.3, alternative solutions about how to aggregate health impacts which fall on different population subgroups are investigated.

Step 4

In the second part of Chapter 6, an analysis from an integrated perspective that considers both the HC and CJ sectors was conducted. The aim was to inform decisions to be taken by decision makers.

The analysis was based on the framework formulated by Claxton et al. (2007), and further developed by Sculpher et al. (2014) and Walker et al. (2019). Multiple objectives and impacts on multiple budget constraints were included in the evaluation Results of the analysis were used to suggest compensation schemes across decision makers. Further details about the framework can be found in Chapter 3.

Step 4 revised

The cross-sectoral analysis conducted in Chapter 6 (here referred to as 'Step 4') was based on the results of the original Step 3 of the analysis, and was therefore grounded on the value judgment that "a QALY is a QALY". Nevertheless, if the CEA from the HC perspective is underpinned by alternative value judgements, and alternative aggregation functions are used (Step 3 revised), the cross-sectoral analysis must be revised as well.

In this chapter, methods and results of the fourth step of the analysis are revised. The final objective of this chapter is to show how the proposed framework can incorporate health equity concerns in the cross-sectoral analysis, and inform decisions to be taken from an integrated and broader perspective by legitimate decision makers.

7.1.2.2. Health impacts on different population subgroups affected by SIPS

As pointed out in Chapter 3, Sub-section 3.3.2, CEAs are typically underpinned by the assumption that "a QALY is a QALY", irrespectively of who is the beneficiary (Cookson et al., 2017, Drummond, M. et al., 2015). If, on the contrary, alternative value judgements are incorporated in the evaluation, it becomes necessary to specify characteristics of the individuals affected by the interventions and assess the distribution of health impacts across the population.

Health impacts of SIPS fall on three main groups of individuals: offenders treated with the intervention, (direct and unknown) victims of crimes and general population bearing the opportunity costs[41].

Offenders treated with the intervention

The intervention had the aim of reducing alcohol misuse and was aimed at individuals on probation. Therefore, direct effects of SIPS fell on offenders (on probation) with a history of alcohol misuse. Such a population subgroup might be deemed socially marginalised and is characterised on average by a lower SES when compared to the general population (Bramley et al., 2015). Because alcohol related harms follow a social gradient with the most alcohol related harms experienced by deprived socioeconomic groups (Smith & Foster, 2014), the incorporation of health equity concerns might affect substantially the results of the analysis.

Victims

The group of the victims is in turn composed of two subgroups: direct victims of crimes committed by offenders treated with SIPS, and unknown victims who bear the consequences of the OCs generated in the CJS. In fact, when CJS displaces resources to accommodate the costs of a new intervention, unknown offenders are likely to commit crimes (and hurt unknown victims) as the resources for their programmes are displaced.

General population

OCs of HC fall on the general population. These costs include also the additional HC costs to treat victims of crimes. In fact, this additional component falls on unknown patients that cannot be treated due to the resources displaced by the HC system to accommodate the additional HC treatments for the victims of crime. Assuming that resources displaced come from the generic HC budget, consequences are distributed across all population subgroups.

7.1.2.3. Two potential value judgments about health inequalities

As mentioned in Chapter 3, Sub-section 3.3.2, the perceived unfairness of health inequalities can be based on a broad range of other factors. In this chapter, two alternative value judgments are explored, and two alternative solutions to aggregate health impacts falling on different population

[41] The same categorisation can be used to describe health impacts of any other alcohol interventions targeted to offenders.

subgroups are proposed. Both analyses are firstly conducted from the 'narrow' HC perspective[42] (Step 3 revised). Secondly, a cross-sectoral analysis is carried out (Step 4 revised).

First scenario

The first analysis is based on the hypothesis that impacts on victims' health are valued differently from generic health variations. Different weights are therefore assigned to variations in health if these are caused by criminal activity. To incorporate this alternative value judgment in the evaluation, an equity trade-off analysis based on equity weighting is conducted. Methods were introduced in Chapter 3, Sub-section 3.3.3.1, and their operationalisation is illustrated in detail in Section 7.2.

Second scenario

The second analysis is based on the hypothesis that decision makers might want to prioritise interventions that benefit disadvantaged and least healthy groups. The impact of this alternative value judgment is investigated in Section 7.3, where costs and outcomes of SIPS are disaggregated by equity-relevant subgroups, and an equity impact analysis based on a DCEA is conducted.

7.2. Equity trade-off analysis

The equity trade-off analysis was introduced in Chapter 3, Sub-Section 3.3.3. Briefly, in the equity trade-off analysis, 'equity weights' can be used to value health benefits that apply to people with characteristics that reflect considerations of fairness or that are disadvantaged because of factors beyond their control (Cookson et al., 2017). For example, on the basis of vertical equity considerations, the general public might associate a different value (i.e. different weights) to health impacts due to someone else's responsibility, such as medical negligence or malpractice (Brazier et al., 2007). Similarly, a different equity weight might be put on health impacts if caused by being a victim of crime.

In this section the equity weighting method proposed by Round and Paulden (2018) and illustrated in Chapter 3, Sub-section 3.3.3.1, is used to explore the effects of this alternative value judgment.

[42] The analysis from the HC perspective includes the spill-over effects from CJS on victims' health and is therefore still based on cross-sectoral data.

In Sub-section 7.2.1 the methods are described. Results of the analysis from the HC perspective and from an integrated perspective are reported in Sub-sections 7.2.2 and 7.2.3, respectively, and are discussed in Sub-section 7.2.4.

7.2.1. Equity weighting and victims of crime

In Chapter 3, the generic formula to estimate the Incremental Net weighted Health Benefit (INwHB) of an intervention was proposed (see Equation 3.11). The generic formula can be adapted to reflect the specific value judgment investigated in this analysis, as shown in Equation 7.1.

$$INwHB = \omega * (Q_{gv} - Q_{dv}) + (Q_{gp} - Q_{dp}) > 0 \tag{7.1}$$

Where: Q_{gv} and Q_{dv} are the QALYs gained and displaced by victims of violent crimes, respectively; Q_{gp} and Q_{dp} are the QALYs gained and displaced by other patients, respectively; ω represents the weight applied to QALYs gained or lost by victims of violent crimes. Because the weight is applied only to victims, in this formulation it is assumed that health impacts falling on offenders and general population have the same weight.

Equation 7.1 can be linked to the Equation 6.1 from Chapter 6 (reported again below), which describes the full impact of the intervention on HC, including the health spill-over effects on the victims of crime.

$$INHB = (\Delta H_O + \Delta H_V) - \left(\frac{\Delta C_H + \Delta C_{Hv} + \Delta C_{Hvu}}{k_H} + \Delta H_{Vu}\right) \tag{6.1}$$

Similarly, INwHBs can be expressed in terms of health impacts on offenders' health, health impacts on (direct and unknown) victims and the costs falling on HC budget.

Resulting DEs (i.e. QALYs gained by victims of violent crimes and other individuals) of the intervention can be then computed as shown in Equation 7.2. As previously stressed, the equity weight is applied only to victims of crime, therefore the only distinction made is between 'victims' and 'other individuals' (which include also the offenders). Consequently, QALYs gained by victims of violent crimes (Q_{gv}) are the variations in direct victims' health (ΔH_V), while QALYs gained by other individuals (Q_{gp}) are the health impact on offenders' health (ΔH_O).

$$Q_{gv} = \Delta H_V \qquad\qquad Q_{gp} = \Delta H_O \tag{7.2}$$

To compute the OCs (i.e. QALYs displaced in patients who are victims of violent crimes and other patients) of the intervention it is necessary to disentangle the HC resources used to treat generic patients from those used for victims of criminal offences. In other words, how the total costs (which fall on the HC budget) are generated[43] is not relevant, but it is crucial to understand where the resources are displaced. In fact, total displaced resources due to costs falling on HC budget affect the magnitude of resources that would have been alternatively used to treat both other general unknown patients and other unknown victims of violence. Total HC costs are thus disaggregated into variations in the resources available for general patients (ΔC_P) and for victims of crimes (ΔC_V). These components are described more in detail in the next sub-section.

OCs can be therefore calculated as shown in Equation 7.3. QALYs displaced in victims of violent crimes (Q_{dv}) consist of the variations in costs falling on the HC and otherwise used to treat victims of crime (ΔC_V), and the health impacts falling on unknown victims' health (ΔH_{Vu}). This latter component is generated by the OCs in the CJS. QALYs displaced in other patients (Q_{dp}) are the costs falling on the HC and otherwise used to treat general patients (ΔC_P). Both ΔC_V and ΔC_P are 'converted' in QALYs using the marginal productivity of the HC sector (k_H).

$$Q_{dv} = \frac{\Delta C_V}{k_H} + \Delta H_{Vu} \qquad\qquad Q_{dp} = \frac{\Delta C_P}{k_H} \qquad (7.3)$$

To obtain the INwHB associated with the intervention, equity weighting can be applied to DEs and OCs as illustrated in Equation 7.1. Resulting INwHB can be computed as shown in Equation 7.4.

$$INwHB = \left(\Delta H_O - \frac{\Delta C_P}{k_H}\right) + \omega\left(\Delta H_V - \frac{\Delta C_V}{k_H} - \Delta H_{Vu}\right) \qquad (7.4)$$

Two operations are required before being able to operationalise the INwHB formula illustrated in Equation 7.4. Firstly, in order to compute the amount of QALYs displaced in victims of violent crimes and in the general population, the proportions of victims and general population treated by the NHS must be estimated. This issue is discussed in in Sub-section 7.2.1.1. Secondly, it is necessary to find an appropriate weight to be applied to impacts on the health of victims of violent crimes. This is discussed in Sub-section 7.2.1.2.

7.2.1.1. Health care opportunity costs falling on victims of crime

As stressed in Sub-section 3.3.3.1, in order not to violate the principle of horizontal equity, equity weights must also be taken into account when assessing the OCs. The amount of victims' QALYs displaced must be therefore estimated and weighted accordingly. For this reason, it is necessary to determine the proportion of victims in the general population treated by HC services.

[43] These are given by the sum of: costs to treat offenders (ΔC_H) additional costs due to the HC services for victims (ΔC_{Hv}) and unknown victims (ΔC_{Hvu}) (see Chapter 6, Equation 6.1).

Data about finished admission episodes (FAEs)[44] and accident and emergency (A&E) attendances can give an idea of the overall activity of the National Health Service (NHS). A report published by the NHS Digital (2012) provided details on Hospital Episode Statistics (HES) data about admissions and attendances due to assault in the period from April 2011 to March 2012 in England. These are the only data publicly available containing information about admissions due to assaults. It was therefore assumed that proportions observed during this specific interval of time can be generalised to different years. According to the report, 15.0 million FAEs were recorded in that year, and assaults accounted for 38,766 FAEs. Moreover, with regard to the A&E HES, 17.6 million A&E attendances were recorded, and there were 185,941 A&E attendances as a result of assault.

Only data about assault were available. These numbers certainly under-estimate the total FAEs and A&E attendances due to generic violent crimes. In fact, according to the Home Office Statistical Bulletin (Chaplin, Flately & Smith, 2011), 67% of the crimes with consequences on health[45] committed in 2010 were assaults, while 33% were other offences with consequences on health. Therefore, the 38,766 FAEs and 185,941 A&E attendances due to assaults represent approximately 67% of the total criminal offences with consequences on health. A crude estimate of the total number of FAEs and A&E events due to violent criminal offences can be obtained by multiplying the FAEs and A&E events due to assault for a factor of 1.5. Results are summarised in Table 7.1.

Category	**No. of events**	**% of total events**
Finished admission episodes (FAEs)	15,000,000	
FAEs due to assaults	38,766	0.26%
FAEs due to violent criminal offences	57,840	0.39%
A&E attendances	17,600,000	
A&E attendances as a result of assault	185,941	1.06%
A&E attendances due to violent criminal offences	277,429	1.58%

Table 7.1 FAEs and A&E attendances due to crime

There is a higher percentage (1.58%) of A&E attendances due to violent criminal offences compared to FAEs (0.39%). Such a discrepancy might be due to the types of injuries. It is probably more likely for a victim to end up in A&E rather than being admitted. Assuming that these proportions are constant in time, and keeping in mind that this work is mostly illustrative, it can be concluded that

[44] A finished admission episode is the first period of inpatient care under one consultant within one health care provider (NHS Digital, 2012).

[45] According to the report, 12.2% of all crimes committed in 2010 had consequences on health, and 87.8% were other offences without (estimated) consequences on health, such as criminal damage, violence without injury and theft. It is acknowledged that even these crimes might have actually health consequences on the victims. This is a simplifying assumption, due also to the lack of data.

on average approximately 1% of the activity of the NHS is devoted to treating victims of criminal offences.

When introducing equity weights, health impacts on direct and unknown victims are weighted differently. The same equity weight applied to health outcomes for DEs can be now used also to adjust the OCs. In fact, 1% of the OCs of an intervention displacing HC resources are estimated to fall on victims of crime. The equity weight must be therefore applied. The remaining part (99%) of the OCs falls on generic patients and no equity weight is applied.

7.2.1.2. Equity weight for impacts on victims' health

As mentioned at the end of Sub-section 7.2.1, to implement the equity weighting summarised in Equation 7.4 it is necessary to find an appropriate weight (ω) to be applied to impacts on the health of victims of violent crimes.

Such a parameter is not available in the literature. Given the exploratory purpose of this analysis, it has been decided to use a hypothetical weight of 2. A threshold analysis will be conducted to investigate weight values that would be required to change the recommended decisions.

Incorporating an equity weight of 2 in the analysis implies that QALYs gained or lost by victims of violent crimes are valued twice as important as health impacts falling on other individuals. It must be stressed again that in this analysis only consequences falling on victims were valued differently. All other health impacts were treated equally. For instance, consequences on offenders treated with the intervention were compared to any other health impacts falling on the general population. This limitation is further discussed at the end of this section, before introducing the equity impact analysis.

7.2.2. CEA of SIPS from the HC perspective including equity weights

As explained in Sub-section 7.1.2, the first objective of this analysis is to revise the results of the original CEA of SIPS conducted from the HC perspective without equity weighting (i.e. Step 3). The second objective, illustrated in the next sub-section, is to revise the cross sectoral evaluation of the intervention (i.e. Step 4).

According to the 'traditional' CEA (conducted in Chapter 3), both interventions are associated with negative INHBs when compared to CIL. On the contrary, when the additional components (i.e. QALY losses avoided by victims and additional HC costs savings) are included in the CEA (Chapter 6), while INHBs of BA remain negative, BLC appears to be cost-effective from the HC perspective.

Equity weights can be incorporated in the CEA from the HC perspective using the methods illustrated in Sub-section 7.2.1. Impacts of BA and BLC are compared to CIL and have been disentangled into 'traditional' estimates and additional components. The 'traditional' estimate is the result of the CEA without considering the impacts on victims' health. Results are presented in Table 7.2, where costs are expressed in QALYs using the marginal productivity of the HC sector (k_H) to convert monetary impacts in QALY variations. Cost savings are expressed as negative OCs and INwHBs are obtained as the difference between DEs and OCs.

Group affected	'Traditional' estimate		Additional component		Total	
	DE	**OC**	**DE**	**OC**	**DE**	**OC**
Brief advice (BA) vs CIL						
Offenders	-0.010				**-0.010**	
Victims			0.033		**0.033**	
Unknown victims				-0.001		**-0.001**
General population		0.029		-0.020		**0.009**
INwHB	**-0.039 QALYs**		**0.054 QALYs**		**0.015 QALYs**	
Brief Lifestyle Counselling (BLC) vs CIL						
Offenders	-0.018				**-0.018**	
Victims			0.043		**0.043**	
Unknown victims				-0.001		**-0.001**
General population		-0.018		-0.026		**-0.044**
INwHB	**0.000 QALYs**		**0.070 QALYs**		**0.070 QALYs**	

Table 7.2 CEA of SIPS from HC perspective including equity weights (Step 3 revised)

Compared to original estimates without equity weighting (reported in Chapter 6, Table 6.7), the only difference in the 'traditional' analysis is that now OCs take into consideration the impact of displacing resources for victims. In fact, even if health impacts on (direct and unknown) victims are not included, 1% of the health care OCs still fall on other unidentifiable victims. Results of the 'traditional' CEA are therefore slightly different. INHBs associated with BLC previously estimated were negative (-0.001 QALYs). This is because BLC was associated with negative health consequences, but also cost savings. Nevertheless, cost savings could not offset the negative impact on offenders' health. On the contrary, in this analysis, INHBs associated with BLC are zero. This is because the negative OCs associated with BLC become more valuable due to the higher value associated with additional resources available for victims of violence. Cost savings offset the negative impact on offenders' health.

When the additional component is incorporated in the analysis, DEs on (both unknown and direct) victims and 1% of the additional OCs are weighted using ω. Overall INwHBs are higher when compared to the original analysis. BA and BLC are now associated with INwHBs of 0.015 QALYs (previous estimate: -0.002 QALYs) and 0.070 QALYs (previous estimate: 0.048 QALYs), respectively. According to this analysis, BLC is still the overall cost-effective option, being the intervention that provides the highest INwHBs. However, compared to the results without equity weighting, the main difference is that INwHBs associated with BA become positive (even though BA is still dominated by BLC).

In this analysis, a hypothetical weight of 2 was used to value impacts on victims' health. A threshold analysis was conducted in order to investigate the minimum equity weight necessary to make INwHBs associated with BA positive. It was found that the tipping point for the equity weight to switch the sign of the average INwHBs associated with BA was 1.1. Incremental Net weighted Benefits (INwBs), expressed both in natural and monetary units, corresponding to an equity weight of 1.1 are reported in Appendix P.

7.2.3. Incorporating equity weights in the cross-sectoral analysis of SIPS

In the analysis carried out in Sub-section 7.2.2, the perspective adopted was still the narrow viewpoint of the HC system. Cross-sectoral data were only used to estimate health impacts on victims. The ultimate goal of the proposed framework is to incorporate equity concerns in a full cross-sectoral analysis conducted from an integrated perspective (Step 4 revised). New estimates from the analysis conducted in Sub-section 7.2.2 (which now incorporate also equity concerns) can be used to inform a cross-sectoral analysis from an integrated perspective that includes DEs and OCs falling on both HC and CJ sectors.

As illustrated in Chapter 6, impacts on health and crime can be aggregated using the social valuation (V_H) of a health outcome (i.e. a QALY) and the social valuation (V_R) of a CJ outcome (i.e. an avoided episode of recidivism). Simplifying assumptions such assuming a constant ratio of *v* over *k* across sectors can be employed to compensate the lack of data for sectors other than HC[46]. Health and CJ benefits can be aggregated to obtain the societal Incremental Net Monetary Benefits (INMBs) as shown in Equation 6.2 (see Chapter 6, Sub-section 6.3.1). If equity weights are incorporated in the analysis, the same formulation can be used, but using INwHBs instead of INHBs. Resulting Incremental Net weighted Monetary Benefits (INwMBs) including equity weights can be therefore obtained as illustrated in Equation 7.5.

$$INwMB = V_H * INwHB + V_R * INRB \quad (7.5)$$

[46] Borrowing evidence from the HC sector, being $V_H \approx$ £60,000 (Glover & Henderson, 2010) and $k_H \approx$ £13,000 (Claxton, Martin, et al., 2015), it can be assumed that all *v*'s are about 4.6 times bigger than *k*'s across sectors.

Cross-sectoral INwMBs of BA and BLC compared to CIL are reported in Table 7.3.

Group affected	Impact on health		Impact on reconviction		Total	
	DE	OC	DE	OC	DE	OC
Brief advice (BA)						
Offenders	-130 £ (-0.010 QALYs)		5610 £ (0.51 R)		**25208 £**	
Unknown offenders				-496 £ (-0.05 R)		**-2280 £**
Victims	431 £ (0.033 QALYs)				**1985 £**	
Unknown victims		-9 £ (-0.001 QALYs)				**-44 £**
General population		112 £ (0.009 QALYs)				**515 £**
INwMB	**199 £ (0.015 QALYs)**		**6106 £ (0.56 R)**		**29001 £**	
Brief Lifestyle Counselling (BLC)						
Offenders	-234 £ (-0.018 QALYs)		7260 £ (0.66 R)		**32320 £**	
Unknown offenders				-631 £ (-0.06 R)		**-2903 £**
Victims	558 £ (0.043 QALYs)				**2568 £**	
Unknown victims		-12 £ (-0.001 QALYs)				**-56 £**
General population		-572 £ (-0.044 QALYs)				**-2631 £**
INwMB	**908 £ (0.070 QALYs)**		**7891 £ (0.72 R)**		**40478 £**	

Table 7.3 Cross-sectoral INwMBs of BA and BLC compared to CIL (Step 4 revised)

Results resemble the original cross-sectoral analysis, conducted in Chapter 6. Similarly to the findings presented in Sub-section 6.3.2 and summarised in Table 6.11, from an integrated and broader perspective BLC appears to be the overall cost-effective treatment option. Nevertheless, when equity weights are incorporated in the analysis, the implementation of both BA and BLC would not require any compensation across sectors. This is because BA and BLC are on average cost-effective from both the HC and CJS perspective.

7.2.4. Discussion

Equity-weighting analysis should not be used as an algorithm for making decisions, but as an aid to deliberation, used to assist decision-makers and stakeholders in exploring the implications of alternative value judgements about equity (Cookson et al., 2017). In this section the focus is on victims' health, but alternative equity considerations could be investigated using the same methods. For example, empirical evidence suggested that some inequalities in health might be considered less inequitable than others if they are attributed to individual responsibility. Less priority might be assigned to those who are considered to be in some way responsible for their ill health (Dolan & Tsuchiya, 2009). Consequently, it might be argued that a lower equity weight could be put on offenders' health, to distinguish them from general population. However, evidence is not conclusive. The priority could still be to treat those who have less lifetime health, independently of their responsibility (Edlin, Tsuchiya & Dolan, 2012). Moreover, the ethical implications of introducing such an adjustment would have to be discussed.

The equity weighting approach illustrated in this section has the advantage of being relatively easy to implement. Nonetheless, the risk is that the evaluation becomes context dependent and undermines the comparability of the results across different public health areas. For example, the value judgment explored in this section (i.e. the prioritisation victims' health) might be deemed appropriate for this very specific context, but not relevant in many other cases. In fact, in this analysis, any other potentially disadvantaged group (e.g. due to age, burden of illness etc.) did actually receive a lower weight compared to victims of crime.

The objective of this work is to develop a framework which is flexible, but context independent in its general formulation. For this reason, if equity weighting is incorporated in the evaluation, then it would be necessary to introduce a set of weights to be applied to all categories of individuals with characteristics that reflect considerations of fairness. Incorporating these 'universal principles' in the analysis would allow to implement the same methods to different context and compare the results across different areas of public health. Nevertheless, the operationalisation of such a framework would be much more complex and hindered by data availability.

Alternatively, an overarching equity principle applicable to all evaluations could be based on the potentially inequitable association between SES and health (Love-Koh et al., 2015). The effect of incorporating in the economic evaluation of SIPS a value judgment that aims at prioritising benefits falling on more disadvantaged population subgroups is explored in the next section.

7.3. Equity impact analysis

Introduced in Chapter 3, Sub-section 3.3.3.2, the equity impact analysis quantifies the distribution of costs and effects by equity-relevant variables (Cookson et al., 2017). For example, as illustrated in Chapter 3, Sub-section 3.3.2.1, the positive correlation between SES and health (in terms of QALEs) might be deemed unfair and unjust (Love-Koh et al., 2015, Marmot et al., 2010). An equity impact analysis can be therefore carried out to examine the distribution of DEs and OCs broken down by SES subgroups, and assess the impact of the intervention on heath equity.

In the proposed framework, the DCEA (Asaria, Griffin & Cookson, 2016) was chosen as the analytical technique to conduct the equity impact analysis. Methods to conduct a DCEA were already illustrated in Chapter 3, Sub-section 3.3.3.2. The objective of this section is to carry out a DCEA of SIPS, and show how to incorporate its results in the cross-sectoral evaluation. The analysis plan of the DCEA is described in the next sub-section.

7.3.1. Analysis plan of the DCEA of SIPS

This operationalisation of the DCEA reflects the approach employed by Love-Koh et al. (2018) for the analysis of NICE public health guidelines. Similarly to their analysis, it is assumed that the DEs of the intervention follow the distribution of prevalence of the target population by SES subgroup. In other words, in order to assess the distributional impact on offenders' health, firstly the prevalence of offenders with alcohol issues (i.e. the target population) by SES subgroup is estimated. Secondly, the treatment effect is simply applied to all potential recipients of the intervention. The estimation of the DEs on offenders' health is explained in detail in Sub-section 7.3.2.

In Sub-section 7.3.3, population costs are converted into population health OCs, and their distribution by SES is estimated. In Sub-section 7.3.4, the distributional impact of DEs and OCs falling on direct and unknown victims is estimated. Due to the lack of data, it is assumed that the prevalence of victims of crime by SES is the same as offenders' prevalence.

Once estimates of health DEs and population health OCs by SES are available, the distribution of population INHBs by SES can be calculated as the DEs net of the OCs. INHBs of SIPS by SES are reported at the beginning of Sub-section 7.3.5. Moreover, using the distribution of pre-intervention lifetime health available in the literature (Love-Koh et al., 2015), whether net health impacts of the intervention affect the unequal distribution of health at baseline can be assessed. Variations in inequality indexes and results of the DCEA of SIPS are illustrated in Sub-section 7.3.5.

Finally, in Sub-section 7.3.6, results of the DCEA are incorporated in the cross-sectoral analysis of SIPS.

7.3.2. Distributional impact on offenders

Due to the limited sample size of the SIPS trial and the lack of detailed information about the offenders involved in the study, it was not possible to estimate the heterogeneous effects of the intervention for different SES subgroups. For this reason, in this analysis it was assumed that intervention effects follow the distribution of prevalence of offenders who misuse alcohol.

Prevalence rates were converted into absolute numbers using population data to obtain the number of potential recipients for each SES subgroup (assessed using the income deprivation at local authority level). Subgroup impacts were calculated by multiplying intervention DEs for the number of potential recipients in each SES subgroup.

7.3.2.1. Prevalence of offenders with alcohol issues by Index of Multiple Deprivation (IMD)

Classification based on the Offender Assessment System (OASys)

The National Offender Management Service usually assigns offenders to one of four 'tiers', based on a number of factors such as the risk of reoffending and risk of serious harm. As the tier number rises (Tier 1 is the lowest, and offenders on community sentences belong to Tiers 2–4), there is an increase in risk, the needs of the offender, demands of the sentence and the level of resource needed to manage them (Ministry of Justice, 2013)[47]. Additional information about the prevalence of needs and risks among offender groups is provided by the Offender Assessment System (OASys). OASys is a system for assessing offenders' risks and needs in terms of the likelihood of reconviction, risk of harm to the public, and other eight criminogenic needs (i.e. offending-related factors), including alcohol misuse (Ministry of Justice, 2013)[48].

In 2013, the Ministry of Justice reported that 35% of the offenders with a community or suspended sentence order were identified as having an alcohol misuse need, according to the OASys classification. Moreover, 16% of offenders under post-release supervision were assessed as having issues related to alcohol misuse. However, OASys assessments are not required for all offenders.

---

[47] As reported in the Probation Circular 08/2008 (National Offender Management Service, 2008), the four tiers represent different levels of intervention: Tier 1 = Punish; Tier 2 = Punish and Help; Tier 3 = Punish and Help and Change; Tier 4 = Punish and Help and Change and Control.

[48] The full list includes: accommodation, education, training and employment, relationships, lifestyle and associates, drug misuse, alcohol misuse, and thinking and behaviour and attitudes.

Therefore, OASys data should not be read as representative of the entire offending population (Ministry of Justice, 2013) . These numbers can help to provide a broad idea of the magnitude of the prevalence of alcohol misuse among offenders, but reliable estimates of potential recipients per SES subgroup could not be calculated.

Severe and Multiple Disadvantage (SMD) classification

An alternative option is to use the 'severe and multiple disadvantage' (SMD) classification, developed by the analysts of the Lankelly Chase Foundation (Bramley et al., 2015). SMD index identifies adults involved in the homelessness, substance misuse and CJS in England. To obtain such classification, various datasets from offender, substance misuse and homeless services were employed, and information was integrated with additional administrative datasets. The population was divided into three categories: SMD1, experiencing one disadvantage domain only (i.e. 'homelessness only', 'offending only', or 'substance misuse only'); SMD2, experiencing two out of three disadvantage domains (i.e. 'homelessness & offending'; 'substance misuse & offending'; 'substance misuse & homelessness'); SMD3, experiencing all three disadvantage domains.

Bramley et al. (2015) reported 58,000 cases of SMD3, with a prevalence rate of 0.15%, calculated as percentage of the adult working age (16-64) population of England. The association between SMD3 prevalence rates and IMD was also investigated. Figure 7.2 is an adaptation of the original picture from the report, and shows that a higher relative prevalence of SMD3 is associated with lower IMD quantiles. Approximately one third of the total cases of SDM3 are observed in the subgroup with the lowest IMD.

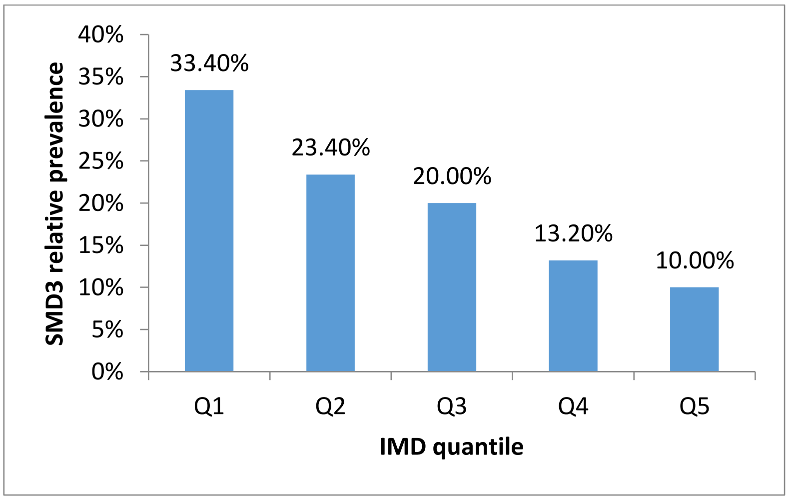


Figure 7.2 Relative prevalence of SMD3, adapted from Bramley et al. (2015)

Specifically, for SMD2, 164,000 cases were reported, with a prevalence rate of 0.42%. Of these, 99,000 cases were associated with the SMD2 subgroup 'substance misuse & offending'. The prevalence rate of this subgroup (0.25%) was not available in the original report and was estimated from the number of observations. Assuming that SMD2 (and in particular to the subgroup 'substance misuse & offending') has the same distribution across IMD quantiles as SMD3, resulting prevalence (i.e. percentage of the total 16-64 English population) is illustrated in Table 7.4.

IMD quantile	SMD2 (substances & offending) relative prevalence	SMD2 (substances & offending) prevalence in the 16-64 population
Q1 – poorest	33.40%	0.08%
Q2	23.40%	0.06%
Q3	20.00%	0.05%
Q4	13.20%	0.03%
Q5 – richest	10.00%	0.03%
Total	100.00%	0.25%

Table 7.4 Prevalence of SMD2 subgroup 'substance misuse & offending'

In the SMD classification, 'substance misuse' was defined as *"participating in publicly funded treatment for dependence on drugs or alcohol"* (Bramley et al., 2015). For this reason, on the one hand the prevalence rate of 0.25% might overestimate the population that could be treated with SIPS. The SMD2 subgroup 'substance misuse and offending' includes in fact all types of substance misuses, and not only alcohol. SIPS is instead targeted specifically and only to offenders with alcohol misuse issues. On the other hand, the administrative data-led approach used for the estimation of the SMD categories considers only individuals who are already in touch with relevant services. Estimated prevalence rate might be therefore a reasonable estimate of the true prevalence of offenders who misuse alcohol.

Using the prevalence of SMD2 for the adult working age population of England, the number of potential recipients of the SIPS could be therefore estimated and compared to the total population in 2010 (Office for National Statistics, 2017). Results are reported in Table 7.5.

IMD	Total population	Adult working age 16-64	Potential recipients
Q1	10,541,505	6,824,097	28,892
Q2	10,663,393	7,102,625	20,241
Q3	10,560,965	6,886,727	17,300
Q4	10,483,579	6,748,755	11,418
Q5	10,395,075	6,555,841	8,650
Overall	52,644,517	34,118,045	86,502

Table 7.5 Potential recipients of SIPS by IMD quantile

Prisoners and alcohol misusers are characterised by higher deprivation levels, compared to the general population. Therefore, more potential recipients of SIPS can be found in lower IMD subgroups.

7.3.2.2. Impacts on offender population: results

In the analysis of SIPS from the HC perspective (see Chapter 4, Table 4.13) it was found that the DEs falling on offenders were -0.010 QALYs and -0.018 QALYs for BA and BLC, respectively. Impacts on offender population (expressed in QALYs) were obtained by multiplying the DEs of the intervention for the number of potential recipients. Results are shown in Table 7.6.

IMD	BA vs CIL	BLC vs CIL
Q1	-289	-520
Q2	-202	-364
Q3	-173	-311
Q4	-114	-206
Q5	-87	-156
Total	-865 QALYs	-1557 QALYs

Table 7.6 Impacts (measured in QALYs) on offender population

DEs of SIPS fall on the particularly disadvantaged group of individuals on probation with alcohol issues. Population negative health impacts of BA and BLC are therefore unequally distributed and cause more harm to lower SES subgroups. It must be remembered however that such estimate of the distributional impact of intervention's DEs is a crude estimate that does not take into account potential heterogeneous treatment effects.

7.3.3. Distributional impact on general population

Health OCs of SIPS falling on general population (including the additional HC costs to treat victims of crime) for BA and BLC are on average 0.009 QALYs and -0.044 QALYs, respectively (see Chapter 6, Table 6.7). These estimates can be multiplied for the number of potential recipients of SIPS (86,502, as reported in Table 7.5), to compute the total population health OCs. Resulting population health OCs of BA and BLC were estimated to be approximately 738 QALYs and -3,769 QALYs, respectively.

From the total population health OCs of the intervention, the distribution of population health OCs by SES can be computed. Love-Koh et al. (2016) explored the socioeconomic distribution of OCs in the English NHS, and provided the distribution of the impact on the population bearing the OCs according to their SES, measured by IMD quintile group. As illustrated in the first column of Table

7.7, individuals in the lower IMD quantiles bear a higher proportion of the health OCs (Love-Koh et al., 2016).

Using the information on the distributional impacts of the OCs of NHS activity, it is possible to map where the population health OCs of SIPS fall. The resulting distribution of the OCs of BA and BLC on the five IMD population's subgroups is illustrated in Table 7.7.

IMD	Proportion of total OCs falling on IMD subgroup*	Population health OCs (QALYs)	
		BA	BLC
Q1	26%	192	-980
Q2	22%	162	-829
Q3	22%	162	-829
Q4	16%	118	-603
Q5	14%	103	-528
Total	100%	738	-3769

* Adapted from Love-Koh (2016)

Table 7.7 Distributional impact of population health OCs of SIPS by IMD quantiles

OCs associated with BA and BLC are unequally distributed. BA is associated with higher OCs falling on lower IMD subgroups. On the contrary, higher cost savings (negative OCs) associated with BLC are experienced by individuals with lower IMD levels. For example, 192 QALYs and 103 QALYs are displaced from the first and fifth quantile of the BA group, respectively. Individuals with lower SES bear higher OCs associated with BA. On the contrary, 980 QALYs and 528 QALYs are gained by the first and fifth quantile of the BLC group, respectively.

7.3.4. Distributional impacts on (direct and unknown) victims

Reductions in criminal activity generate positive health consequences for the victims due to avoided QALY losses. Two last components must be included in the analysis, namely: the DEs on victims of crimes committed by offenders treated with SIPS, and the OCs falling on unknown victims.

Net health impacts on the victims are given by the DEs on victims net of OCs falling on unknown victims. Impacts on direct and unknown victims associated with BA were found to be 0.017 QALYs (DEs) and -0.0004 QALYs (OCs), respectively. Impacts on direct and unknown victims associated with BLC were 0.021 QALYs (DEs) and -0.0005 QALYs (OCs), respectively (see Chapter 6, Table 6.7). Therefore, INHBs on (direct and unknown) victims associated with BA and BLC were 0.017 and 0.021 QALYs, respectively. Given the total number of potential recipients of SIPS (86,502, see Table 7.5), population health effects falling on victims were estimated to be 1,467 QALYs and 1,898 QALYs for BA and BLC, respectively.

Due to the lack of information about the distribution of the effects falling on victims, it was assumed that their distribution follows offenders' prevalence. It might be in fact sensible to hypothesise that victims share some characteristics with the offenders (e.g. they share similar SES). It was thus assumed that impacts on victims were distributed as reported in Figure 7.2. If the prevalence of victims and offenders is similar across IMD subgroups, reductions in criminal activity disproportionately benefit individuals with lower SES. The resulting population INHBs falling on victims by IMD subgroup is illustrated in Table 7.8.

IMD	BA	BLC
Q1	490	634
Q2	343	444
Q3	293	380
Q4	194	251
Q5	147	190
Total	1467 QALYs	1898 QALYs

Table 7.8 Distributional impacts on victims

As expected, greater health benefits (associated with reductions in criminal activity) are experienced by individuals with lower IMD levels. For example, 490 (637) QALYs are gained by the BA (BLC) group with the lowest SES, compared to 147 (190) QALYs gained by the group with the highest SES.

7.3.5. Results of the DCEA of SIPS

Total population INHBs are given by the DEs on offender population net of the population health OCs, plus the INHBs on victims. Total population INHBs by IMD quantile associated with BA and BLC were calculated and are summarised in Table 7.9.

IMD	DEs on offender population	Population health OCs	INHBs on victims	Total population INHBs
Brief advice (BA)				
Q1	-289	192	490	**9**
Q2	-202	162	343	**-22**
Q3	-173	162	293	**-42**
Q4	-114	118	194	**-39**
Q5	-87	103	147	**-43**
Total	-865 QALYs	738 QALYs	1467 QALYs	**-136 QALYs**
Brief Lifestyle Counselling (BLC)				
Q1	-520	-980	634	**1094**
Q2	-364	-829	444	**909**
Q3	-311	-829	380	**897**
Q4	-206	-603	251	**648**
Q5	-156	-528	190	**562**
Total	-1557 QALYs	-3769 QALYs	1898 QALYs	**4110 QALYs**

Table 7.9 Population INHBs by SES

Looking at the DEs falling on offenders, greater negative health impacts falling on lower SES subgroups are associated with both BA and BLC. Focusing on the impacts on victims' health, greater benefits are experienced by more disadvantaged subgroups treated with both BA and BLC. The main difference between BA and BLC is in the distribution of OCs falling on the general population. Due to negative OCs (cost savings), BLC is associated with positive impacts health, especially for low IMD subgroups. On the contrary, BA is associated with higher OCs, falling particularly on more deprived subgroups.

Compared to CIL, total population INHBs of BA and BLC are -136 QALYs and 4,110 QALYs, respectively. If the focus of the analysis is on cost-effectiveness, it can be therefore concluded that on average total INHBs associated with BA are negative, and total INHBs associated with BLC are positive. If the focus is on health equity (i.e. health inequalities across SES subgroups), the distribution of total health impacts of SIPS can be assessed. The distribution of population INHBs for BA and BLC is illustrated in Figure 7.3.

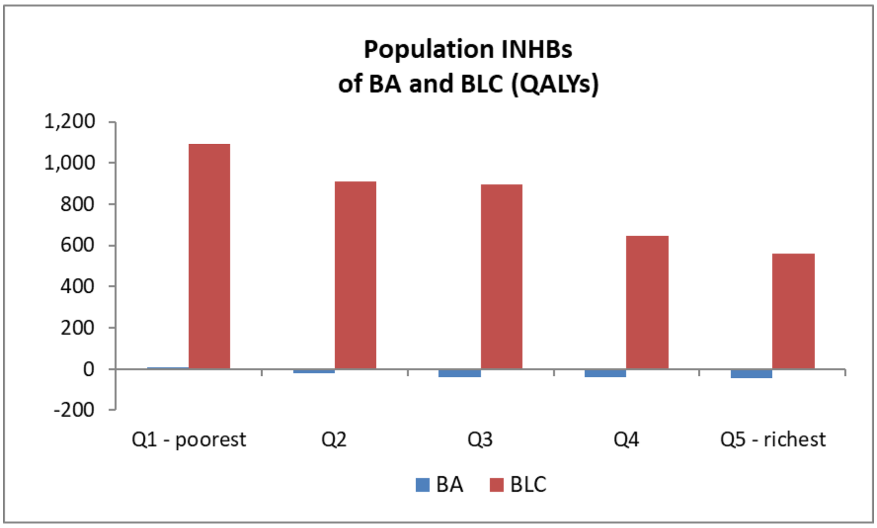


Figure 7.3 Incremental population health impact of BA and BLC compared to CIL, by IMD

Visually, it appears that both BA and BLC reduce health inequalities and are therefore equity-focused. Greater benefits fall on the most disadvantaged sub-groups, and benefits gradually declines for richer subgroups. In particular, BA is associated with (small but) negative impacts falling on less disadvantaged subgroups.

Impact on health equity can be expressed using the Atkinson index of inequality A(ε) (Atkinson, A. B., 1970), which is a function of population's inequality aversion (See Chapter 3, Sub-section 3.3.3.2, for further details). The value of the inequality aversion parameter ε was estimated to be on average 10.95 for the general public in England. Intuitively, such a value means that a marginal improvement of health for the individual in lowest IMD quantile is associated with a weight of 5.6 when compared to a marginal improvement of health for the individual in the highest IMD quantile[49] (Robson et al., 2017).

Both BA and BLC are associated with reductions in *A(ε)* and therefore to an improvement in health equity. The Health Equity Impact Plane can be used to summarise the results of the DCEA, in terms of both cost-effectiveness and equity impact. Cost-effectiveness (population INHBs) is represented on the vertical axis, and net health equity impact on the horizontal axis (Cookson et al., 2017). The sign of net equity impact is reversed (i.e. increases in -A(ε) represent an improvement in health

[49] This indirect equity weight was obtained taking the partial derivative of the EDEH with respect to health at baseline of the individual in the most disadvantaged group divided by the partial derivative of the EDEH with respect to health at baseline of the individual in the least disadvantaged group.

equality) to resemble the cost-effectiveness plane (Love-Koh et al., 2018). The health equity impact plane for BA and BLC is illustrated in Figure 7.4.

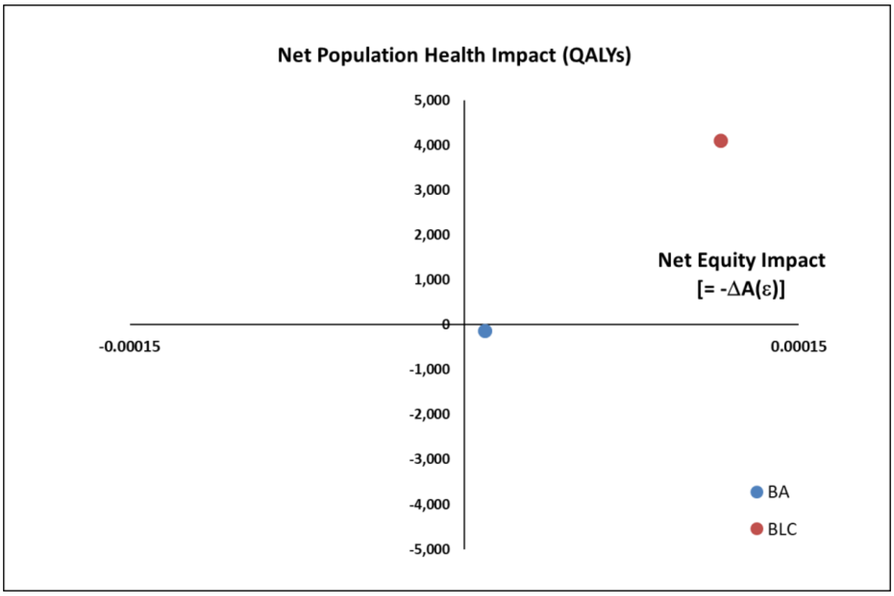


Figure 7.4 Health Equity Impact Plane for SIPS

Average impact associated with BLC falls in the 'win-win' quadrant of delivering health gains and reducing health inequity insofar as it disproportionately benefits socially disadvantaged groups (Cookson et al., 2017). BLC is therefore not only cost-effective, but also provides a greater improvement in health equity, compared to BA. BA is associated on average to negative population INHBs and a smaller improvement in health equity.

Regardless of the degree of inequality aversion, BLC is always worthwhile. With regard to BA, a sensitivity analysis around the inequality aversion parameter can be performed. Figure 7.5 reports the EDEH associated with BA compared to CIL for inequality aversion parameter values ranging from 0 to 40.

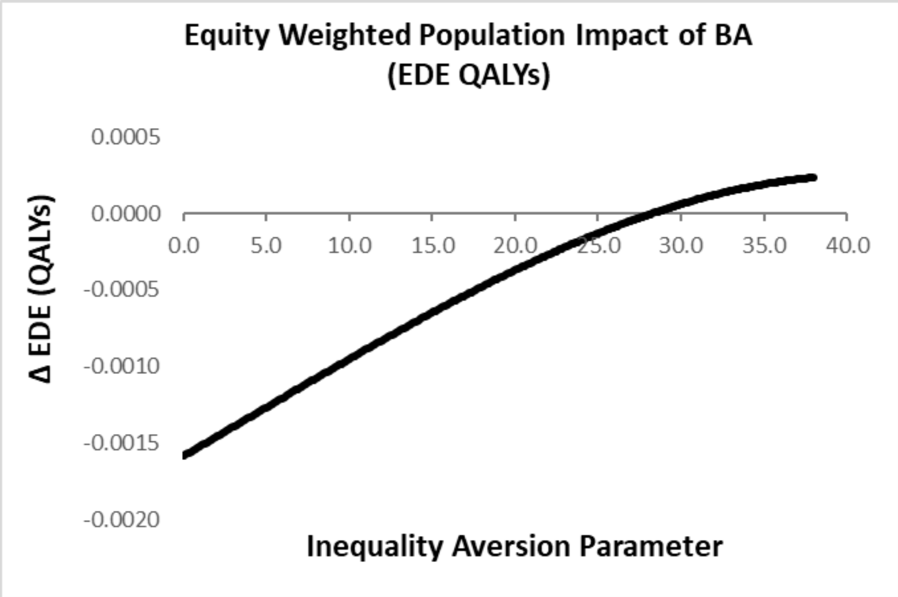


Figure 7.5 Sensitivity analysis around the inequality aversion parameter

The larger the degree of inequality aversion, the more worthwhile BA becomes compared to CIL. However, the tipping point where impact on EDEH becomes positive is associated with an inequality aversion parameter of approximately 28, which is unlikely to be reasonable. In fact, an indirect equity weight of 28 would imply weighting marginal health impacts falling on the most deprived individuals' health 80 times more than health impacts falling on the least deprived individuals.

7.3.5.1. Alternative DCEA excluding health and HC impacts falling on victims

Results presented in this section take into account also the impacts on victims of crime. To illustrate the importance of considering the full impact of the intervention, the same analysis was performed from the 'traditional' narrow HC perspective (i.e. without incorporating also health spill-overs on victims through the CJS and the associated additional HC costs).

As reported in Appendix Q, results were drastically different. Both BA and BLC appeared to be neither cost-effective nor likely to improve health equity. Conclusions and recommendations from the DCEA would thus differ according to the inclusion or exclusion of cross-sectoral health spill-over effects.

7.3.6. Incorporating DCEA results in the cross-sectoral analysis of SIPS

In the analysis conducted in Sub-section 7.3.5, the perspective adopted was the narrow viewpoint of the HC system. The aim of this sub-section is to use the results of the DCEA to inform a cross-sectoral analysis conducted from an integrated perspective that incorporates health equity concerns (Step 4 revised).

In the original cross-sectoral analysis conducted in Chapter 6, Section 6.3, total impact on the HC sector was summarised using INHB. In the revised cross-sectoral analysis conducted in Sub-section 7.2.3, equity weights were incorporated and the impact on the HC sector was summarised using INwHB instead of INHB. In this analysis, in order to take into account impacts in both total health and health distribution, variations in the EDEH were employed to summarise the impact of the intervention from the HC perspective. In fact, as explained previously in Chapter 3, Sub-section 3.3.3, EDEH represents the level of health per person which, if equally distributed, would give the same level of social welfare as the original distribution (Asaria, Griffin & Cookson, 2016).

The original equation to compute the INMBs (see Equation 6.2 in Chapter 6) can be adapted using variations in EDEH instead of INHB. The resulting cross-sectoral monetary benefits include therefore equity weights associated with health impacts falling on different subgroups, and are called INwMBs. It must be stressed that this estimate of the INwMB cannot be compared to the results presented in the previous section because those were based on different equity weights and on a different method to assess health impact (i.e. equity weighting).

Recalling the usual assumptions employed in this case study (i.e. constant ratio v/k), resulting INwMBs can be obtained as illustrated in Equation 7.6.

$$INwMB = V_H * \Delta EDE + V_R * INRB \tag{7.6}$$

When the inequality aversion parameter ε is equal to 0, the distribution of health impacts is not relevant and results are the same as the cross-sectoral analysis carried out in Chapter 6. This is because no weights are associated with health impacts falling on different population subgroups, and variations in EDE correspond to the INHBs reported previously in Table 6.11 (i.e. -0.002 QALYs for BA and 0.048 QALYs for BLC). Resulting INwMBs are therefore identical to the INMBs calculated in Chapter 6.

As ε increases, higher weights are put on the more disadvantaged subgroup. Assuming ε to be equal to 10.95 (Robson et al., 2017), variations in EDEH associated with BA and BLC are -0.001 QALYs and 0.055 QALYs, respectively. Both values are higher than the INHBs originally calculated. This is because both BA and BLC contribute to reduce health inequalities. Cross-sectoral INwMBs (based on a DCEA with ε = 10.95) of BA and BLC compared to CIL are reported in Table 7.10.

	Health Care Sector	Criminal Justice System	INwMB
BA	-12 £ (-0.001 QALYs)	6,106 £ (0.56 R)	28,032 £
BLC	713 £ (0.055 QALYs)	7,891 £ (0.72 R)	39,578 £

Table 7.10 Cross-sectoral INwMBs (based on DCEA) of BA and BLC compared to CIL

Compared to the original cross-sectoral analysis, INwMB values associated with BA and BLC are slightly higher than the INMBs estimated in Chapter 6. This is because INwMBs now reflect the beneficial impact of the intervention in reducing health inequities. Nevertheless, in terms of decision making, no differences can be observed. BLC is still the overall cost-effective option from an integrated perspective and requires no compensation across sectors to be implemented. BA is associated with positive (but lower) INwMBs, but to be implemented it would require the CJS to compensate the HC sector.

7.4. Final remarks

In this chapter, the impacts of two alternative value judgments were assessed. In the first part, using an equity weighting analysis, how to put higher value on victims' health was investigated. In the second part, it was assumed that the decision maker aims at prioritising interventions that benefit disadvantaged and least healthy groups, and a DCEA was carried out.

Alternative equity considerations could have been applied. Nevertheless, the main objective of this chapter was not to evaluate specifically the SIPS programme, but to show how to operationalise the incorporation of health equity concerns in the cross-sectoral evaluation of a public health intervention with effects on HC and CJ.

7.4.1. Within-dimension and within-individual approach to inequalities

In the proposed framework, the within-dimension approach was followed. As explained in Chapter 3, Sub-section 3.2.2.1, this approach aggregates outcomes firstly within and secondly across dimensions. The focus of this chapter was on the HC dimension. Nevertheless, concern about inequalities could be potentially introduced in aggregations within dimensions other than HC. Different functional forms representing alternative normative judgements could be used for example also when aggregating impacts on CJ, and the same methods could be potentially used to incorporate inequality concerns in other dimensions.

However, two potential limitations to the extension of this analysis emerge. Firstly, from a pragmatic point of view, the incorporation of equity concerns in sectors other than HC would be hindered by the lack of data, especially with regard to OCs. The second limitation is related to the aggregation approach, that firstly aggregates across individuals (within the dimension) and secondly across dimensions. Inequalities are therefore only assessed within each dimension, and the potential link between inequalities in multiple and related dimensions would be missed.

From a theoretical point of view, the adoption of a within-individual approach to aggregation could help addressing the latter issue and take into account the whole combination of inequalities across dimensions. As illustrated in Chapter 3, Sub-section 3.2.2.1, the within-individual approach consists of aggregating first within-individuals across all dimensions, and then across individuals (Walker et al., 2019). However, such an extension of the framework would require an adaptation of the methods that goes beyond the scope of this thesis.

7.4.2. Potential cross-temporal extension of the SIPS case study

In this chapter, the focus was placed on the 'health equity' extension of the cross-sectoral analytical framework. Nevertheless, the proposed framework can also potentially accommodate the inclusion of long-term consequences. However, this would lead to issues related to how to extrapolate effects and costs to a more appropriate time horizon, and assess the impacts of hard and soft budgets.

The potential cross-temporal extension of the cross-sectoral economic evaluation of SIPS is discussed in the next chapter, Sub-section 8.4.2.

Chapter 8: Discussion and conclusions

8.1. Summary of objectives and results of the previous chapters

Before discussing contributions, strengths and limitations of this thesis, a brief summary of objectives and findings of each chapter is provided here. Sub-section 8.1.1 recaps the theoretical part of this thesis; in Sub-section 8.3 results of the case study are summarised.

Afterwards, Section 8.2 highlights the key contributions of this thesis to the literature. Strengths and limitations of the framework are discussed Section 8.3. A discussion on the results of the case study is provided in Section 8.4. The same section explores also potential extensions of the analysis, and compares the proposed case study with other case studies taken from the literature. In Section 8.5, recommendations and further research are illustrated. Lastly, Section 8.6 draws the conclusions and points out the policy implications of this research.

8.1.1. Theoretical part

8.1.1.1. Chapter 2

In Chapter 2, an overview of the theoretical background for economic evaluation of HC interventions was provided to set the scene for the thesis. A taxonomy of the main theoretical underpinnings and analytical techniques was developed. Finally, the appropriateness and limitations of the different approaches and methods when applied to the evaluation of public health interventions were investigated.

8.1.1.2. Chapter 3

Chapter 3 covered the development of the proposed analytical framework and described its theoretical underpinnings (i.e. the normative approach taken) and the analytical techniques employed.

8.1.2. Case study

In Chapters 4, 5, 6 and 7, the SIPS case study was used as a vehicle to illustrate the operationalisation of the proposed framework. A summary of objectives and findings of each chapter is provided here.

8.1.2.1. Chapter 4

The cost-effectiveness of SIPS was investigated from the ‘traditional’ HC perspective (i.e. including exclusively costs falling on health and social care budget, and effects on trial participants). It was found that the highest NBs were associated with CIL, and it was concluded that BA and BLC did not appear to be cost-effective compared to CIL.

8.1.2.2. Chapter 5

The cost-effectiveness of SIPS was investigated from the CJS perspective, and incremental costs falling on the CJS budget were compared to avoided reconvictions, which were used as a measure of effectiveness. Results of the CEA showed that from the CJS perspective it was worth investing in either BA or BLC when compared to CIL, but the overall cost-effective intervention was BLC.

8.1.2.3. Chapter 6

The CEA from the ‘traditional’ HC perspective was expanded to incorporate also spill-over effects on (direct and unknown) victims’ health and the additional costs falling on the HC budget to treat them. According to this revised CEA conducted from the HC perspective, BLC was found to be the cost-effective option.

A cross-sectoral analysis from a joint health and criminal justice perspective was also conducted. It was assumed that decision makers were interested in both health and criminal justice, and impacts on both HC and CJ sectors were included. According to this cross-sectoral CEA, both BA and BLC were associated with positive INMBs when compared to CIL, and the overall cost-effective option was BLC.

8.1.2.4. Chapter 7

Two independent analyses were carried out to investigate the impacts of incorporating alternative value judgments into the economic evaluation. The first analysis was based on the hypothesis that impacts on victims’ health (i.e. variations in health caused by criminal activity) are valued differently from generic health variations. To incorporate this value judgment in the evaluation, an equity trade-off analysis based on equity weighting was conducted. Resulting INwHBs (i.e. INHBs that incorporate equity concerns) were higher when compared to the original INHBs estimates from the ‘revised’ CEA, and both BA and BLC were found to be on average cost-effective, compared to CIL. Nevertheless, from the HC perspective, final results did not change. BLC was still on average the overall cost-effective option. Similarly, results of the cross-sectoral analysis showed that BLC was also the overall cost-effective treatment option from an integrated and broader perspective.

The second analysis was based on the hypothesis that decision makers want to prioritise interventions that benefit disadvantaged and least healthy groups. An equity impact analysis based on a DCEA was carried out. From the HC perspective, BLC was found to be not only cost-effective, but also associated with a greater improvement in unfair health inequalities, compared to both BA and CIL. From a broader perspective, cross-sectoral INwMBs (i.e. INMBs that include health equity concerns) values associated with BA and BLC were slightly higher than the original cross-sectoral INMBs. From an integrated perspective, BLC was the overall cost-effective option.

8.2. Key contributions to the literature

Public health interventions are often characterised by costs and outcomes falling on multiple sectors, with impacts that might occur far in the future. Furthermore, such interventions frequently aim to address health equity issues. The lack of a consistent framework to evaluate their cost-effectiveness has been emphasised in the literature (Weatherly et al., 2009, Wanless, 2004). The call for the development of methods for the economic evaluation of public health interventions is particularly strong in the context of alcohol misuse (Barbosa, Godfrey & Parrott, 2010).

With a particular focus on cross-sectoral aspects, the importance of investigating broader perspectives of CEAs and considering outcomes across different sectors has been recently highlighted by a panel of experts (Sanders et al., 2016). The panel emphasised that changes in health may lead to positive externalities (e.g. reduction in crime). In order for the analysis to reflect the full consequences and opportunity costs of an intervention, a valuation of these externalities is required. Furthermore, the panel stressed that it is not sufficient to include in the evaluation only the total reduction in non-health care costs (e.g. police and CJ costs). Changes in non-health outcomes (e.g. feeling safer and experiencing less crime) must also be taken into account. Therefore, further development of methods for capturing the value of broader impacts from real-world health interventions is needed (Neumann et al., 2018).

The aim of this thesis was to begin to address these gaps in the literature. More precisely, this work had two objectives. Firstly, to develop an analytical framework for the evaluation of public health interventions. The framework consists of a cross-sectoral analysis, with the potential incorporation of health equity and cross-temporal concerns. Secondly, to show how to operationalise the framework, using an intervention to reduce alcohol consumption among offenders as a case study.

8.3. Strengths and limitations of the framework

In theory, the proposed framework could be used for the economic evaluation of any intervention. However, there might be cases where all costs are likely to fall on the health care sector, there are no obvious impacts on outcomes other than health, there are no substantial equity issues, and no problems with intertemporal concerns arise. In those cases it would be reasonable to proceed with a conventional analysis conducted from the narrow health care system perspective. However, particularly in the context of public health, it would be difficult to identify such cases.

For all the other cases, a first step would be to assess whether relevant impacts are expected to fall outside the health care sector. As recommended by the 2nd Cost-Effectiveness Panel, analysts should attempt to quantify non-health consequences in the impact inventory unless those consequences are likely to have a negligible effect on the result of the analysis (Sanders et al., 2016). Similarly, the relevance of cross-temporal and equity impacts could be also investigated. Evidence on public health interventions that have impacts on inequity has been recently provided (Griffin et al., 2019). However, the choice of including or excluding cross-sectoral, equity and cross-temporal concerns in an evaluation is likely to always be disputable.

For this reason, the strength of the proposed framework resides in its flexibility. The framework can be easily adapted. In the simplest way it can be just used as a traditional CEA from the health care system perspective. Then it is possible to build up the analysis to consider other sectors and impacts. It can be used for the evaluation of any public health intervention, and, when deemed necessary, it can focus on cross-sectoral aspects, and/or health equity issues, and/or cross-temporal aspects. Further, assumptions (e.g. choice of dimensions and outcomes, time constraints, and characteristics of the budget) and value judgments (e.g. about equity) are explicitly stated and can be updated or modified. Impacts of these decisions can be tested by looking at alternative choices.

Not all the information necessary to operationalise the framework are currently available (e.g. marginal productivity estimates in sectors other than health care are rare, as further discussed in Sub-section 8.5.2). Nevertheless, even in absence of precise evidence, sensitivity analyses examining alternative values can be conducted. Moreover, results could be easily updated when such estimates become available. Another contribution of the framework is therefore to identify the gaps in the evidence that need to be addressed.

The proposed framework is flexible to alternative approaches for the identification and aggregation of outcomes. However, one pragmatic limitation of the framework is that it still requires the specification of dimensions and outcomes to be included in the impact inventory. In fact, the field is still far from reaching an agreement on which elements to include in such a summary, how these should be determined, and how to value them (Neumann et al., 2018, Sanders et al., 2016). For this reason, the inclusion or exclusion of dimensions and outcomes from the evaluation is likely to be open to debate.

For example, as mentioned already in Chapter 3, Sub-section 3.2.1.1, and Chapter 4, Sub-section 4.4.4.1, in the usual practice, economic evaluations conducted from the HC perspective typically include both health care and social care costs (NICE, 2013). Nevertheless, health spill-over effects on family members and informal caregivers are very rarely considered in practice (Al-Janabi, Van Exel, Brouwer, Trotter, et al., 2016). Recent literature appears in favour of the inclusion of health spill-over effects on caregivers in the evaluation (Sanders et al., 2016), and methods for quantifying health spill-overs and including them in economic evaluation have already been proposed (Al-Janabi, van Exel, Brouwer & Coast, 2016, Al-Janabi, Van Exel, Brouwer, Trotter, et al., 2016). Nevertheless, no final agreement about the appropriateness of including these effects in economic evaluations has been reached (Brouwer, 2018, McCabe, 2018). Because of this ongoing debate, the inclusion of other cross-sectoral health spill-over effects (e.g. impacts on victims' health due to spill-over effects through the CJS) could be disputed. In fact, if health spill-over effects on carers are not considered from the HC perspective, the inclusion of other health impacts (e.g. falling on the victims of crime) could be deemed difficult to justify.[50]

Two more general limitations of the framework must be pointed out. Firstly, the inclusion of spill-over effects might introduce potential double counting issues (Sanders et al., 2016). Secondly, assumptions about decision makers' budgets might not be realistic. The framework assumes a single payer for each sector, while in reality there are likely to be more payers with similar objectives, but separate budget constraints (Remme, Martinez-Alvarez & Vassall, 2017). The implications of these two potential issues are explored in the next sub-section, using an illustrative example in the context of social care.

[50] Moreover, if impacts on victims are incorporated in the analysis, should health impacts of offenders' families be included as well? However, this is not a limitation of the framework itself. These are issues of ethics in social decision making.

8.3.1. Budgets and decision makers' objectives: the case of social care

The SIPS case study was grounded on the assumption of non-overlapping and non-correlated outcomes. It was assumed that the intrinsic value of crime was independent from the impact on victims' health. Consequently, the cross-sectoral aggregation of the outcomes could be simply additive without resulting in any double counting. The cross-sectoral outcome was thus given by the sum of the CJ outcome (re-convictions) and HC outcome (QALYs), when expressed in monetary terms. Moreover, HC and CJS budgets were assumed to be separate and independent. The structure of budgets and outcomes considered in the SIPS case study is illustrated in Figure 8.1.

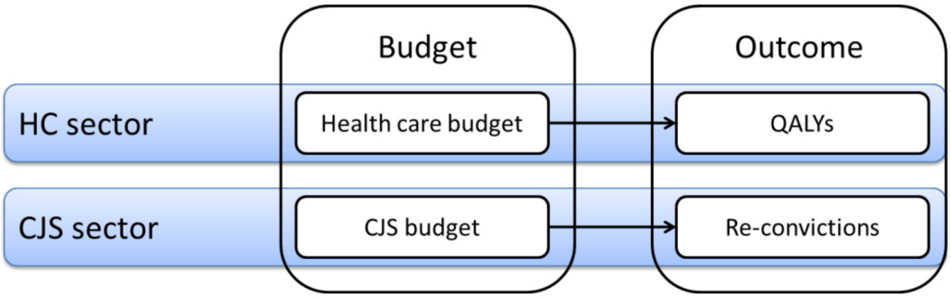


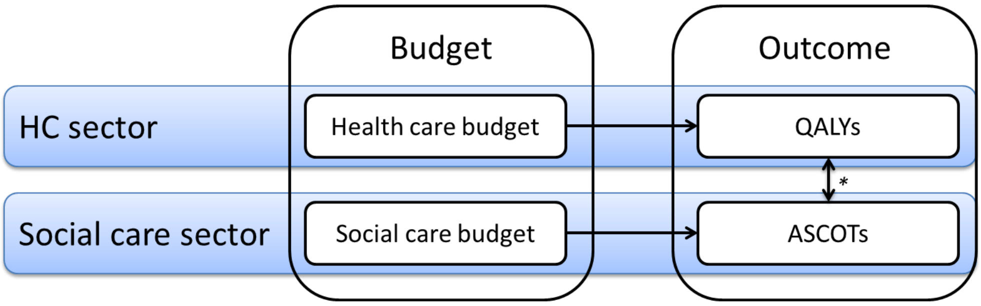
Figure 8.1 HC and CJ sectors: budgets and outcomes

Potential issues might arise when the additivity assumption does not hold (because outcomes are overlapping or correlated), or when decision makers are not responsible for only one budget and one outcome. As pointed out in Chapter 4, Sub-section 4.4.4.1, this might be the case if the framework is applied in the health care and social care context.

The aim of this sub-section is to describe potential scenarios related to characteristics of budgets and outcomes, and highlight practical consequences for the proposed framework of analysis. Using the example of an evaluation including the two dimensions of health care and social care, two potential structures of outcomes and budgets are explored.

8.3.1.1. Assumption 1: separate budgets, but overlapping and/or correlated outcomes

In this first scenario it is assumed that health care and social care are considered as two separate dimensions, with two independent budgets. It is also assumed that QALY and Adult Social Care Outcomes Toolkit (ASCOT) are the two measures to assess health and social care outcomes, respectively. QALYs are in fact widely used in the economic evaluation of health, and ASCOTs are frequently used to evaluate social care interventions (Brazier & Tsuchiya, 2015). The structure of health care and social care budgets and outcomes is illustrated in Figure 8.2.



** The double arrow highlights the correlation between the outcomes*

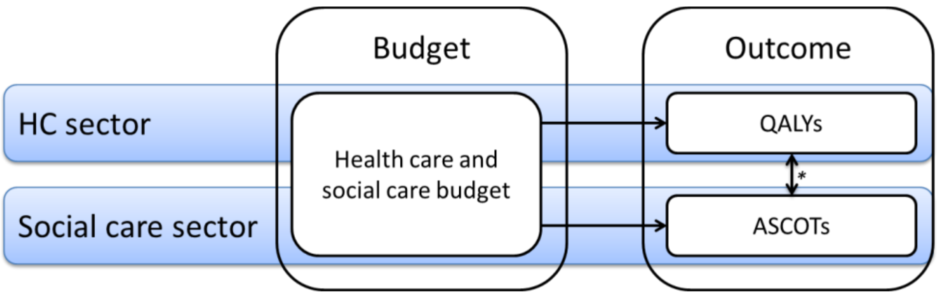
Figure 8.2 HC and social care sectors: budgets and outcomes (assumption 1)

In such a situation, health care budget is invested to maximise QALYs and social care budget to maximise ASCOTs. The problem would be that QALYs and ASCOTs are not conceptually independent (the correlation is highlighted with a double arrow in Figure 8.2), because they capture some of the same benefits (Brazier & Tsuchiya, 2015). Therefore, there would be a risk of double-counting and the additivity assumption would no longer hold (Walker et al., 2019).

This sub-section does not aim to provide methods to solve these issues. Rather the aim is to highlight these potential problems and point to relevant literature addressing them. Recent evidence about the relationship between ASCOT and EQ-5D-3L (Stevens, Brazier & Rowen, 2018) might help addressing the double counting issue. Alternatively, a broad generic measure of quality of life that focuses on health and social care might be used, such as the E-QALY, which is currently under development (Brazier, 2018).

8.3.1.2. Assumption 2: shared budget, and distinct but overlapping and/or correlated outcomes

In this second scenario, it is assumed that health care and social care decision makers share the same budget. In other words, the same resources are available to fund both health care and social care interventions. The structure of health care and social care budgets and outcomes is illustrated in Figure 8.3.



** The double arrow highlights the correlation between the outcomes*

Figure 8.3 HC and social care sectors: budgets and outcomes (assumption 2)

Health care and social care could be then considered as one sector producing two outcomes, and it would be then necessary to have two thresholds (one for QALYs and one for ASCOTs) for the same budget[51]. However, health care and social care could also be considered as two separate sectors with two independent decision makers having different remits. In such a situation, it would not be realistic to assume the existence of one budget, one outcome and one decision maker for each dimension. This is because not only would outcomes (QALYs and ASCOTs) be potentially correlated and partially overlapping, but also the existence of a shared budget would make the compensation scheme across sectors more difficult to implement. Moreover, because outcomes are not directly linked to distinct decision makers' budgets, it would be more complicated to calculate opportunity cost estimates based on the marginal productivities of the sectors.

8.4. Discussion about the case study

Methods and results of each step of the analysis were discussed in each chapter. However, in this section, some key issues are explored more in detail. In Sub-section 8.4.1, the main limitations of case study results are described. In Sub-section 8.4.2, an analysis plan for the potential cross-temporal extension of the case study is provided. Such an extension is left as potential further research. Lastly, in Sub-section 8.4.3, the original contribution of this case study is highlighted by comparing it with other case studies taken from the literature.

8.4.1. Comments on the results of the case study

8.4.1.1. Potential spurious effect of SIPS on criminal activity

Results of the case study should be interpreted with caution. When originally assessing the effectiveness of SIPS, Newbury-Birch et al. (2014) pointed out that the reduction in criminal recidivism was found in the absence of significant differences in drinking consumption between the groups. Nevertheless, even in absence of differences in alcohol consumption across interventions, increased awareness might have resulted in a change in offending behaviour rather than consumption *per se*, or offending might be linked to particular patterns of drinking rather than overall consumption.

Nevertheless, the main objective of the case study was not to provide a conclusive recommendation specifically about the implementation of SIPS. Instead, the case study was used as a vehicle to

[51] For further details, see Walker et al. (2019).

illustrate how to operationalise the theoretical analytical framework (proposed in Chapter 3) in the specific context of a public health intervention having impacts on HC and CJ. The importance of using a real intervention instead of a fictional example is further discussed in Section 8.5.

8.4.1.2. Cost-effectiveness of SIPS from the CJS perspective

The CEA conducted from the CJS perspective was grounded on the assumption that the reduction in crime has a value *per se*[52]. As explained in Sub-section 5.2.2, the decision maker in the HC sector has the objective of maximising health because health has a value *per se* (Drummond, M. et al., 2015). Similarly, the aim of the CJS decision maker is to reduce crime because a negative value is associated with crime *per se*. In other words, when a treatment reduces crime rates, the full valuation should not only include the total reduction in CJ costs, but also the increase in non-health outcomes, such as feeling safer and experiencing less crime (Neumann et al., 2018)

However, a brief review of the previous literature of CEAs in CJS was conducted (see Chapter 5, Sub-section 5.2.1) and it was found that frequently no value was actually assigned to reductions in crime *per se*. Often analyses only compared costs and cost savings falling on CJS. Moreover, even assuming that society has a positive WTP to reduce crime *per se* (Cohen et al., 2004), the WTP value used in the case study might actually overestimate the intrinsic value of reducing crime. As mentioned in Chapter 5, Sub-section 5.4.2.3, it would be necessary to investigate the WTP value of reducing crime *per se*, with the explicit exclusion of consequences on victims' health. Alternatively, for example, a potential outcome measure for CJS activity based on perceived safety could be developed, as discussed in Sub-section 8.5.2.

8.4.1.3. Different recommendations from different perspectives

In the case study it was shown that conclusions and recommendations differ according to the perspective adopted for the evaluation. CEAs provided different results if conducted from: the 'traditional' HC perspective (Chapter 4); the CJS perspective (Chapter 5); a 'revised' HC perspective which includes spill-over effects via CJS (first part of Chapter 6); an integrated perspective considering both HC and CJ impacts (second part of Chapter 6).

Results of the 'revised' CEA from the HC perspective could be further investigated. As suggested in Chapter 6, Sub-section 6.2.2, broader impacts on victims' health, such as mental distress and fear

[52] As discussed in Chapter 5, it was assumed that reduction in crime is not only associated with cost savings and positive health consequences, but has a value per se due to safety perceived and less crime experienced.

of crime could be also included in the evaluation (Dolan & Peasgood, 2007). Furthermore, in the case study only effects on direct victims' health were considered. Health consequences on individuals not directly involved in the crime, but still experiencing health consequences (i.e. indirect victims) could be included as well. Therefore, in the case study, health impacts that were associated with each crime might actually underestimate the full impact on victims of crime. If these additional components were included in the analysis, health spill-over effects generated in the CJS would be higher, and the inter-sectoral marginal productivity of the CJS in generating health would be affected. Nevertheless, the theoretical framework would still be valid.

With regard to the integrated cross-sectoral analysis, results also depend on crucial assumptions, concerning for example the marginal productivity (that reflects the opportunity costs) of the CJS. To address the lack of estimates in sectors other than HC, the ratio of *v* over *k* was assumed to be constant across sectors. Such an assumption has been already suggested in the literature (Claxton et al., 2019). However, if new empirical evidence is produced then alternative assumptions about the allocation of resources across sectors could be made. The value of the ratio *v/k* could be modified and the framework would still be valid. This work highlights the lack of data and need of further research, which are further discussed in Section 8.5.

8.4.1.4. Health inequities and social determinants of health

A discussion of results and limitations of the proposed solutions to address health equity issues was provided in Chapter 7, Sub-sections 7.2.4 and 7.4.1. In brief, compared to equity weighting, DCEA based on SES subgroups appears to be the most suitable method to incorporate health equity concerns in the proposed framework. The identification of population subgroups based on their SES is an overarching equity principle that can be used in different contexts. Furthermore, from a pragmatic point of view, data that are necessary to conduct DCEAs are already available.

However, it must be pointed out that, according to the results of the DCEA of SIPS, the impacts of the intervention in reducing unfair health inequalities is very small in absolute terms. Even if delivered to all potential users, effects on QALEs of each subgroup are negligible. This might be because the intervention focused more on lifestyle choices, and not on broader social and economic circumstances, known as the 'social determinants of health' (Dahlgren & Whitehead, 1993, 1991). In fact, strong evidence linking broader social determinants of health to good and poor health has been provided, and, if the main objective of the decision maker is to address health inequalities, action is required across all the wider determinants of health (Marmot et al., 2010).

According to Marmot et al. (2010), persisting inequalities across key domains such as early child development and education, employment and working conditions, housing and neighbourhood conditions, standards of living, and, more generally, the freedom to participate equally in the benefits of society, provide ample explanation of health inequalities. Therefore, action taken by the DHSC and the NHS alone will not reduce health inequalities. By contrast, acting on education and sense of community would be probably more effective in order to address alcohol issues. Individuals with less education report being in poorer health, they are more likely to smoke, more likely to be obese and suffer alcohol harm. Similarly, social networks have been shown to be as powerful predictors of mortality as common lifestyle and clinical risks such as excessive alcohol consumption (Buck & Gregory, 2013). In fact, one of the key purposes of relocating public health into local authorities was precisely to address the wider determinants of health and well-being in a setting that was perceived to be more sympathetic to such a structural approach than the NHS had been (Marks et al., 2015).

8.4.2. Potential extension of the case study: cross-temporal analysis

Public health interventions can be particularly difficult to evaluate because significant effects might occur far in the future (WHO, 2015, Weatherly et al., 2009). It might therefore be necessary to extend the time horizon for the estimation of costs and effects (Bojke et al., 2017). Two aspects then become crucial for the analysis of public health interventions: the extrapolation of costs and effects (e.g. beyond the available trial data), and the impacts of budgetary policies and investment strategies.

In the case study, the focus was mainly on the cross-sectoral aspects of the analysis, and on the incorporation of health equity concerns. However, the framework can also potentially accommodate an intertemporal extension of the evaluation, with the inclusion of long-term consequences and the assessment of impacts of alternative budgetary policies. Extrapolation and budgetary impact issues are addressed in Sub-sections 8.4.2.1 and 8.4.2.2, respectively. Sub-section 8.4.2.3 illustrates how to operationalise the cross-temporal extension of the framework.

8.4.2.1. Extrapolation of SIPS costs and effects

Even though public health programmes may impact on health over the longer term, long-term consequences are frequently neglected due to the difficulties in measuring them (Weatherly et al., 2009). Data collected often have a follow-up duration which is too short to allow for a robust quantification of the likely costs and benefits over the patient's entire lifetime. To estimate the full long-term impacts of a public health intervention, extrapolation techniques based on assumptions regarding the behaviour of the quantities of interest beyond the time horizon supported by the clinical evidence might be required. Decision-analytic modelling can be employed to extrapolate costs and effects of the programme which are too distant in time to be observed and directly measured (Bojke et al., 2017).

When carrying out economic evaluations of alcohol interventions, the need to assess and model long-term consequences is widely recognised. Alcohol treatments have the potential to improve alcohol-related mortality and long-term morbidity through a change in drinking behaviour (Hoang et al., 2016, Barbosa, Godfrey & Parrott, 2010). In particular, screening and brief alcohol interventions (such as SIPS) may produce long-term HC and CJ resource use cost savings. However, it is not clear how long the impact of brief interventions can be expected to last, and there is considerable uncertainty as to whether such an interventions will be cost saving in the long-term (Latimer et al., 2009).

The use of decision analytic modelling permits examination of longer time frames, preferable to the typically short-term economic evaluations of health interventions for alcohol problems conducted alongside RCTs. At the expense of introducing additional assumptions, modelling offers several advantages, such as extrapolating beyond the data observed in a trial, and linking intermediate clinical endpoints to final outcomes (Hoang et al., 2016). Examples of modelling used for the extrapolation of health care costs and effects associated with brief alcohol interventions can be found in the literature. For example, Barbosa et al. (2010) developed a framework based on a Markov model with a lifetime horizon to evaluate brief alcohol treatments taking into account longer term health outcomes and costs. Similarly, Tariq et al. (2009) used a state-transition Markov-type model to estimate impacts on costs and QALYs of a brief alcohol intervention in the long run. As the long-term effects of alcohol brief interventions are uncertain, the model took into account this uncertainty regarding the long-term persistence of the effect on alcohol consumption. Purshouse et al. (2013) and Angus et al. (2014) used the Sheffield Alcohol Policy Model (SAPM) (see Chapter 7, Sub-section 7.1.1) to extrapolate also health spill-over effects of a brief alcohol intervention through transport (due to road traffic accidents) and CJ (due to assaults).

However, the perspective of the economic evaluation in all the aforementioned studies was that of the health care system. Only health care costs and effects were considered. The need to develop a theoretical model to conduct a similar analysis from a broader perspective has been recognised (Barbosa, Godfrey & Parrott, 2010). Nevertheless, no long-term alcohol models that extrapolate non-health care costs and consequences were found in the literature (Hoang et al., 2016). Therefore, it would be interesting to investigate the long-term impact of SIPS from a broader perspective, and develop a model (or adapt the existing SAPM in order to record CJ outcomes) where cross-sectoral consequences are also taken into account. The short-term economic evaluation of SIPS carried out in this thesis might have underestimated the expected health or non-health outcomes and overestimated the costs of the intervention.

8.4.2.2. Impact of budgetary policies on SIPS

Let us suppose that the extrapolation process has been conducted and lifetime cross-sectoral costs and effects of SIPS are available. Long-term (HC and CJ) costs and effects could be then included in the evaluation, and the cross-sectoral cost-effectiveness estimate could be updated in the light of the new evidence.

Nevertheless, in the context of public health interventions, a CEA based on the assumption of the existence of soft budgets may be unrealistic. This is because standard decision rules in CEA based on expected net benefits usually assume the existence of 'soft' budgets constraints. In other words, as long as the health care decision maker plans to meet the budget at expectation, any deficit will be indemnified (McKenna et al., 2010). Consequently, the timing of costs and benefits is only relevant for discounting purposes. In reality, decision makers are constrained in their ability to commit resources today to obtain long-term benefits. For example, local authorities operate within tightly regulated annual budget cycles that fit within a 4- to 5-year planning cycle (Frew, 2016). Moreover, policy makers could prioritise short-term impacts, or ignore consequences happening in a distant future beyond the electoral cycle (Hoang et al., 2016).

As illustrated by McKenna et al. (2010), optimisation solutions of the health maximisation problem faced by the decision maker (and therefore investment decisions) can change according to the existence of soft or rigid budget constraints. For example, if the decision maker cannot run a deficit (i.e. is provided with a 'hard' budget), substantial opportunity costs could be imposed. This could ultimately lead to corner solutions, when no health care is provided due to the risk of exceeding the constraint, or restrictive provision of health care and large expected budget surpluses. Consequently, it might happen that an intervention which is theoretically cost-effective cannot be implemented because of a rigidity of the decision makers' budget constraints. For this reason, when evaluating public health interventions, rigidities in decision makers' budgets constraints should be included.

8.4.2.3. Implications for decision making

The cross-sectoral analysis was based on the decision rule that if INBs are positive, then the intervention should be funded[53]. Nonetheless, such a decision rule would be a true reflection of the real decision problem only if the timing of costs and effects was not relevant. In reality, costs and effects might not occur all at once and, in the presence of rigidities in budgets, their timing might be crucial. More rigid budgetary policies might determine a prioritisation of those investments which guarantee short-term effects.

The cross-sectoral INMB formula shown in Chapter 3, Equation 3.8, can therefore be adapted: impacts can be redefined as time series in order to make the cross-temporal aspects explicit and assess the effects of budgetary policies. When impacts (in terms of costs and effects) of the intervention are defined as a stream of costs and effects in time, two new variables play a role in the decision making process: the discount rate and the time horizon. In fact, not only decision makers can place a higher discount rate on outcomes realised far in the future, but also completely ignore consequences happening after a specified time horizon (e.g. they might not be interested in consequences after their political mandate is over). Moreover, costs and effects might accrue with different timings, and decision makers might have different preferences about costs and effects. In other words, their time horizon and discount rate for costs might be different from the way effects are traditionally valued. The revised cross-sectoral INMB formula, including the cross-temporal extension, is shown in Equation 8.1. Two time horizons (*T1* and *T2*) and discount rates (*r* and *d*) are introduced, one for the costs, and one for the effects.

$$INMB = V_H\left(\sum_{t=0}^{T1}\frac{\Delta H_t}{(1+r)^t} - \sum_{t=0}^{T2}\frac{\frac{\Delta C_{Ht}}{k_H}+\frac{\Delta C_{Nt}}{k_{NH}}}{(1+d)^t}\right) + V_N\left(\sum_{t=0}^{T1}\frac{\Delta N_t}{(1+r)^t} - \sum_{t=0}^{T2}\frac{\frac{\Delta C_{Nt}}{k_N}+\frac{\Delta C_{Ht}}{k_{HN}}}{(1+d)^t}\right) + \\ + \left(\sum_{t=0}^{T1}\frac{\Delta C_t}{(1+r)^t} - \sum_{t=0}^{T2}\frac{\frac{\Delta C_{Nt}}{k_{NC}}+\frac{\Delta C_{Ht}}{k_{HC}}}{(1+d)^t}\right) \quad (8.1)$$

Equation 8.1 could be used to explore alternative scenarios for the discount rate and the time horizon. The cross-sectoral economic evaluation of SIPS (with the potential incorporation of health equity concern) could be then tailored according to decision makers' constraints and objectives. For example, the discount rate might be set to reflect the higher priority associated with short-term impacts, and the time horizon set to reflect the electoral cycle. Because the choice of alternative timeframes might significantly influence the results (Hoang et al., 2016), this cross-temporal

[53] It is assumed that compensation schemes across decision makers can be enforced.

analysis could emphasise the potential net benefit that is loss due to political constraints and budgets rigidities. Moreover, by also providing a long-term evaluation based on soft budgets, the framework can be used to show explicitly the missed investment opportunities (in terms of potential INBs not realised) due to budgetary policies. By incorporating long-term impacts, these results would help to make the case for investment in public health interventions.

8.4.3. A comparison with other case studies from the literature

When developing the analytical framework proposed in this thesis, a list of crucial aspects in the economic evaluation of public health interventions emerged, and a checklist was developed. The checklist proposed in Table 8.1 includes: the choice of the normative approach and the theoretical underpinnings behind the evaluation; the inclusion of cross-sectoral costs and outcomes; the identification of opportunity costs; the incorporation of equity concerns; and the assessment of cross-temporal issues. This list of topics also reflects the open questions and research priorities in the approach to the economic evaluation of public health interventions pointed out by Chalkidou et al. (2008) and Edwards et al. (2013), and the set of considerations for researchers conducting economic evaluations with complex outcomes developed by Wildman et al. (2016).

Normative approach
1. What normative approach (i.e. theoretical underpinning) was used for analysis?
Cross-sectoral costs and effects
2. Was the perspective of the economic evaluation clearly defined? 3. Were the main beneficiaries of the intervention clearly identified? 4. Were the agencies bearing the costs of implementing the intervention clearly identified? 5. Was the health attribute separated from other attributes? 6. Were multiple outcome measures used (e.g. listed in an impact matrix)? 7. Were all benefits expressed using the same unit (e.g. in monetary terms)?
Opportunity costs
8. Were the opportunity costs of implementing the intervention identified? 9. Were estimates of the marginal productivity within and across the sectors used?
Equity concerns
10. Were any equity judgments explicitly stated? 11. Was the impact of the intervention on health inequity investigated? 12. Was the impact of the intervention on other inequities investigated?
Cross-temporal aspects
13. Were costs and outcomes extrapolated meaningfully over lengthy time horizons? 14. Was the impact of the choice of time horizon on the cost-effectiveness explored? 15. Was the impact of budgetary policies on the decision making process analysed?

Table 8.1 Checklist for the development of the framework

Compared to the methods proposed in this thesis (and summarised in Chapter 3, Sub section 3.4.1), different approaches and analytical techniques could have been adopted to address all these aspects. The checklist reported in Table 8.1 can be therefore used when conducting economic evaluations of public health interventions. Such an assessment would encourage analysts to be transparent in the approach they have taken, and explicit about assumptions and value judgments underpinning the economic evaluation. A few examples are provided in the following sub-sections.

8.4.3.1. Co-financing approach

The importance of allocating resources efficiently across sectors has been recently highlighted by Remme et al. (2017), who developed the so-called ‘co-financing approach’, which is also based on the work by Claxton et al. (2007). The authors used Culyer’s (2016) bookshelf metaphor to describe how budget redistributions among health and non-health payers can help to achieve a more optimal resource allocation[54]. An intervention with multi-sectoral benefits might end up being under-financed when either one of the other sector has to pay for the whole intervention. By contrast, if the cost of intervention is shared between the sectors, the productivity of the intervention per unit of expenditure increases, making it better value for money.

In a previous work, Remme et al. (2014) used the co-financing approach to evaluate a structural intervention to reduce HIV vulnerability by keeping adolescent girls in school in Malawi. Impacts of the intervention on both health and education sectors were assessed, and the approach to the evaluation appears to be very similar to the analysis proposed in this thesis. Nevertheless, looking at the main aspects pointed out in the checklist in Table 8.1, a few differences can be highlighted. Firstly, in the cross-sectoral analysis, the rule for cost-effectiveness was based on a WTP threshold at a cost per disability-adjusted life year (DALY) averted below country's gross domestic product (GDP) per capita. Even though such a threshold is recommended by WHO (Edejer et al., 2003) and is commonly used in economic evaluations of HIV interventions, it is not based on an empirical assessment of the likely health OCs (Ochalek, Lomas & Claxton, 2018). Similarly, no attempt was made to estimate the OCs for the education sector, and the cost-effectiveness threshold was simply based on the highest ICER per education outcome found in previous economic evaluations in sub-Saharan Africa. Secondly, even if Remme et al. (2014) acknowledged the importance of investigating disparities in health gains between different groups and mentioned the potential use of extended ICERs[55], no attempt was made to incorporate equity objectives in the evaluation. Thirdly, with regard to the cross-temporal aspects of the analysis, long-term benefits and cost were identified, but excluded from the analysis based on the co-financing approach. The authors argued that CEA-based approach focuses on immediate intervention outcomes, and did not discuss relevant time horizons for the analysis or the presence of potential budgetary constraints.

8.4.3.2. Sufficient capabilities approach

Another alternative approach was proposed by Goranitis et al. (2017) in the context of the treatment for drug addiction. The authors developed a methodological case study designed to explore the impact of: changing the evaluative space within an economic evaluation from health to

---

[54] In the bookshelf framework, each book represents an intervention. The height of the book shows its effectiveness, and its thickness depends on its total cost. Books are ranked in order of their height, and investments are undertaken until the health budget is exhausted (Remme, Martinez-Alvarez & Vassall, 2017, Culyer, 2016).

[55] See Chapter 3, Sub-section 3.3.3.2.

capability well-being (i.e. using a broader perspective for the analysis); changing the decision-making rule from health maximisation to the maximisation of sufficient capability (i.e. incorporating alternative value judgments in the decision making process). The case study explored a broader perspective of analysis both in terms of costing and evaluative space. Broader costing was conducted from a government perspective and included health and social care costs, criminal justice, housing (e.g. supported accommodations), private consumption (e.g. state benefits). As regards the evaluative space, capability instruments were used to measure broader outcomes than health alone. The authors investigated also the incorporation of alternative equity concerns by introducing a more egalitarian decision making rule, based on sufficiency instead of maximisation. The 'sufficient capability' approach prioritises individuals below a normatively sufficient level of capability and aims to maximise the number of people in society that achieve sufficient capability.

The analysis showed that different evaluative spaces and decision-making rules have the potential to offer opposing treatment recommendations. Nevertheless, final results of the analysis depend on some unspecified decision maker's WTP per additional year of full capability equivalent (or year of sufficient capability equivalent, if equity concerns are included) that does not refer to opportunity cost. Moreover, because all costs were aggregated and capability was used as the unique outcome measure, the results of such analysis might not reflect the actual decision making processes in the various sectors and so may not meet the decision makers' requirements.

8.4.3.3. Broader costing perspective

Other examples from the literature of analyses conducted from a 'societal' perspective broadly consisted of aggregating costs across sectors and comparing them to an incremental unit of effect (Drost et al., 2016, Byford et al., 2013)

For example, to evaluate an intervention for reducing alcohol use and binge drinking, Drost et al. (2016) calculated the total incremental costs (falling on HC, education, private consumption, housing and CJS) per incremental reduction of one unit of alcohol consumption. Nevertheless, because no guidelines are available that provide a reference cost-effectiveness threshold for reducing the consumption of alcohol, it was not possible to make recommendations on whether the intervention should be implemented. More generally, the issues related to using only one threshold instead of considering different budgets with different opportunity costs have been discussed in Chapter 2, Sub-section 2.4.3.1. Further, similarly to the sufficient capabilities approach previously discussed, a limitation of such an approach is that the analysis risks not meeting decision makers' requirements. In fact, it would be difficult to identify a decision maker that is responsible for an aggregated budget across sectors and a unique measure of outcome.

8.5. Recommendations and further research

In this thesis, instead of using a stylised or fictional example, data from an actual intervention were used. The objective of this choice was to show what data was available and what further evidence may be needed. The development of the case study helped to identify what currently unavailable data should be collected for future studies, and what other wider research (e.g. outcomes of interest to decision makers and marginal productivities) would be valuable.

This section provides an overview of the practical suggestions and recommendations for future work and for future evaluations in this field. The need of improving the collection of cross-sectoral costs and outcomes is described in Sub-section 8.5.1. Sub-section 8.5.2 illustrates potential methods to estimate cross-sectoral marginal productivities, which are currently not available. Because this thesis focused specifically on HC and CJS sectors, this discussion covers these two areas. However, it would be equally important to address the same issues in other public sectors (e.g. education, housing, environment, transport) potentially impacted by public health interventions.

8.5.1. The need of collecting cross-sectoral costs and outcomes

Time and resource constraints often prevent a comprehensive collection of all data required to expand the evaluation perspective (Hill et al., 2017). Nevertheless, as pointed out in Chapter 5, Sub-section 5.8.2, to capture the real world economic consequences of the intervention, a more accurate procedure to record CJ outcomes would be beneficial. The CEA conducted from the CJS perspective requires reliable data about the types of crimes committed in order to associate unit costs accurately. Moreover, the analysis would be more accurate if CJ resource use consumption were collected following each offender, as is commonly done for HC resource use.

The identification and classification of inter-sectoral costs and benefits (ICBs) would offer decision makers supporting information on how to optimally allocate scarce resources (Drost et al., 2013). Reliable and valid instruments to measure ICBs such as CJ outcomes and resources use are pivotal, and lists of resource-use measurement (RUM) instruments in the CJ sector have been recently proposed (Mayer et al., 2017, Drost et al., 2013). Because CJ service resource use (e.g. legal assistance, police custody, court appearance, police contacts, and use of correction services) is particularly important in the field of alcoholism (Mayer et al., 2017), all future trial-based economic evaluations of alcohol interventions should consider the routine adoption of RUM instruments to measure ICBs falling on CJS (Drost et al., 2013).

With regard to cross-sectoral unit costs, the development of cost catalogues including estimates for both HC and non-HC (e.g. the cost of an arrest or a prison sentence) costs was recently suggested (Neumann et al., 2018). The UK has already developed such catalogues for HC costs. Data on CJS unit costs are also available (Dubourg, Hamed & Thorns, 2005, Brand & Price, 2000), but an update of such estimates would be advisable (Bhattacharya, 2017).

8.5.2. The lack of marginal productivity estimates within and across sectors

The lack of estimates of marginal productivities within and across sectors hampers the operationalisation of the framework. Marginal productivities are necessary to assess the opportunity costs of the intervention. To overcome the lack of data in sectors other than health care, in this analysis it was necessary to make an assumption. Following Claxton et al. (2019), it was assumed that the ratio of marginal productivity and consumption value of each outcome was constant across sectors, even though different ratios could be potentially used for different sectors. Such an assumption would be not necessary if estimates of marginal productivities for all sectors (e.g. the CJS) were available.

In Sub-section 8.5.2.1, considerations about the estimation of the marginal productivity of the CJS in preventing crime are presented. Marginal productivities across HC and CJ sectors are discussed in Sub-section 8.5.2.2.

8.5.2.1. Estimate of the marginal productivity of the CJS

To obtain an estimate of the marginal productivity for the CJS, it would be necessary to assess the variations in the CJS budget over time and variations in the outcome(s) chosen to represent the effectiveness of the activity of the CJS in preventing crime.

CJS budget

In the CEA conducted from the CJS perspective, only costs in response to and as a consequence of crime were considered. Nevertheless, a third category of costs was also provided by the Home Office (Dubourg, Hamed & Thorns, 2005, Brand & Price, 2000): the costs in anticipation of crime. The three categories of costs and their components are summarised Table 8.2.

In response to crime	As a consequence to crime	In anticipation of crime
CJS costs (including police)	Consumption	Security expenditure
	Health	Insurance administration

Table 8.2 CJS cost categories

Costs in response to crime were used to estimate the effects of variations in criminal activities in terms of CJ resource use. Costs as a consequence of crime were used to estimate spill-over effects on other sectors (e.g. property stolen and damaged, lost output, tangible and intangible health consequences). If the objective is instead to assess the activity of the CJS in preventing crime, costs in anticipation of crime would be the cost category to take into consideration.

CJS outcomes

Two potential alternatives are suggested to estimate the effect of CJS activity in reducing crime. The first solution is to assess number of crimes and recidivism, using reconvictions as a proxy. This was the method proposed in Chapter 5. Alternatively, as discussed in Sub-section 8.4.1.2, it could be argued that the ultimate goal of the CJS is to increase safety. Therefore, instead of measuring criminal events, a new composite measure based on the sense of safety could be developed to measure the effectiveness of CJS activity.

For example, a year of life adjusted for the perceived sense of safety could be developed. This could be called Safety-Adjusted Life Year (SALY). Similarly to a QALY, a SALY could be obtained by weighting a year for a safety loss coefficients for each crime experienced, directly or indirectly. Safety loss could be related to both the severity of the crime experienced, and the duration of its consequences. The value of a SALY would therefore range from 1 (a year when the individual feels completely safe for the entire time) to 0 (if the whole year is perceived as completely unsafe by the individual). However, such a measure would have potential overlapping dimension with a QALY, and might compromise the applicability of the framework (see Section 8.3). Moreover, to be useful in terms of decision making, such a measure would need to be accepted by decision makers when assessing the effectiveness of a CJ intervention.

Assessing variations in CJS budget and outcomes

Once costs in anticipation of crime and appropriate outcome measures are identified, it is necessary to assess their variation over time to determine the marginal productivity of the CJS.

Variations in criminal activity could be estimated by investigating the yearly statistics about criminal events (Chaplin, Flately & Smith, 2011). By contrast, estimates of the amount of resources invested to prevent crime are only available for the years 2000 (Brand & Price, 2000) and 2003 (Dubourg, Hamed & Thorns, 2005). However, to overcome the lack of data, it could be possible to assume that variations in expenditures follow the trend of HC expenditures, which are available (Lomas, Martin & Claxton, 2018).

Results

The comparison of variations in CJS expenditures in anticipation of crime and variations in criminal activity would provide an estimate of the marginal productivity of the CJS. Such an analysis should be then adjusted for potential secular trends in criminal behaviour (i.e. potential reduction of crime over time not due to differential expenditure in anticipation of crime). Moreover, estimates should be corrected for the endogeneity of spending (i.e. CJS spending is partly determined by the level of needs, which are correlated with criminal justice outcomes).

The adoption of alternative outcome measures for the CJ output would imply alternative solutions for the estimation of its marginal productivity. If the framework proposed in this thesis is used, such an analysis would provide the investment necessary for the CJS to prevent one reconviction. Alternatively, a cost per additional SALY could be estimated.

8.5.2.2. Estimate of the cross-sectoral marginal productivities

CJS activity is associated with spill-over effects on health care and consumption. Cross-sectoral productivity of the CJS can be then quantified in terms of health (e.g. measured in QALYs) and consumption (e.g. property stolen or damaged, lost output).

An estimate of the marginal productivity of the CJS in generating health was already calculated in Chapter 6, Sub-section 6.2.3.3. The inter-sectoral marginal productivity of CJS in producing health was estimated to be approximately equal to 870,000 £/QALY. However, this value could be revised if better evidence about the health consequences of crime were provided, and more accurate data about CJS activity in preventing crime were available. In Sub-section 6.3.3.3, the procedure to estimate the marginal productivity of the CJS in protecting individual consumption was also illustrated. No estimate was computed, but data are potentially available (Dubourg, Hamed & Thorns, 2005).

With regard to the health care sector, estimates for the marginal productivity of the health care system in generating consumption already exist (Claxton, Sculpher, et al., 2015, Roberts, G., 2015). However, there are no estimates of the impact of HC investments on criminal activity.

8.6. Conclusions and policy implications

Even though the importance of economic evaluation to support decision making is widely acknowledged, there is a limited evidence of it being used in practice in decisions related to public health. This may be partly due to economic evidence typically not addressing the priorities of multiple stakeholders and not reflecting local authorities' needs of capturing wider non-health outcomes. Further, economic evaluations should better reflect the financial and political context and take into consideration decision making criteria that incorporate equity (e.g. minimising unfair health inequalities) (Frew & Breheny, 2019).

The framework proposed in this thesis aims to address these issues. It can inform decision makers by offering assessments of the benefits and opportunity costs for those dimensions (other than health care) which are considered most important (Walker et al., 2019). Moreover, the analysis can accommodate different decision-making criteria (based on alternative value judgments), and potentially accommodate political constraints on budgets and timeframes. For these reasons, economic evaluations based on the proposed framework can potentially provide better support for decision making in local authorities.

The proposed framework could be potentially employed for the evaluation of all public health interventions. However, because the case study illustrated in this thesis considered specifically HC and CJ sectors, it would be straightforward to apply the same approach to other public health interventions having impacts on the same dimensions. For example, it could be used to evaluate programmes for the treatment of drug addiction. Such programmes may have the primary objective of enhancing population health, however costs and benefits may fall also on CJS, through less crime attributable to drug use (Goranitis et al., 2017, Sculpher et al., 2014). The framework could be also be used in the evaluation of interventions for the treatment of opiate addiction (Byford et al., 2013).

Recent evidence has shown support from local authorities for methods that incorporate multiple sector costs and benefits (Frew, 2016). However, a cross-sectoral analysis can only show the potential cost-effectiveness of an intervention. To actually implement the intervention, the compensation across-sectors would need to be real (Remme, Martinez-Alvarez & Vassall, 2017). In reality, such a compensation mechanism currently does not exist. Furthermore, if compensation schemes across sectors were created, incentives to prevent gaming of the system would be necessary (Remme et al., 2014). This work however contributes to emphasising the importance of breaking down silos between decision makers' budgets (Butler, 2018, Frew, 2017).

Appendices

A. SIPS population characteristics at baseline, adapted from Newbury-Birch et al. (2014)

	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
Age in years	184	31.6 (10.9)	177	30.8 (10.8)	163	30.5 (11.1)	524	31.0 (10.9)
EQ-5D	178	0.84 (0.24)	171	0.82 (0.22)	157	0.88 (0.21)	506	0.85 (0.23)
Baseline AUDIT score	181	15.40 (8.33)	178	16.29 (8.80)	161	16.58 (8.58)	520	16.07 (8.57)
	N	%	N	%	N	%	N	%
Baseline AUDIT status								
Negative	32	17.7	22	12.4	16	9.9	70	13.5
Positive	149	82.3	156	87.6	145	90.1	450	86.5
AUDIT score category								
Lower risk (0–7)	32	17.7	22	12.4	16	9.9	70	13.5
Increasing risk (8–15)	75	41.4	80	44.9	72	44.7	227	43.7
Harmful/Possible dependence (16–40)	74	40.9	76	42.7	73	45.3	223	42.9
Gender								
Male	160/184	87	145/178	81.5	142/163	87.1	447/525	85.1
Ethnicity								
White	140/184	76.1	138/178	77.5	122/163	74.8	400/525	76.2
Marital status								
Single	125/184	67.9	125/178	70.2	103/163	63.2	353/525	67.2
Education after 16 years								
Yes	84/182	46.2	77/178	43.3	74/163	45.4	235/523	44.9
Possess degree or equivalent								
Yes	25/184	13.6	26/178	14.6	30/161	18.6	81/523	15.5
Smoke tobacco								
Current smoker	150/184	81.5	139/177	78.5	126/162	77.8	415/523	79.3
Readiness Ruler								
Never think about drinking less	39	21.3	35	19.7	43	26.5	117	22.4
Sometimes think about drinking less	37	20.2	37	20.8	34	21	108	20.7
Decided to drink less	39	21.3	33	18.5	18	11.1	90	17.2
Already trying to cut down	68	37.2	73	41	67	41.4	208	39.8

B. EQ-5D-3L questionnaire

Please place a cross in one box in each group.

1. Mobility

I have no problems in walking about ☐

I have some problems in walking about ☐

I am confined to bed ☐

2. Self-care

I have no problems with self-care ☐

I have some problems washing or dressing myself ☐

I am unable to wash or dress myself ☐

3. Usual activities (e.g. work, study, housework, family or leisure activities)

I have no problems with performing my usual activities ☐

I have some problems with performing my usual activities ☐

I am unable to perform my usual activities ☐

4. Pain or discomfort

I have no pain or discomfort ☐

I have moderate pain or discomfort ☐

I have extreme pain or discomfort ☐

5. Anxiety or depression

I am not anxious or depressed ☐

I am moderately anxious or depressed ☐

I am extremely anxious or depressed ☐

C. A focus on the multilevel structure of the data

SIPS study was a cluster trial. As illustrated in Section 3.4.3, the 525 offenders included in the study were recruited by 131 offender managers. For this reason, the impact of the clusters must be taken into account, otherwise results could be biased. The question to be addressed here is whether could be worthwhile to perform the analysis using a multilevel model (Roberts, J. K., 2004).

For the sake of performing a CEA, there is little difference in simply adjusting the standard errors of the parameter estimates taking into account the clusters, or building a multilevel model. However, in the perspective of performing a subsequent multiple imputation procedure, there are some technical differences if data are multilevel or not. Multilevel modelling can be computationally intensive, and in particular performing a multilevel multiple imputation would add a consistent computational hurdle (e.g. Stata software does not support this operation) (Quartagno & Carpenter, 2018).

The following table provides a summary of the clusters (i.e. the groups of offenders recruited by each OM) allocated to each intervention allocation grouped according to their size. The first part of the table refers to the whole population sample, whereas the second group of columns describe the size of the clusters including only those offenders who provided information about QALYs.

Offenders recruited (cluster size)	All patients				Only if reported QALYs			
	Intervention allocation			Total	Intervention allocation			Total
	CIL	BA	BLC		CIL	BA	BLC	
1	15	3	20	38	13	9	19	41
2	8	2	5	15	13	7	7	27
3	4	7	7	18	6	9	5	20
4	6	4	5	15	1	0	3	4
5	9	8	5	22	0	5	1	6
6	1	3	4	8	0	0	1	1
7	1	2	1	4	0	1	0	1
8	0	3	0	3	0	0	0	0
9	0	0	0	0	0	1	0	1
10	0	1	0	1	1	0	1	2
11	0	1	1	2	0	0	0	0
12	0	0	0	0	1	0	0	1
16	1	0	0	1	0	0	0	0
17	0	1	0	1	0	0	0	0
18	1	0	0	1	0	0	0	0
25	1	0	1	2	0	0	0	0
Total	47	35	49	131	35	32	37	104

Considering the whole trial population, a total of 131 clusters were identified. Of these, almost 30% include only one offender, and more than half include less than three offenders. 27 clusters do not actually provide any information about QALYs, because offenders in that group did not report the outcome of interest. Focusing on the clusters where QALYs are reported (104 clusters and 255 offenders), almost 40% are composed of only one individual; roughly 25% of the clusters have two observations, and 20% have three.

It is apparent that it would be difficult to draw conclusions about clusters' impact if the majority of clusters have only three or less observations and the mean cluster size is 2.45 (SD 2.04). A summary showing the proportions of small size samples is reported in below.

	All	**Only if reported QALYs**
Clusters with only 1 observation	29%	39%
Clusters with 2 or less observations	40%	65%
Clusters with 3 or less observations	54%	85%

To formally assess the relatedness of clustered data, the Intracluster Correlation Coefficient (ICC) can be measured. Values of ICC range from zero (no correlation between the observations within clusters) to one (responses within a cluster are identical) (Snijders & Bosker, 2012). A first crude estimation of the ICC was undertaken and then the ICC was estimated again using a regression in order to take into account other factors that might affect the crude estimation. The first crude estimation provided an estimate of 0.13. Given that usually only values smaller than 0.01 are considered negligible (Snijders & Bosker, 2012), it seems that observations do have a degree of correlation within the cluster. However, it must be taken into account that the positive ICC value is not statically significant at 5% level (95% CI goes from 0 to 0.32). Moreover, this was a crude estimation of the ICC, and it is likely to be biased by other factors, such as the baseline health score.

Alternative regression models can be employed to assess the impact of the clusters. Recalling the analysis of QALYs in Section 4.3.2, baseline EQ-5D scores must be included in the regression, whereas other covariates are negligible. Four alternative regression models are then compared:

- Model A: OLS regression where standard errors are assumed to be Normally distributed;
- Model B: OLS regression with robust standard error (called Model 1 in Section 4.3.2, this was the model chosen for the analysis);
- Model C: multilevel regression without robust SE;
- Model D: multilevel regression with robust SE.

As illustrated in the table below, when comparing Model A and Model B, coefficients are identical (as expected), whereas CIs are broader in Model A. Nevertheless, the difference in CIs is marginal. Results from Model C are almost identical to the previous OLS regression models. Multilevel regressions provide also the residual ICC. Once controlled for baseline EQ-5D scores in a multilevel regression, the residual estimated ICC is zero. It seems that there is thus no actual impact of clusters.

OLS regression models						
	Model A: OLS without taking into account clusters			Model B: OLS with SE adjusted for clusters		
Dependent variable: QALYs	Coefficient	95% CI		Coefficient	95% CI	
Intervention allocation:						
Brief Advice	-0.023	-0.056	0.011	-0.023	-0.060	0.014
Brief Lifestyle Counselling	-0.028	-0.062	0.007	-0.028	-0.056	0.001
Baseline EQ-5D	0.675	0.610	0.740	0.675	0.594	0.756
Constant	0.296	0.236	0.356	0.296	0.217	0.375
Multilevel models						
	Model C: Multilevel model			Model D: Multilevel model with SE adjusted for clusters		
Dependent variable: QALYs	Coefficient	95% CI		Coefficient	95% CI	
Intervention allocation:						
Brief Advice	-0.023	-0.056	0.010	-0.023	-0.060	0.014
Brief Lifestyle Counselling	-0.028	-0.062	0.007	-0.028	-0.056	0.001
Baseline EQ-5D	0.675	0.611	0.740	0.675	0.595	0.755
Constant	0.296	0.237	0.355	0.296	0.218	0.374
ICC	0			0		

Given the structure of the clusters and with the support of these results, it appears that the impact of the clusters on QALYs (the main health outcome) is negligible. Moreover, the analysis of SIPS from the health care perspective has mainly the aim of illustrating a method of economic evaluation through a case study and the focus is not on multilevel modelling and multilevel imputation.

D. Assumptions used in analysis of health care resource use data

For each resource category, the resource use questionnaire was composed by two questions: a first generic 'yes-no' question about whether or not the resource was used in the previous six months, and a second question recording the quantity used[56].

If the individual answered 'yes' to the first generic question, then the resource intensity use should also have been recorded. Nevertheless, some individuals answered 'yes' to generic health care resource consumption questions, but did not report a specific value for resource use. Over a total of 18,139 resource use values reported, 18,083 were correctly recorded (i.e. a 'no' in the first question corresponded to a missing or a null resource consumption, or a 'yes' corresponded to a stated resource use), and an additional 56 were missing.

To better illustrate this problem, a practical example of the issue is reported in the table below, which summarises the calls to the drinkline in the first 6 months. Out of three patients who called the drinkline, only one reported the number of contacts.

ID	Intervention	Have you called the drinkline in the past 6 months?	Number of contacts
51345	Minimal	Yes	Missing
52865	Minimal	Yes	14
50744	Brief Advice	Yes	Missing

At this initial stage of the analysis, the aim was to enlarge the complete cases data set on the basis of alternative assumptions on the missing information about resource use when the generic resource use was instead stated. However, it must be stressed that at both month 6 and 12, for most of the health resource use data categories there was no need for imputing any consumption value (e.g. only 0.6%, 0.5% and 0.2% missing values at 6 months were found). Consequently, the assumptions made about the data are unlikely to affect the final results much.

Two alternative scenarios were explored.

- The 'minimum imputation scenario' was based on the assumption that if the generic resource use was stated (i.e. the answer to the question about resource use has been 'yes'),

[56] For questionnaires where some questions were answered and others were not, missing were treated as negatives where questions before and after had 'yes' ticked.

a minimum consumption value of one was certain. For this reason, the 56 missing observations were replaced with the minimum certain consumption (i.e. 1).

- The 'mean imputation scenario' was based on mean imputation. For those individuals who answered 'yes' to the resource use question, but did not report the resource use consumption, the missing resource use was replaced with the mean consumption of the individuals who had a positive resource use in the same group. In cases where the observation was unique and it was therefore not feasible to compute the mean of the group, the minimum certain consumption (i.e. 1) was imputed. Out of 56 cases, 17 were unique observations in their group, and the minimum consumption was imputed.

Going back to the example illustrated in the previous table, in the 'minimum imputation scenario', the missing values for individuals with IDs 51345 and 50744 were replaced with 1 (the minimum certain value of contacts). In the 'mean imputation scenario', the missing value for individual 51345 was replaced with the mean value of their treatment group (i.e. 14). With regard to the missing value for individual 50744, it was not possible to compute the mean value for the BA group, therefore the missing value was replaced with 1 when performing the mean imputation.

The resulting summary of the resources use at month 6 and 12 in the three scenarios (based on original data, mean and minimum imputation) is reported at the end of this section.

Looking at the categories that include at least one imputed value for the resource use, it is clear in most cases there are no considerable differences in the three scenarios. The main differences can be found in counselling both at six and twelve months, and in drinkline calls, detoxification and the overnight residential programme at six months. As regards the calls to the drinkline, these have a really marginal impact on the resource use cost; therefore, there is no real difference in using the scenario based on mean or minimum imputation. Detoxification instead, has a bigger impact on costs, but again the two scenarios are equivalent because the same values are imputed and the resulting resource use is the same. This is because in all these cases the observations were unique and there was no mean consumption to be calculated for the group. On the contrary, looking at the overnight residential programme, resource uses are considerably different between the two scenarios. The same applies to counselling, where the resource use consumption was higher at both 6 and 12 months in the 'mean imputation' scenario in all intervention groups.

The main argument in support of the 'minimum imputation scenario' is that the distributions are positively skewed and therefore mean imputation might potentially give higher weight to outliers. We cannot be certain that people reporting the generic use of the resource without the specific quantity used did actually consume the resource; under the assumption that individuals did actually consume the resources, the minimum consumption would be the only certain value we could impute. However, as a counter-argument, especially as regards counselling and overnight residential programme, it would probably make sense that the resource is used more than once, and it might be more realistic to impute the mean value. If with mean imputation there might be a risk of overestimating the resource use consumption, with minimum imputation there might be instead a risk of underestimating it. Moreover, it must be noticed that there are also individuals that answered ‘yes’, but then explicitly reported zero consumption of the resources; these individuals were included when calculating the average group consumption of the resources, and they might balance the presence of outliers with very high consumption values.

Having established these premises, and keeping in mind the aforementioned almost negligible role of imputed values, the mean imputation scenario was taken as the reference case. All the following analyses will be based on the original dataset adjusted for the imputations under the mean case scenario. As a final note, it must be stressed that the treatment of missing values is not the primary purpose of this work.

Health care resource use at month 6

	CIL			BA			BLC		
	Original	Mean	Minimum	Original	Mean	Minimum	Original	Mean	Minimum
Drinkline calls	0.1 (1.2) $_{119}$	0.2 (1.7) $_{1}$	0.1 (1.2) $_{1}$	0 (0) $_{128}$	0 (0) $_{1}$	0 (0) $_{1}$	0 (0) $_{110}$	0 (0) $_{0}$	0 (0) $_{0}$
Counselling	0.1 (0.8) $_{115}$	0.2 (0.9) $_{5}$	0.1 (0.8) $_{5}$	0.5 (2.7) $_{126}$	0.7 (2.8) $_{3}$	0.5 (2.6) $_{3}$	0.3 (2.3) $_{109}$	0.3 (2.3) $_{1}$	0.3 (2.3) $_{1}$
Day care	0 (0.1) $_{118}$	0 (0.3) $_{2}$	0 (0.2) $_{2}$	0 (0.4) $_{129}$	0 (0.4) $_{0}$	0 (0.4) $_{0}$	0 (0) $_{110}$	0 (0) $_{0}$	0 (0) $_{0}$
Detoxification	0 (0) $_{118}$	0 (0.1) $_{2}$	0 (0.1) $_{2}$	0 (0.5) $_{127}$	0.1 (0.7) $_{2}$	0 (0.5) $_{2}$	0 (0) $_{109}$	0 (0) $_{1}$	0 (0) $_{1}$
Outpatient treatment for drinking problems	0 (0) $_{118}$	0 (0.1) $_{2}$	0 (0.1) $_{2}$	0 (0) $_{128}$	0 (0.1) $_{1}$	0 (0.1) $_{1}$	0 (0.1) $_{110}$	0 (0.1) $_{0}$	0 (0.1) $_{0}$
Overnight hospital detoxification	0 (0) $_{120}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{128}$	0 (0) $_{1}$	0 (0) $_{1}$	0.1 (1.6) $_{110}$	0.1 (1.6) $_{0}$	0.1 (1.6) $_{0}$
Overnight after care hostel	0 (0) $_{120}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{129}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{110}$	0 (0) $_{0}$	0 (0) $_{0}$
Overnight alcohol treatment facility	0 (0) $_{120}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{128}$	0 (0) $_{1}$	0 (0) $_{1}$	0 (0) $_{110}$	0 (0) $_{0}$	0 (0) $_{0}$
Overnight residential programme	0 (0) $_{119}$	0 (0) $_{1}$	0 (0) $_{1}$	1.4 (16) $_{128}$	2.8 (22.5) $_{1}$	1.4 (16) $_{1}$	0 (0) $_{110}$	0 (0) $_{0}$	0 (0) $_{0}$
A&E visits leading to not admitted	0.1 (0.5) $_{119}$	0.1 (0.5) $_{0}$	0.1 (0.5) $_{0}$	0.2 (0.6) $_{128}$	0.2 (0.6) $_{0}$	0.2 (0.6) $_{0}$	0.1 (0.5) $_{109}$	0.1 (0.5) $_{0}$	0.1 (0.5) $_{0}$
A&E visits leading to admitted	0 (0.2) $_{119}$	0 (0.2) $_{0}$	0 (0.2) $_{0}$	0 (0.3) $_{128}$	0 (0.3) $_{0}$	0 (0.3) $_{0}$	0 (0.2) $_{109}$	0 (0.2) $_{0}$	0 (0.2) $_{0}$
Inpatient nights	0 (0.4) $_{120}$	0 (0.4) $_{0}$	0 (0.4) $_{0}$	0 (0.5) $_{129}$	0 (0.5) $_{0}$	0 (0.5) $_{0}$	0.1 (0.8) $_{110}$	0.1 (0.8) $_{0}$	0.1 (0.8) $_{0}$
Outpatient visits	0.2 (1) $_{118}$	0.3 (1) $_{2}$	0.2 (1) $_{2}$	0.4 (1.8) $_{129}$	0.4 (1.8) $_{0}$	0.4 (1.8) $_{0}$	0.2 (0.9) $_{108}$	0.2 (0.9) $_{1}$	0.2 (0.9) $_{1}$
Day case visits	0 (0.1) $_{120}$	0 (0.1) $_{0}$	0 (0.1) $_{0}$	0.1 (1.1) $_{129}$	0.1 (1.1) $_{0}$	0.1 (1.1) $_{0}$	0 (0.1) $_{109}$	0 (0.1) $_{0}$	0 (0.1) $_{0}$
Emergency ambulance travels	0.2 (1.8) $_{120}$	0.2 (1.8) $_{0}$	0.2 (1.8) $_{0}$	0 (0.3) $_{129}$	0 (0.3) $_{0}$	0 (0.3) $_{0}$	0 (0.3) $_{110}$	0 (0.3) $_{0}$	0 (0.3) $_{0}$
Ambulance travels	0 (0) $_{120}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{130}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{109}$	0 (0) $_{0}$	0 (0) $_{0}$
Private travels	0 (0) $_{120}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0.1) $_{130}$	0 (0.1) $_{0}$	0 (0.1) $_{0}$	0.2 (2.8) $_{109}$	0.2 (2.8) $_{0}$	0.2 (2.8) $_{0}$
GP visits	1.4 (3.3) $_{119}$	1.4 (3.3) $_{1}$	1.4 (3.3) $_{1}$	1.9 (3.2) $_{125}$	1.9 (3.2) $_{1}$	1.8 (3.2) $_{1}$	1.4 (2.3) $_{107}$	1.5 (2.3) $_{1}$	1.4 (2.3) $_{1}$
Nurse visits	0.2 (0.8) $_{119}$	0.2 (0.8) $_{1}$	0.2 (0.8) $_{1}$	0.5 (1.6) $_{125}$	0.5 (1.6) $_{1}$	0.5 (1.6) $_{1}$	0.2 (1) $_{107}$	0.2 (1) $_{1}$	0.2 (1) $_{1}$
Social worker visits	0.1 (0.8) $_{120}$	0.1 (0.8) $_{0}$	0.1 (0.8) $_{0}$	0.4 (2.7) $_{128}$	0.4 (2.7) $_{0}$	0.4 (2.7) $_{0}$	0.1 (0.7) $_{109}$	0.1 (0.7) $_{0}$	0.1 (0.7) $_{0}$
Home Visit: GP	0 (0) $_{120}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0.1) $_{128}$	0 (0.1) $_{0}$	0 (0.1) $_{0}$	0 (0.2) $_{109}$	0 (0.2) $_{0}$	0 (0.2) $_{0}$
Home Visit: Nurse	0 (0) $_{120}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{128}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{109}$	0 (0) $_{0}$	0 (0) $_{0}$
Home Visit: Community psychiatric nurse	0 (0) $_{120}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0.6) $_{127}$	0 (0.6) $_{1}$	0 (0.6) $_{1}$	0 (0) $_{109}$	0 (0) $_{0}$	0 (0) $_{0}$
Home Visit: Others	0 (0.2) $_{120}$	0 (0.2) $_{0}$	0 (0.2) $_{0}$	0 (0.2) $_{128}$	0 (0.2) $_{0}$	0 (0.2) $_{0}$	0 (0) $_{109}$	0 (0) $_{0}$	0 (0) $_{0}$
Home Visits: Social worker	0 (0.1) $_{120}$	0 (0.1) $_{0}$	0 (0.1) $_{0}$	0.3 (2.4) $_{127}$	0.3 (2.4) $_{0}$	0.3 (2.4) $_{0}$	0.1 (0.5) $_{108}$	0.1 (0.5) $_{1}$	0.1 (0.5) $_{1}$
NHS Direct events	0 (0.2) $_{119}$	0 (0.2) $_{0}$	0 (0.2) $_{0}$	0 (0.2) $_{127}$	0 (0.2) $_{1}$	0 (0.2) $_{1}$	0 (0.2) $_{109}$	0 (0.2) $_{0}$	0 (0.2) $_{0}$
NHS walk in events	0 (0.2) $_{115}$	0 (0.2) $_{0}$	0 (0.2) $_{0}$	0.2 (0.8) $_{124}$	0.2 (0.9) $_{2}$	0.2 (0.8) $_{2}$	0 (0.3) $_{108}$	0 (0.3) $_{0}$	0 (0.3) $_{0}$

For the 'original' scenario, the table reports: mean (SD) $_{\text{observations}}$

For the 'mean imputation' and 'minimum imputation' scenario, the table reports: mean (SD) $_{\text{number of imputations}}$

Health care resource use at month 12

	CIL			BA			BLC		
	Original	Mean	Minimum	Original	Mean	Minimum	Original	Mean	Minimum
Drinkline calls	0 (0.1)$_{99}$	0 (0.1)$_{0}$	0 (0.1)$_{0}$	0 (0)$_{114}$	0 (0)$_{0}$	0 (0)$_{0}$	0 (0)$_{99}$	0 (0)$_{0}$	0 (0)$_{0}$
Counselling	1 (5.8)$_{98}$	1.1 (5.8)$_{1}$	1 (5.8)$_{1}$	1.4 (11)$_{114}$	1.4 (11)$_{0}$	1.4 (11)$_{0}$	1 (5.2)$_{97}$	1.2 (5.3)$_{2}$	1 (5.1)$_{2}$
Day care	0 (0)$_{98}$	0 (0.1)$_{1}$	0 (0.1)$_{1}$	0 (0)$_{114}$	0 (0)$_{0}$	0 (0)$_{0}$	0.1 (0.8)$_{99}$	0.1 (0.8)$_{0}$	0.1 (0.8)$_{0}$
Detoxification	0 (0)$_{99}$	0 (0)$_{0}$	0 (0)$_{0}$	0 (0)$_{113}$	0 (0)$_{1}$	0 (0)$_{1}$	0 (0.8)$_{99}$	0 (0.8)$_{0}$	0 (0.8)$_{0}$
Outpatient treatment for drinking problems	0 (0)$_{98}$	0 (0.1)$_{1}$	0 (0.1)$_{1}$	0 (0)$_{113}$	0 (0)$_{1}$	0 (0)$_{1}$	0 (0.8)$_{99}$	0 (0.8)$_{0}$	0 (0.8)$_{0}$
Overnight hospital detoxification	0.2 (2.8)$_{100}$	0.2 (2.8)$_{0}$	0.2 (2.8)$_{0}$	0 (0)$_{113}$	0 (0)$_{1}$	0 (0)$_{1}$	0 (0)$_{99}$	0 (0)$_{0}$	0 (0)$_{0}$
Overnight after care hostel	0 (0)$_{100}$	0 (0)$_{0}$	0 (0)$_{0}$	0 (0)$_{114}$	0 (0)$_{0}$	0 (0)$_{0}$	0 (0)$_{99}$	0 (0)$_{0}$	0 (0)$_{0}$
Overnight alcohol treatment facility	0 (0)$_{99}$	0 (0.1)$_{1}$	0 (0.1)$_{1}$	0 (0)$_{114}$	0 (0)$_{0}$	0 (0)$_{0}$	0 (0)$_{99}$	0 (0)$_{0}$	0 (0)$_{0}$
Overnight residential programme	0 (0)$_{100}$	0 (0)$_{0}$	0 (0)$_{0}$	0 (0)$_{114}$	0 (0)$_{0}$	0 (0)$_{0}$	0 (0)$_{99}$	0 (0)$_{0}$	0 (0)$_{0}$
A&E visits leading to not admitted	0.1 (0.4)$_{102}$	0.1 (0.4)$_{0}$	0.1 (0.4)$_{0}$	0.2 (0.6)$_{115}$	0.2 (0.6)$_{0}$	0.2 (0.6)$_{0}$	0.1 (0.4)$_{99}$	0.1 (0.4)$_{0}$	0.1 (0.4)$_{0}$
A&E visits leading to admitted	0 (0.2)$_{102}$	0 (0.2)$_{0}$	0 (0.2)$_{0}$	0 (0.4)$_{115}$	0 (0.4)$_{0}$	0 (0.4)$_{0}$	0 (0.2)$_{99}$	0 (0.2)$_{0}$	0 (0.2)$_{0}$
Inpatient nights	0.3 (2.8)$_{102}$	0.3 (2.8)$_{0}$	0.3 (2.8)$_{0}$	0.7 (4.5)$_{115}$	0.7 (4.5)$_{0}$	0.7 (4.5)$_{0}$	0.1 (0.7)$_{99}$	0.1 (0.7)$_{0}$	0.1 (0.7)$_{0}$
Outpatient visits	0.6 (3.6)$_{101}$	0.7 (3.7)$_{1}$	0.6 (3.6)$_{1}$	0.6 (2.2)$_{114}$	0.7 (2.2)$_{1}$	0.6 (2.2)$_{1}$	0.4 (1.6)$_{99}$	0.4 (1.6)$_{0}$	0.4 (1.6)$_{0}$
Day case visits	0 (0.1)$_{102}$	0 (0.1)$_{0}$	0 (0.1)$_{0}$	0 (0.1)$_{115}$	0 (0.1)$_{0}$	0 (0.1)$_{0}$	0 (0.1)$_{98}$	0 (0.1)$_{1}$	0 (0.1)$_{1}$
Emergency ambulance travels	0 (0.3)$_{102}$	0 (0.3)$_{0}$	0 (0.3)$_{0}$	0.4 (4.6)$_{115}$	0.4 (4.6)$_{0}$	0.4 (4.6)$_{0}$	0.1 (0.5)$_{99}$	0.1 (0.5)$_{0}$	0.1 (0.5)$_{0}$
Ambulance travels	0 (0.1)$_{102}$	0 (0.1)$_{0}$	0 (0.1)$_{0}$	0 (0.2)$_{115}$	0 (0.2)$_{0}$	0 (0.2)$_{0}$	0 (0)$_{99}$	0 (0)$_{0}$	0 (0)$_{0}$
Private travels	0.1 (1.5)$_{102}$	0.1 (1.5)$_{0}$	0.1 (1.5)$_{0}$	0 (0.1)$_{115}$	0 (0.1)$_{0}$	0 (0.1)$_{0}$	0 (0)$_{99}$	0 (0)$_{0}$	0 (0)$_{0}$
GP visits	1.4 (2.9)$_{101}$	1.4 (2.9)$_{1}$	1.4 (2.9)$_{1}$	2.4 (5.7)$_{115}$	2.4 (5.7)$_{0}$	2.4 (5.7)$_{0}$	1.5 (3.2)$_{99}$	1.5 (3.2)$_{0}$	1.5 (3.2)$_{0}$
Nurse visits	0.6 (2.2)$_{101}$	0.7 (2.2)$_{1}$	0.6 (2.2)$_{1}$	0.2 (1.1)$_{115}$	0.2 (1.1)$_{0}$	0.2 (1.1)$_{0}$	0.1 (0.6)$_{99}$	0.1 (0.6)$_{0}$	0.1 (0.6)$_{0}$
Social worker visits	0.2 (1.7)$_{102}$	0.2 (1.7)$_{0}$	0.2 (1.7)$_{0}$	0.2 (1.6)$_{114}$	0.3 (1.7)$_{1}$	0.2 (1.6)$_{1}$	0.7 (7)$_{99}$	0.7 (7)$_{0}$	0.7 (7)$_{0}$
Home Visit: GP	0 (0)$_{102}$	0 (0)$_{0}$	0 (0)$_{0}$	0 (0.3)$_{115}$	0 (0.3)$_{0}$	0 (0.3)$_{0}$	0 (0.3)$_{99}$	0 (0.3)$_{0}$	0 (0.3)$_{0}$
Home Visit: Nurse	0 (0.8)$_{102}$	0 (0.8)$_{0}$	0 (0.8)$_{0}$	0 (0.3)$_{115}$	0 (0.3)$_{0}$	0 (0.3)$_{0}$	0 (0.3)$_{99}$	0 (0.3)$_{0}$	0 (0.3)$_{0}$
Home Visit: Community psychiatric nurse	0 (0.5)$_{102}$	0 (0.5)$_{0}$	0 (0.5)$_{0}$	0.1 (1.2)$_{115}$	0.1 (1.2)$_{0}$	0.1 (1.2)$_{0}$	0 (0)$_{99}$	0 (0)$_{0}$	0 (0)$_{0}$
Home Visit: Others	0 (0)$_{102}$	0 (0)$_{0}$	0 (0)$_{0}$	0 (0)$_{115}$	0 (0)$_{0}$	0 (0)$_{0}$	0 (0)$_{99}$	0 (0)$_{0}$	0 (0)$_{0}$
Home Visits: Social worker	0.1 (0.7)$_{101}$	0.1 (0.7)$_{0}$	0.1 (0.7)$_{0}$	0.3 (1.7)$_{113}$	0.3 (1.9)$_{1}$	0.3 (1.7)$_{1}$	0 (0.4)$_{97}$	0 (0.4)$_{0}$	0 (0.4)$_{0}$
NHS Direct events	0 (0.2)$_{102}$	0 (0.2)$_{0}$	0 (0.2)$_{0}$	0 (0.2)$_{115}$	0 (0.2)$_{0}$	0 (0.2)$_{0}$	0.1 (0.4)$_{97}$	0.1 (0.4)$_{0}$	0.1 (0.4)$_{0}$
NHS walk in events	0 (0.3)$_{102}$	0 (0.3)$_{0}$	0 (0.3)$_{0}$	0.1 (1.2)$_{111}$	0.1 (1.2)$_{0}$	0.1 (1.2)$_{0}$	0 (0.3)$_{98}$	0 (0.3)$_{1}$	0 (0.3)$_{1}$

For the 'original' scenario, the table reports: mean (SD)$_{\text{observations}}$

For the 'mean imputation' and 'minimum imputation' scenario, the table reports: mean (SD)$_{\text{number of imputations}}$

Total health care resource use after 12 months (complete cases)

	Original scenario			Mean imputation			Minimum imputation		
	CIL	BA	BLC	CIL	BA	BLC	CIL	BA	BLC
Drinkline calls	0.17 (1.51) $_{86}$	0 (0) $_{100}$	0 (0) $_{87}$	0.17 (1.51) $_{0}$	0 (0) $_{0}$	0 (0) $_{0}$	0.17 (1.51) $_{0}$	0 (0) $_{0}$	0 (0) $_{0}$
Counselling	0.98 (5.82) $_{83}$	1.98(12.29) $_{100}$	1.54 (6.34) $_{85}$	1.15 (5.81) $_{3}$	1.98 (12.29) $_{0}$	1.78 (6.48) $_{2}$	1.01 (5.72) $_{3}$	1.98 (12.29) $_{0}$	1.54 (6.26) $_{2}$
Day care	0.02 (0.21) $_{84}$	0 (0) $_{100}$	0.11 (0.88) $_{87}$	0.05 (0.32) $_{2}$	0 (0) $_{0}$	0.11 (0.88) $_{0}$	0.04 (0.26) $_{2}$	0 (0) $_{0}$	0.11 (0.88) $_{0}$
Detoxification	0 (0) $_{85}$	0.02 (0.2) $_{99}$	0 (0) $_{86}$	0.01 (0.1) $_{1}$	0.03 (0.22) $_{1}$	0.1 (0.96) $_{1}$	0.01 (0.1) $_{1}$	0.03 (0.22) $_{1}$	0.1 (0.96) $_{1}$
Outpatient treatment for drinking problems	0 (0) $_{84}$	0.01 (0.1) $_{100}$	0.11 (0.86) $_{87}$	0.02 (0.15) $_{2}$	0.01 (0.1) $_{0}$	0.11 (0.86) $_{0}$	0.02 (0.15) $_{2}$	0.01 (0.1) $_{0}$	0.11 (0.86) $_{0}$
Overnight hospital detoxification	0 (0) $_{87}$	0 (0) $_{100}$	0.19 (1.82) $_{87}$	0 (0) $_{0}$	0 (0) $_{0}$	0.19 (1.82) $_{0}$	0 (0) $_{0}$	0 (0) $_{0}$	0.19 (1.82) $_{0}$
Overnight after care hostel	0 (0) $_{87}$	0 (0) $_{100}$	0 (0) $_{87}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{0}$
Overnight alcohol treatment facility	0 (0) $_{86}$	0 (0) $_{100}$	0 (0) $_{87}$	0.01 (0.1) $_{1}$	0 (0) $_{0}$	0 (0) $_{0}$	0.01 (0.1) $_{1}$	0 (0) $_{0}$	0 (0) $_{0}$
Overnight residential programme	0 (0) $_{86}$	0 (0) $_{100}$	0 (0) $_{87}$	0.01 (0.1) $_{1}$	0 (0) $_{0}$	0 (0) $_{0}$	0.01 (0.1) $_{1}$	0 (0) $_{0}$	0 (0) $_{0}$
A&E visits leading to not admitted	0.27 (0.72) $_{88}$	0.51 (1) $_{101}$	0.37 (0.86) $_{86}$	0.27 (0.72) $_{0}$	0.51 (1) $_{0}$	0.37 (0.86) $_{0}$	0.27 (0.72) $_{0}$	0.51 (1) $_{0}$	0.37 (0.86) $_{0}$
A&E visits leading to admitted	0.06 (0.25) $_{88}$	0.09 (0.59) $_{101}$	0.09 (0.42) $_{86}$	0.06 (0.25) $_{0}$	0.09 (0.59) $_{0}$	0.09 (0.42) $_{0}$	0.06 (0.25) $_{0}$	0.09 (0.59) $_{0}$	0.09 (0.42) $_{0}$
Inpatient nights	0.07 (0.4) $_{89}$	0.66 (4.16) $_{102}$	0.26 (1.05) $_{87}$	0.07 (0.4) $_{0}$	0.66 (4.16) $_{0}$	0.26 (1.05) $_{0}$	0.07 (0.4) $_{0}$	0.66 (4.16) $_{0}$	0.26 (1.05) $_{0}$
Outpatient visits	0.95 (4.34) $_{87}$	1.06 (3.22) $_{102}$	0.68 (1.88) $_{85}$	1.06 (4.37) $_{2}$	1.06 (3.22) $_{0}$	0.72 (1.91) $_{1}$	0.97 (4.29) $_{2}$	1.06 (3.22) $_{0}$	0.7 (1.89) $_{1}$
Day case visits	0.03 (0.18) $_{89}$	0.15 (1.2) $_{102}$	0.03 (0.18) $_{85}$	0.03 (0.18) $_{0}$	0.15 (1.2) $_{0}$	0.04 (0.21) $_{1}$	0.03 (0.18) $_{0}$	0.15 (1.2) $_{0}$	0.04 (0.21) $_{1}$
Emergency ambulance travels	0.38 (2.26) $_{89}$	0.08 (0.44) $_{102}$	0.18 (0.82) $_{87}$	0.38 (2.26) $_{0}$	0.08 (0.44) $_{0}$	0.18 (0.82) $_{0}$	0.38 (2.26) $_{0}$	0.08 (0.44) $_{0}$	0.18 (0.82) $_{0}$
Ambulance travels	0.02 (0.21) $_{89}$	0 (0.09) $_{102}$	0 (0) $_{86}$	0.02 (0.21) $_{0}$	0 (0.09) $_{0}$	0 (0) $_{0}$	0.02 (0.21) $_{0}$	0 (0.09) $_{0}$	0 (0) $_{0}$
Private travels	0.17 (1.69) $_{89}$	0.02 (0.16) $_{102}$	0.34 (3.23) $_{86}$	0.17 (1.69) $_{0}$	0.02 (0.16) $_{0}$	0.34 (3.23) $_{0}$	0.17 (1.69) $_{0}$	0.02 (0.16) $_{0}$	0.34 (3.23) $_{0}$
GP visits	2.67 (4.18) $_{87}$	4.44 (8.49) $_{99}$	2.97 (4.65) $_{85}$	2.77 (4.2) $_{2}$	4.43 (8.44) $_{1}$	2.97 (4.62) $_{1}$	2.73 (4.16) $_{2}$	4.41 (8.45) $_{1}$	2.95 (4.62) $_{1}$
Nurse visits	0.94 (2.33) $_{87}$	0.73 (2.02) $_{99}$	0.54 (1.41) $_{85}$	1.05 (2.52) $_{2}$	0.73 (2.01) $_{1}$	0.54 (1.4) $_{1}$	1.05 (2.54) $_{2}$	0.74 (2.01) $_{1}$	0.54 (1.4) $_{1}$
Social worker visits	0.47 (2.35) $_{89}$	0.75 (4.21) $_{100}$	0.2 (0.94) $_{86}$	0.47 (2.35) $_{0}$	0.8 (4.22) $_{1}$	0.2 (0.94) $_{0}$	0.47 (2.35) $_{0}$	0.75 (4.18) $_{1}$	0.2 (0.94) $_{0}$
Home Visit: GP	0.01 (0.1) $_{89}$	0.06 (0.43) $_{101}$	0.06 (0.45) $_{86}$	0.01 (0.1) $_{0}$	0.06 (0.43) $_{0}$	0.06 (0.45) $_{0}$	0.01 (0.1) $_{0}$	0.06 (0.43) $_{0}$	0.06 (0.45) $_{0}$
Home Visit: Nurse	0.1 (0.95) $_{89}$	0.04 (0.4) $_{101}$	0.03 (0.32) $_{86}$	0.1 (0.95) $_{0}$	0.04 (0.4) $_{0}$	0.03 (0.32) $_{0}$	0.1 (0.95) $_{0}$	0.04 (0.4) $_{0}$	0.03 (0.32) $_{0}$
Home Visit: Community psychiatric nurse	0 (0) $_{89}$	0.23 (1.47) $_{100}$	0.01 (0.1) $_{86}$	0 (0) $_{0}$	0.25 (1.48) $_{1}$	0.01 (0.1) $_{0}$	0 (0) $_{0}$	0.23 (1.47) $_{1}$	0.01 (0.1) $_{0}$
Home Visit: Others	0.03 (0.31) $_{89}$	0.02 (0.29) $_{101}$	0 (0) $_{86}$	0.03 (0.31) $_{0}$	0.02 (0.29) $_{0}$	0 (0) $_{0}$	0.03 (0.31) $_{0}$	0.02 (0.29) $_{0}$	0 (0) $_{0}$
Home Visits: Social worker	0.15 (0.85) $_{88}$	0.84 (3.82) $_{99}$	0.2 (0.76) $_{83}$	0.15 (0.85) $_{0}$	0.93 (3.88) $_{1}$	0.23 (0.8) $_{1}$	0.15 (0.85) $_{0}$	0.85 (3.8) $_{1}$	0.21 (0.76) $_{1}$
NHS Direct events	0.11 (0.41) $_{88}$	0.11 (0.44) $_{100}$	0.1 (0.49) $_{84}$	0.11 (0.41) $_{0}$	0.11 (0.45) $_{1}$	0.1 (0.49) $_{0}$	0.11 (0.41) $_{0}$	0.11 (0.45) $_{1}$	0.1 (0.49) $_{0}$
NHS walk in events	0.12 (0.37) $_{85}$	0.38 (1.08) $_{93}$	0.15 (0.56) $_{85}$	0.12 (0.37) $_{0}$	0.42 (1.1) $_{2}$	0.16 (0.57) $_{1}$	0.12 (0.37) $_{0}$	0.4 (1.07) $_{2}$	0.16 (0.57) $_{1}$

For the 'original' scenario based on assumptions, the table reports: mean (SD) $_{\text{observations}}$

For the 'mean imputation' and 'minimum imputation' scenario, the table reports: mean (SD) $_{\text{number of imputations}}$

E. SIPS cost categories and budgets

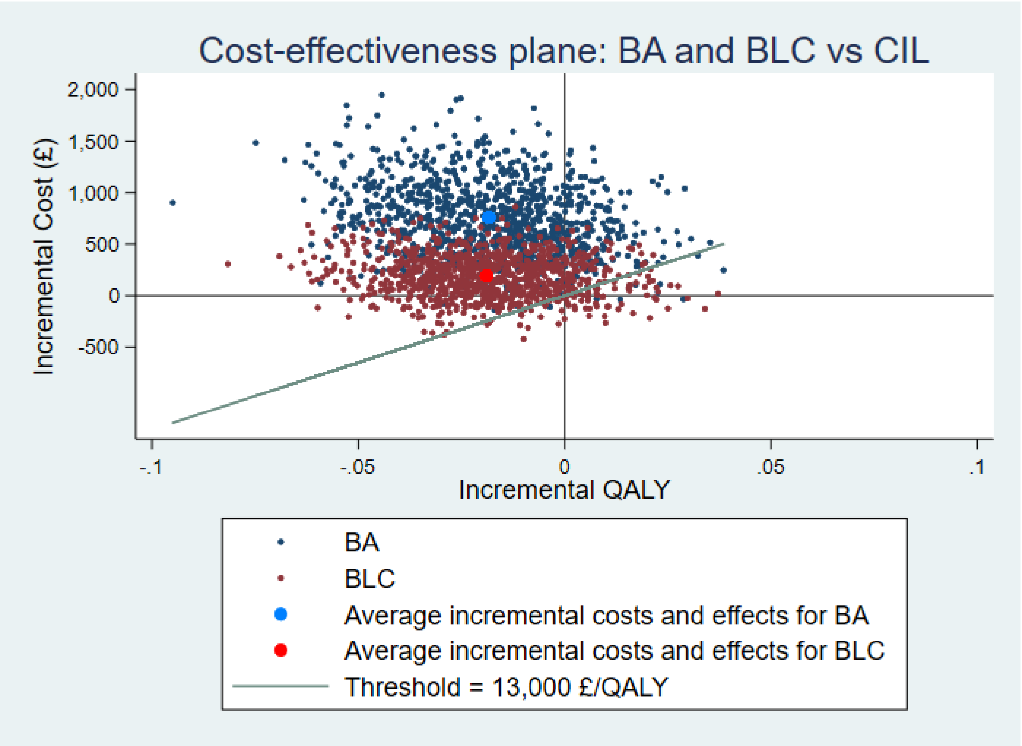
Cost category	Primary care	Secondary care	Non-residential alcohol treatment	Residential alcohol treatment	Social care	Other home visits	Ambulances	Private consumption	CJS
Intervention costs									X
Drinkline calls			X						
Counselling			X						
Day care			X						
Detoxification			X						
Outpatient treatment for drinking problems			X						
Overnight hospital detoxification				X					
Overnight after care hostel				X					
Overnight alcohol treatment facility				X					
Overnight residential programme				X					
A&E visits leading to not admitted		X							
A&E visits leading to admitted		X							
Inpatient nights		X							
Outpatient visits		X							
Day case visits		X							
Emergency ambulance travels							X		
Ambulance travels							X		
Private travels								X	
GP visits	X								
Nurse visits	X								
Social worker visits					X				
Home Visit: GP	X								
Home Visit: Nurse	X								
Home Visit: Community psychiatric nurse						X			
Home Visit: Others						X			
Home Visits: Social worker					X				
NHS Direct events		X							
NHS walk in events		X							

F. CEA results based on SURs without adjusting for baseline costs

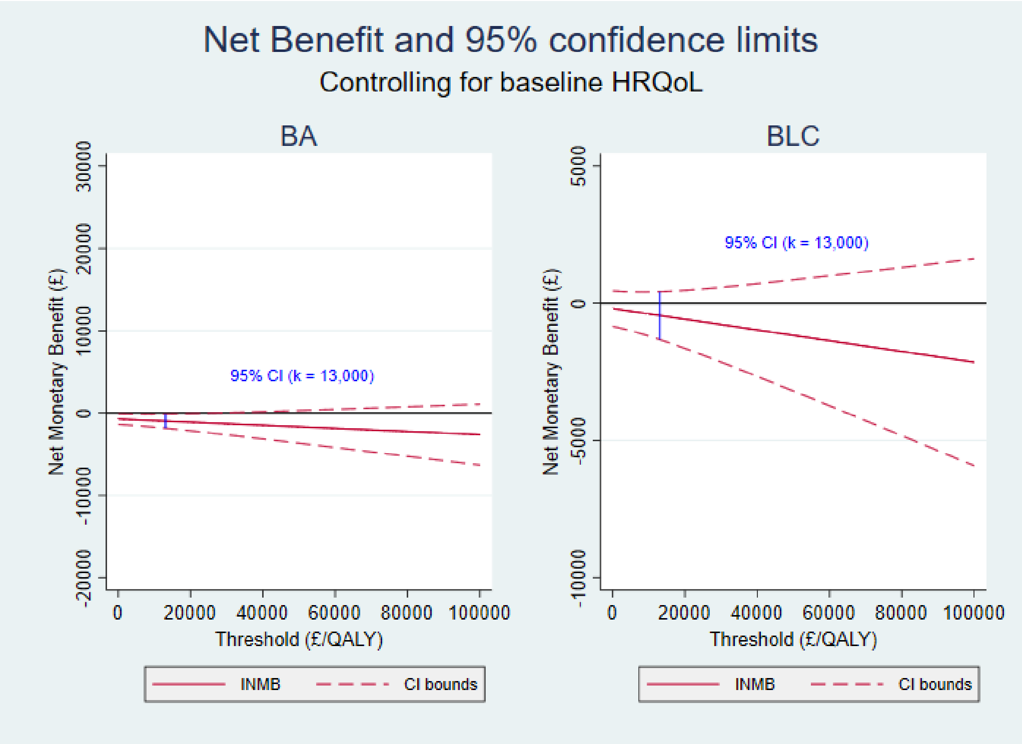
ICERs and INMBs

	BA vs CIL	BLC vs CIL
ICER *(£/QALY)*	-39492	-9966
INMB, k = £13,000 per QALY *(£)*	-987	-449
INMB, k = £20,000 per QALY *(£)*	-1119	-586
INMB, k = £30,000 per QALY *(£)*	-1307	-782

CE plane

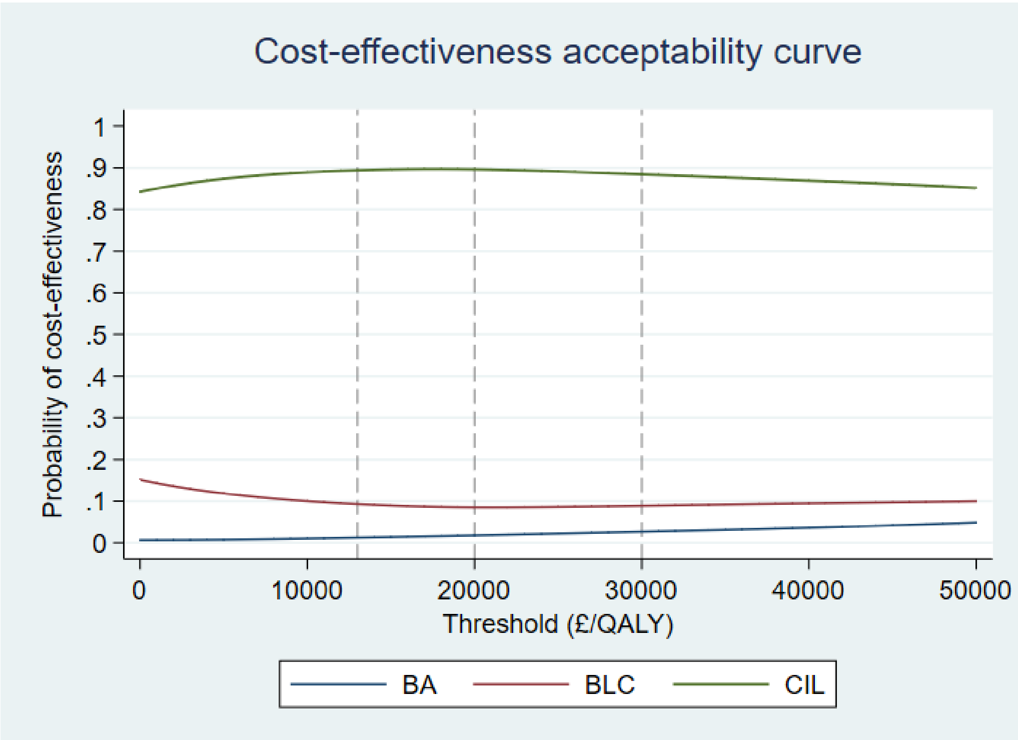


INMBs



Net Benefit and 95% confidence limits
Controlling for baseline HRQoL
BA
BLC
95% CI (k = 13,000)
95% CI (k = 13,000)
Net Monetary Benefit (£)
Net Monetary Benefit (£)
-20000
-10000
0
10000
20000
30000
-10000
-5000
0
5000
0
20000
40000
60000
80000
100000
0
20000
40000
60000
80000
100000
Threshold (£/QALY)
Threshold (£/QALY)
INMB
CI bounds
INMB
CI bounds

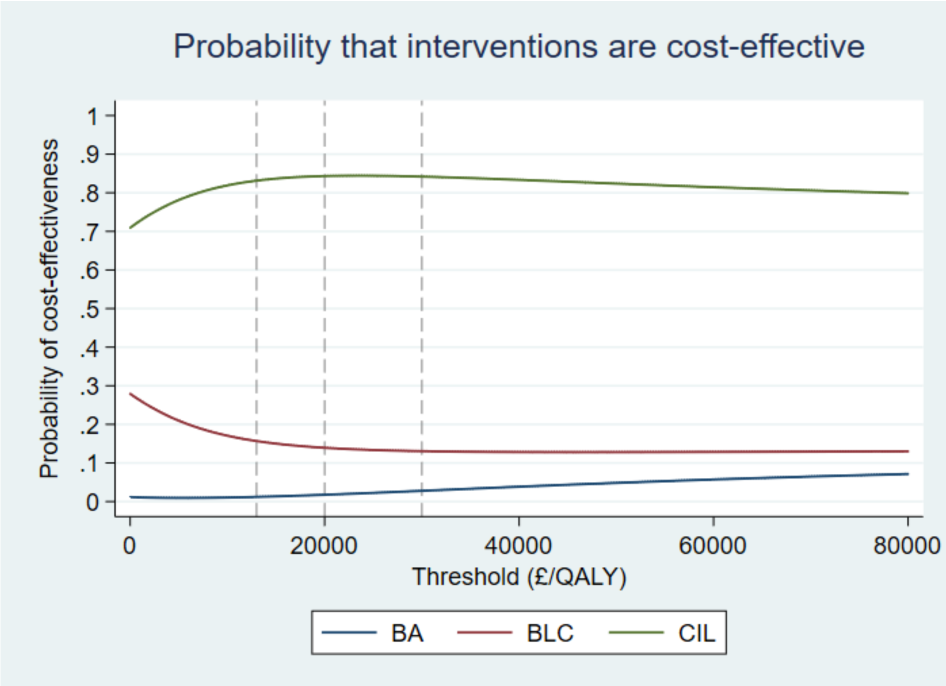
CEAC



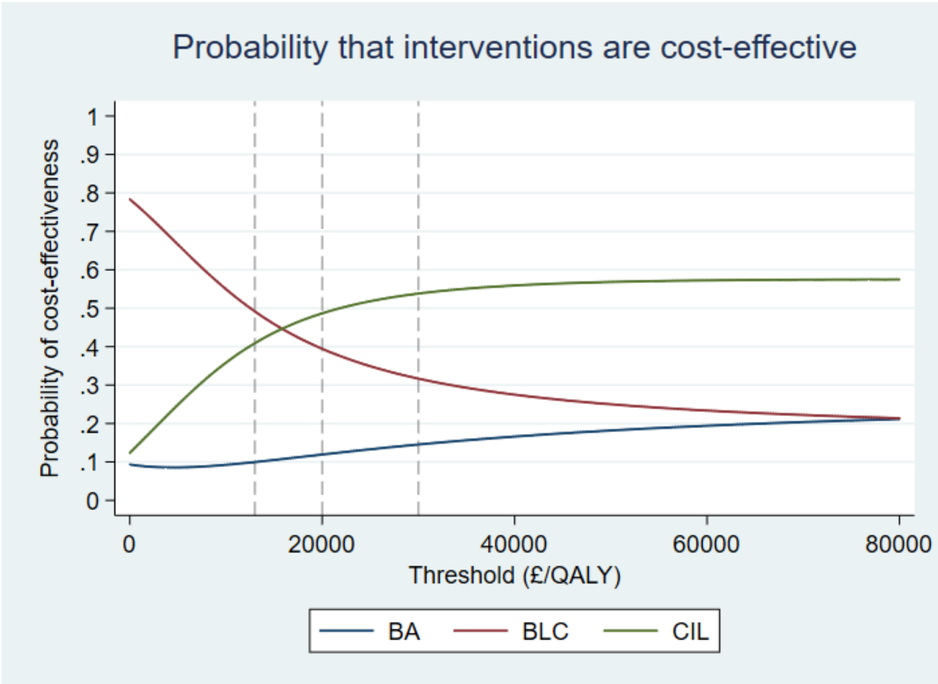
Cost-effectiveness acceptability curve
Probability of cost-effectiveness
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1
0
10000
20000
30000
40000
50000
Threshold (£/QALY)
BA
BLC
CIL

G. Parametric CEACs

Parametric CEACs without adjusting for baseline health care costs



Parametric CEACs adjusting for baseline health care costs



H. Multiple imputation

Proportions of missing data for all health care related variables included in the analysis

Description	Baseline Missing values, %				6 months Missing values, %				12 months Missing values, %			
	Total	CIL	BA	BLC	Total	CIL	BA	BLC	Total	CIL	BA	BLC
Age in years	0	0	1	0								
Gender	0	0	0	0								
Ethnicity	0	0	0	0								
Marital Status	0	0	0	0								
Education after 16 years	0	1	0	0								
Possess degree or equivalent	0	0	0	1								
Smoke tobacco	0	0	1	1								
Location reference (cluster)	0	0	0	0								
Intervention allocation	0	0	0	0								
Baseline and outcome variables for health-related quality of life												
EQ-5D index score	4	3	4	4	34	37	29	36	41	45	37	40
Total QALYs over 1 year*									51	55	49	50
Other baseline and outcome variables												
AUDIT score	1	2	0	1	33	36	29	34	41	45	38	40
Readiness to change	0	1	0	1	33	37	28	35	42	45	37	42
APQ score					31	35	26	33	40	44	35	39
Outcome variables for costs												
Drinkline calls					32	35	28	33	41	46	36	39
Counselling					32	35	28	33	41	46	36	39
Day care					32	35	28	33	41	46	36	39
Detoxification					32	35	28	33	41	46	36	39
Outpatient treatment for drinking problems					32	35	28	33	41	46	36	39
Overnight hospital detoxification					32	35	28	33	40	46	36	39
Overnight after care hostel					32	35	28	33	40	46	36	39
Overnight alcohol treatment facility					32	35	28	33	40	46	36	39
Overnight residential programme					32	35	28	33	40	46	36	39
A&E visits leading to not admitted					32	35	28	33	40	45	35	39
A&E visits leading to admitted					32	35	28	33	40	45	35	39
Inpatient nights					32	35	28	33	40	45	35	39
Outpatient visits					32	35	28	33	40	45	35	39
Day case visits					32	35	28	33	40	45	35	39
Emergency ambulance travels					32	35	28	33	40	45	35	39
Ambulance travels					32	35	27	33	40	45	35	39
Private travels					32	35	27	33	40	45	35	39
GP visits					33	35	29	34	40	45	35	39
Nurse visits					33	35	29	34	40	45	35	39
Social worker visits					32	35	28	33	40	45	35	39
Home Visit: GP					32	35	28	33	40	45	35	39
Home Visit: Nurse					32	35	28	33	40	45	35	39
Home Visit: Community psychiatric nurse					32	35	28	33	40	45	35	39
Home Visit: Others					32	35	28	33	40	45	35	39
Home Visits: Social worker					32	35	29	33	41	45	36	40
NHS Direct events					32	35	28	33	40	45	35	40
NHS walk in events					34	38	29	34	41	45	38	39
Intervention costs									0	0	0	0
Total costs over 1 year*									52	56	49	50

* Total QALYs and total costs over 1 year refer to the sum of QALYs and costs over the individuals with complete data for the relevant variables (EQ-5D for QALYs and cost components for costs).

Associations of health outcomes missingness with baseline variables and observed outcomes

	EQ-5D month 6			EQ-5D month 12			EQ-5D month 12		
	Coeff.	95% CI		Coeff.	95% CI		Coeff.	95% CI	
Baseline EQ-5D	0.339	-0.659	1.337	0.843	-0.123	1.808	0.985	-0.644	2.613
Age	0.025	0.002	0.048	0.039	0.017	0.062	0.044	0.008	0.079
Gender									
Male	-0.417	-1.013	0.180	-0.501	-1.069	0.068	-0.039	-0.884	0.807
Ethnicity									
Black	-0.006	-0.591	0.580	-0.385	-0.941	0.171	-0.651	-1.469	0.167
Asian	-0.029	-1.305	1.247	-0.980	-2.196	0.236	-2.462	-4.314	-0.611
Mixed	0.017	-1.131	1.164	-0.815	-1.907	0.277	-2.018	-3.505	-0.531
Other	-0.002	-1.855	1.852	-0.033	-1.925	1.860	-1.179	-3.713	1.354
Marital Status									
Living with partner	0.202	-0.368	0.772	-0.112	-0.656	0.433	-0.014	-0.884	0.857
Married	-0.709	-1.560	0.142	0.151	-0.751	1.054	1.314	-1.034	3.662
Separated	0.038	-0.865	0.942	-0.509	-1.348	0.329	-0.720	-1.950	0.510
Divorced	1.401	-0.139	2.941	0.001	-1.089	1.090	-1.014	-2.353	0.325
Widowed	0.000			0.000			0.000		
Education after 16 years									
Yes	-0.111	-0.571	0.348	-0.161	-0.604	0.281	-0.014	-0.712	0.684
Possess degree or equivalent									
Yes	0.261	-0.365	0.888	0.550	-0.060	1.160	0.559	-0.396	1.514
Smoke tobacco									
Ex-smoker	0.509	-0.258	1.276	0.182	-0.533	0.897	0.110	-1.013	1.233
Never smoked	0.582	-0.085	1.248	0.219	-0.398	0.835	0.191	-0.716	1.098
Baseline AUDIT score	-0.026	-0.051	0.000	-0.016	-0.041	0.009	0.010	-0.039	0.060
Baseline RCR									
Sometimes think about drink..	0.392	-0.214	0.999	0.198	-0.383	0.779	-0.620	-1.544	0.305
I have decided to drink less	0.331	-0.302	0.964	0.292	-0.317	0.900	-0.121	-1.139	0.897
Already trying to cut down	0.393	-0.137	0.922	0.657	0.144	1.171	0.156	-0.737	1.048
Month 6 EQ-5D							-1.203	-3.096	0.691
Month 6 AUDIT score							-0.046	-0.103	0.011
Month 6 APQ score							-0.016	-0.108	0.076
Month 6 RCR									
Sometimes think about drink..							0.888	-0.121	1.896
I have decided to drink less							0.581	-0.387	1.549
Already trying to cut down							0.948	0.056	1.839
Constant	0.002	-1.467	1.472	-1.085	-2.508	0.338	0.418	-2.149	2.985

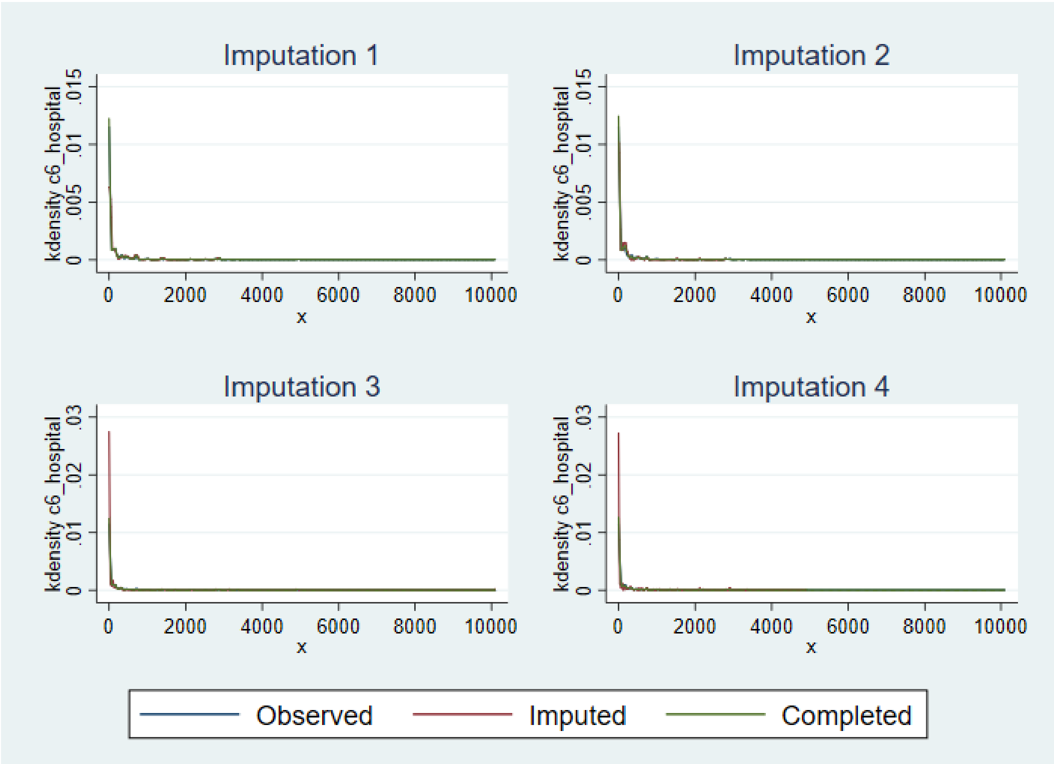
Associations of health care costs missingness with baseline variables

	Hospital and AED			Visits and home visits			Treatment for drinking problems		
	Coeff.	95% CI		Coeff.	95% CI		Coeff.	95% CI	
Baseline EQ-5D	0.689	-0.287	1.665	0.689	-0.287	1.665	0.653	-0.324	1.630
Age	0.041	0.019	0.064	0.041	0.019	0.064	0.043	0.020	0.066
Gender									
Male	-0.588	-1.169	-0.007	-0.588	-1.169	-0.007	-0.605	-1.185	-0.025
Ethnicity									
Black	-0.399	-0.960	0.161	-0.399	-0.960	0.161	-0.377	-0.937	0.184
Asian	-1.036	-2.254	0.181	-1.036	-2.254	0.181	-1.016	-2.235	0.202
Mixed	-0.909	-2.006	0.188	-0.909	-2.006	0.188	-0.873	-1.968	0.223
Other	-0.093	-1.989	1.803	-0.093	-1.989	1.803	-0.065	-1.959	1.828
Marital Status									
Living with partner	-0.131	-0.680	0.419	-0.131	-0.680	0.419	-0.110	-0.659	0.438
Married	0.076	-0.834	0.987	0.076	-0.834	0.987	0.081	-0.830	0.991
Separated	-0.621	-1.465	0.223	-0.621	-1.465	0.223	-0.606	-1.450	0.238
Divorced	-0.118	-1.216	0.980	-0.118	-1.216	0.980	-0.127	-1.225	0.970
Widowed	0.000			0.000			0.000		
Education after 16 years									
Yes	-0.141	-0.587	0.305	-0.141	-0.587	0.305	-0.124	-0.570	0.321
Possess degree or equivalent									
Yes	0.565	-0.054	1.183	0.565	-0.054	1.183	0.490	-0.123	1.104
Smoke tobacco									
Ex-smoker	0.124	-0.595	0.843	0.124	-0.595	0.843	0.149	-0.569	0.867
Never smoked	0.253	-0.372	0.879	0.253	-0.372	0.879	0.284	-0.341	0.910
Baseline AUDIT score	-0.017	-0.042	0.009	-0.017	-0.042	0.009	-0.015	-0.041	0.010
Baseline RCR									
Sometimes think about drink..	0.222	-0.361	0.805	0.222	-0.361	0.805	0.171	-0.411	0.753
I have decided to drink less	0.392	-0.221	1.005	0.392	-0.221	1.005	0.380	-0.233	0.993
Already trying to cut down	0.729	0.212	1.246	0.729	0.212	1.246	0.691	0.175	1.207
Constant	-0.911	-2.352	0.529	-0.911	-2.352	0.529	-0.942	-2.383	0.500

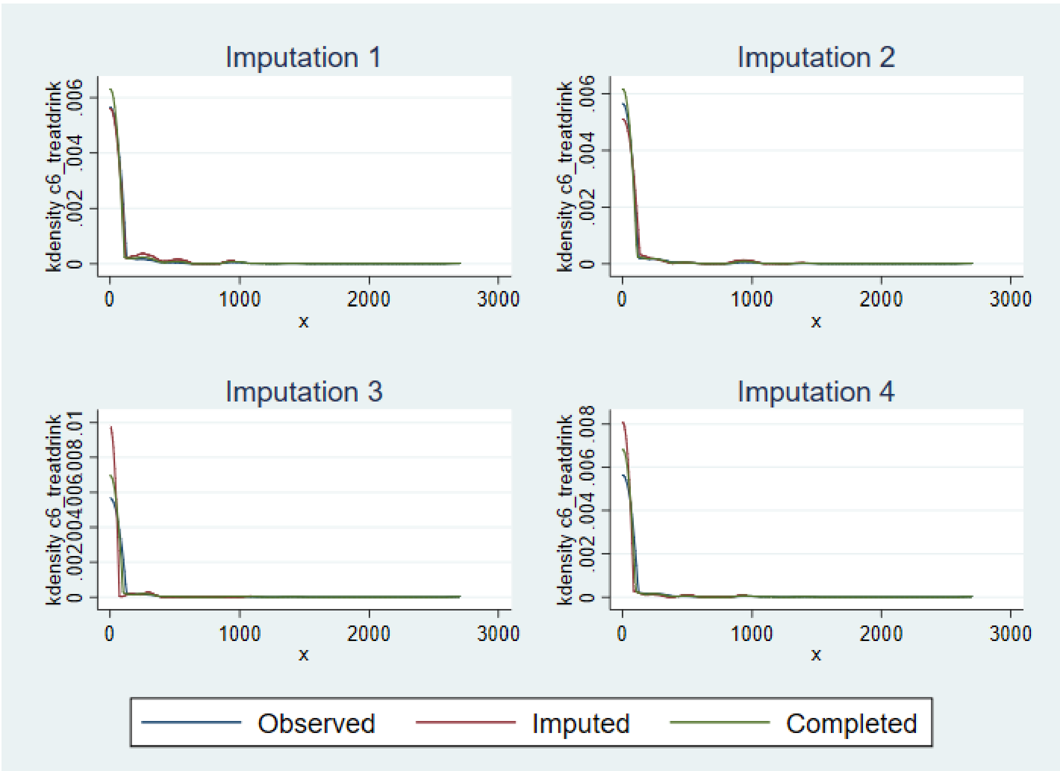
Associations of health outcomes and health care costs missingness with baseline variables and observed outcomes

	Hospital and AED			Visits and home visits			Treatment for drinking problems			EQ-5D month 12		
	Coeff.	95% CI		Coeff.	95% CI		Coeff.	95% CI		Coeff.	95% CI	
Baseline EQ-5D	0.226	-1.378	1.831	0.226	-1.378	1.831	0.084	-1.522	1.689	4.391	-1.600	10.382
Age	0.056	0.021	0.090	0.056	0.021	0.090	0.059	0.025	0.094	0.026	-0.119	0.172
Gender												
Male	-0.475	-1.330	0.379	-0.475	-1.330	0.379	-0.511	-1.359	0.337	1.502	-1.087	4.090
Ethnicity												
Black	-0.510	-1.313	0.293	-0.510	-1.313	0.293	-0.454	-1.253	0.346	0.135	-4.191	4.460
Asian	-1.958	-3.557	-0.359	-1.958	-3.557	-0.359	-1.906	-3.511	-0.302	0.000		
Mixed	-1.682	-3.007	-0.356	-1.682	-3.007	-0.356	-1.585	-2.911	-0.260	0.000		
Other	-0.872	-3.313	1.569	-0.872	-3.313	1.569	-0.815	-3.248	1.618	0.000		
Marital Status												
Living with partner	-0.428	-1.211	0.356	-0.428	-1.211	0.356	-0.375	-1.154	0.404	0.094	-3.212	3.401
Married	1.015	-0.802	2.833	1.015	-0.802	2.833	1.020	-0.786	2.826	0.000		
Separated	-1.047	-2.190	0.097	-1.047	-2.190	0.097	-1.021	-2.163	0.122	0.000		
Divorced	-1.289	-2.592	0.015	-1.289	-2.592	0.015	-1.285	-2.588	0.017	0.000		
Widowed	0.000			0.000			0.000			0.000		
Education after 16 years												
Yes	-0.154	-0.802	0.494	-0.154	-0.802	0.494	-0.113	-0.757	0.531	-0.136	-2.814	2.542
Possess degree or equivalent												
Yes	0.555	-0.335	1.445	0.555	-0.335	1.445	0.395	-0.469	1.260	2.874	-3.433	9.180
Smoke tobacco												
Ex-smoker	-0.158	-1.150	0.834	-0.158	-1.150	0.834	-0.111	-1.100	0.878	0.000		
Never smoked	-0.030	-0.875	0.815	-0.030	-0.875	0.815	0.026	-0.817	0.868	-0.780	-4.086	2.526
Baseline AUDIT score	-0.003	-0.042	0.037	-0.003	-0.042	0.037	-0.001	-0.040	0.039	-0.040	-0.200	0.120
Baseline RCR												
Sometimes think about drink..	-0.206	-1.023	0.611	-0.206	-1.023	0.611	-0.305	-1.117	0.506	1.956	-2.234	6.145
I have decided to drink less	0.549	-0.395	1.493	0.549	-0.395	1.493	0.539	-0.403	1.481	0.736	-2.163	3.634
Already trying to cut down	0.651	-0.123	1.426	0.651	-0.123	1.426	0.575	-0.193	1.343	0.000		
Costs at 6 months												
Hospital and AED	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.002	-0.003	0.000
Visits and home visits	0.001	-0.001	0.002	0.001	-0.001	0.002	0.001	-0.001	0.002	-0.008	-0.014	-0.001
Treatment for drinking problems	-0.002	-0.003	0.000	-0.002	-0.003	0.000	-0.002	-0.003	0.000	0.026	-0.003	0.055
Costs at 12 months												
Hospital and AED										0.000	-0.002	0.002
Visits and home visits										0.006	0.000	0.011
Treatment for drinking problems										0.000	-0.002	0.002
Constant	-0.087	-2.270	2.097	-0.087	-2.270	2.097	-0.075	-2.259	2.109	-0.802	-9.259	7.654

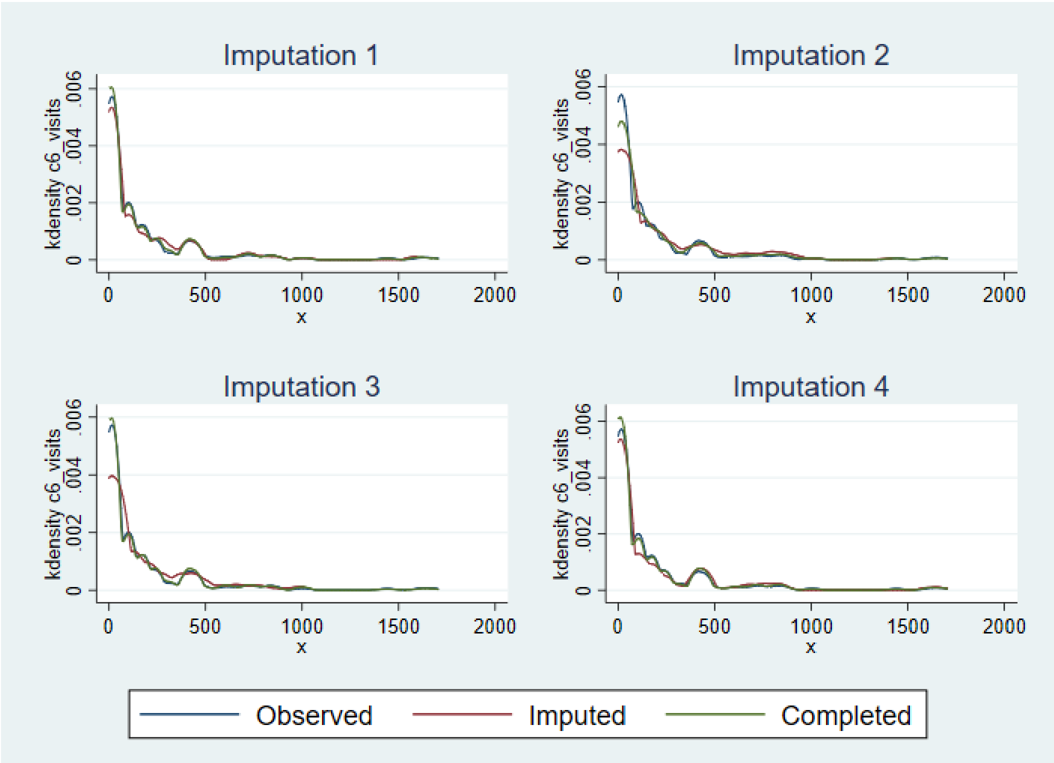
Post-imputation diagnostics for hospital and AED costs at 6 months



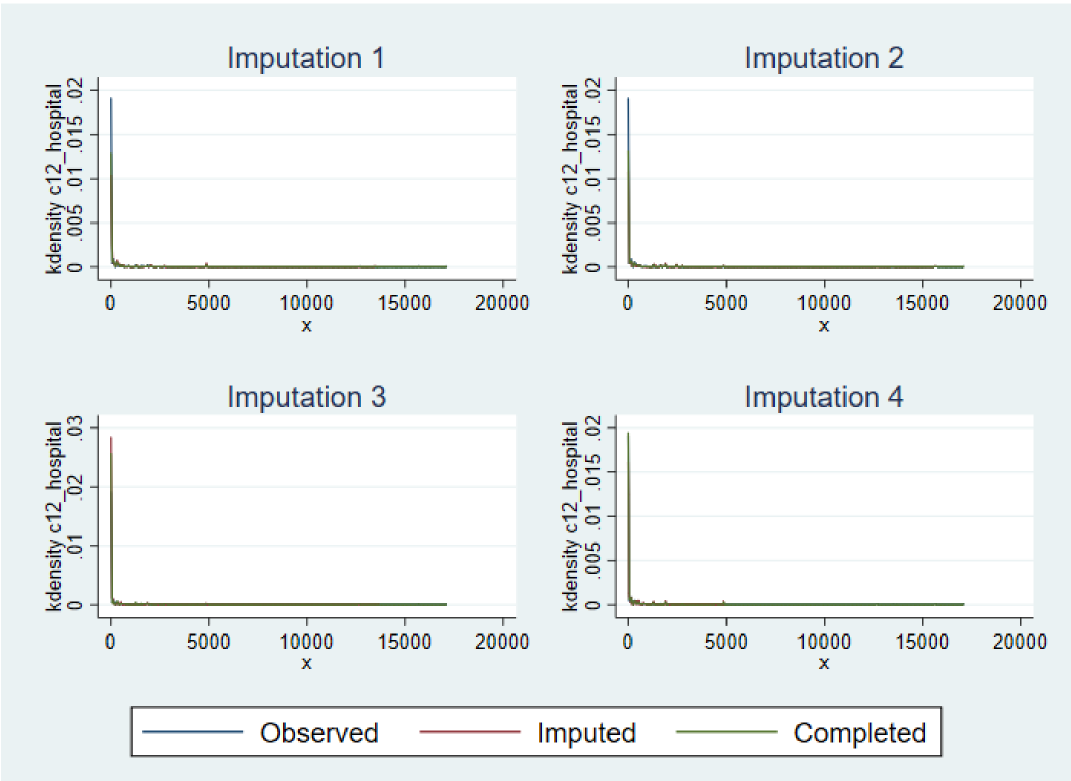
Post-imputation diagnostics for treatment for drinking problems costs at 6 months



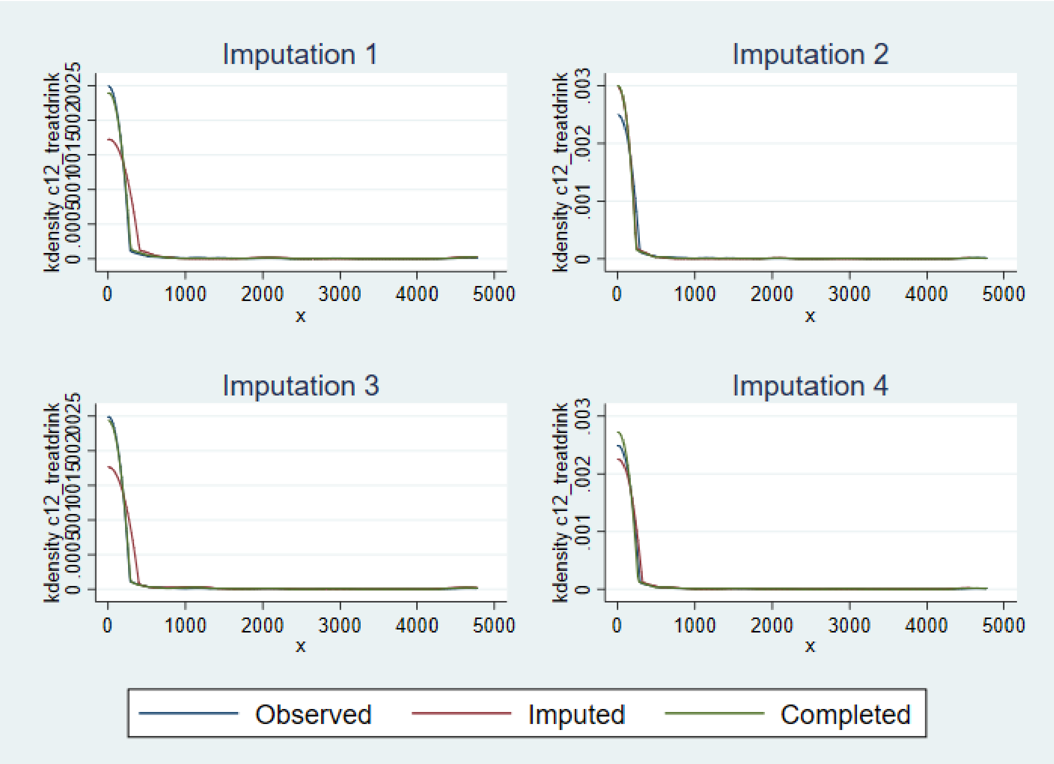
Post-imputation diagnostics for visits and home visits costs at 6 months



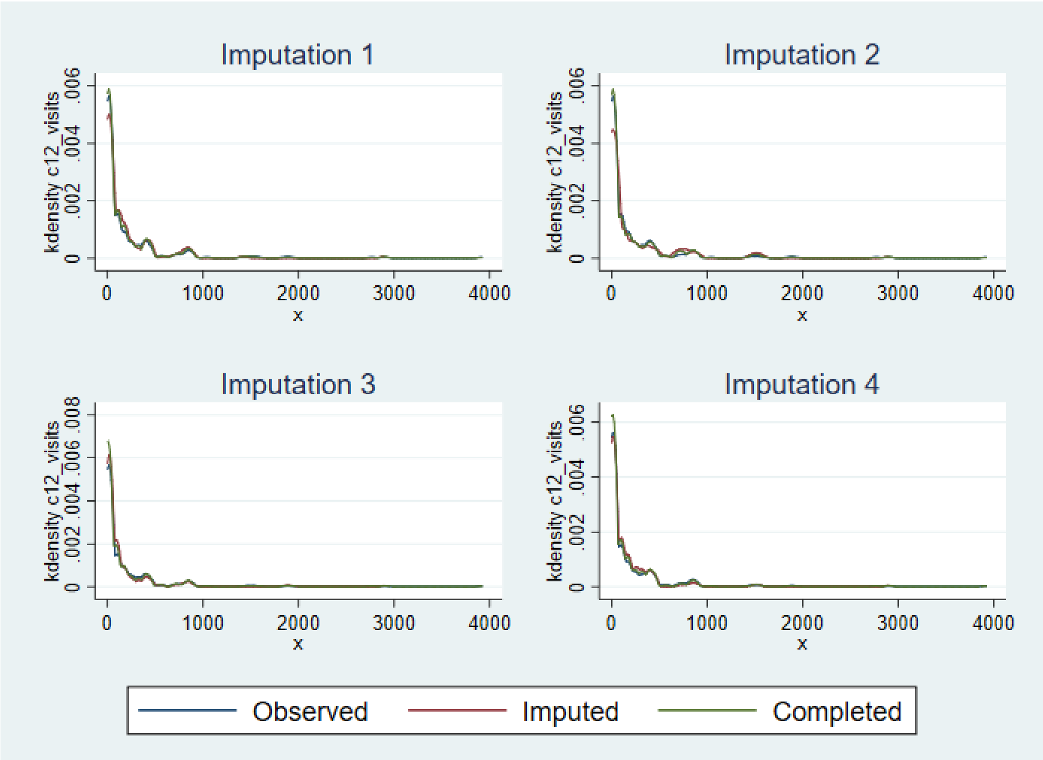
Post-imputation diagnostics for hospital and AED costs at 12 months



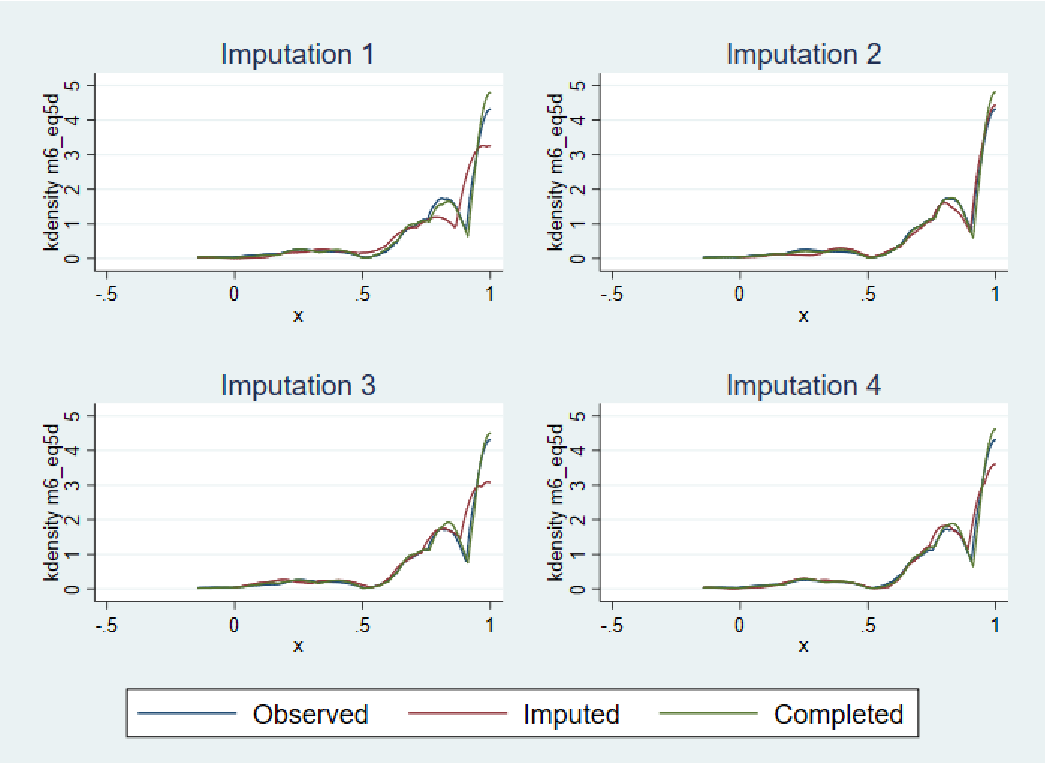
Post-imputation diagnostics for treatment for drinking problems costs at 12 months



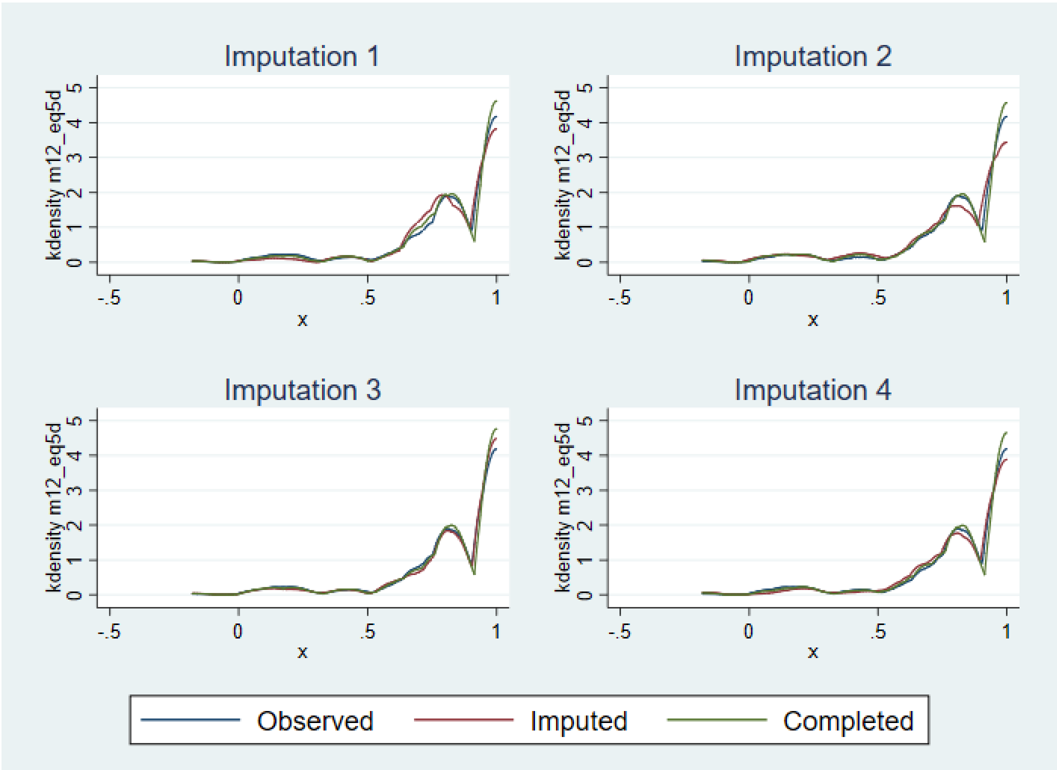
Post-imputation diagnostics for visits and home visits costs at 12 months



Post-imputation diagnostics for EQ-5D scores at 6 months



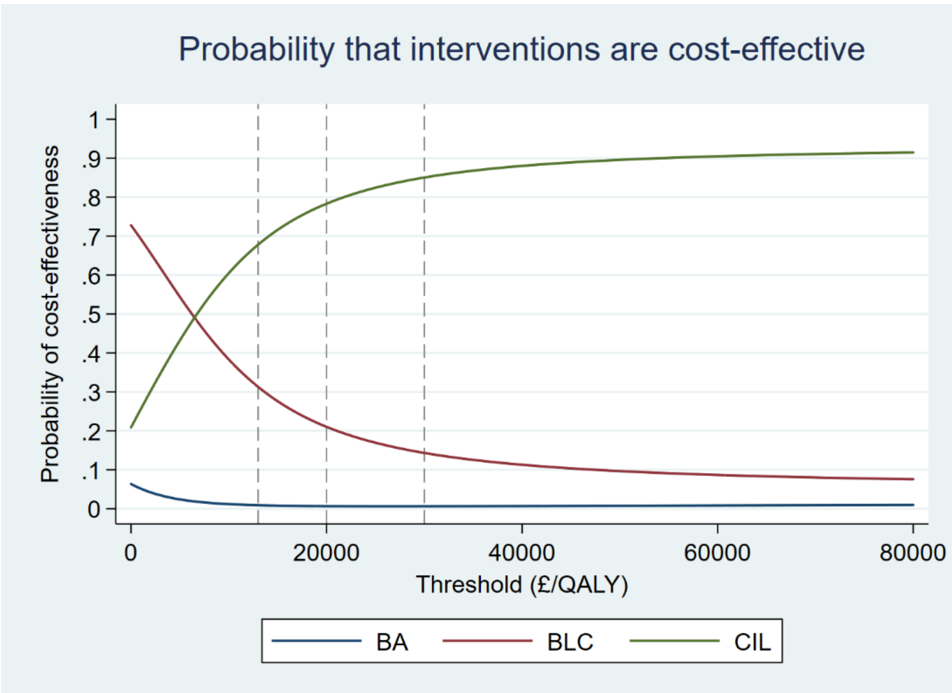
Post-imputation diagnostics for EQ-5D scores at 12 months



Results of the alternative CEA based on SUR without baseline costs adjustment

	QALYs			Costs		
	Coeff.	95% CI		Coeff.	95% CI	
Intervention allocation:						
BA	-0.031	-0.061	-0.002	427.905	-122.943	978.754
BLC	-0.025	-0.055	0.005	-154.452	-653.863	344.958
Baseline EQ-5D	0.584	0.518	0.650			
Constant	0.376	0.315	0.437	912.858	553.939	1271.777

Compared to CIL	**ICER**	**INMB** (for k=13,000)
BA	Dominated	-837 £
BLC	6227 £/QALY	-168 £



I. Examples of CEAs conducted in the CJS sector

Authors	Description	Measure of effectiveness	Results	Potential issues
(Griffith et al., 1999)	Examination of the cost-effectiveness of corrections-based treatment for drug abuse	3-year re-incarceration	Approximately $500 was needed to reduce the re-incarceration rate by 1% for low-risk parolees who completed treatment, compared to $165 for high risk parolees	Only relative cost-effectiveness was determined (i.e. the option with the lowest ICER)
(McCollister, French, Inciardi, et al., 2003)	CEA of Delaware's CREST Outreach Center, a work release therapeutic community (TC) and aftercare program for criminal offenders	Number of days re-incarcerated during an 18-month, post-release follow-up period	CREST program reduced incarceration for criminal offenders at an average cost of $65 per day. By adding an aftercare component to the CREST work release program, a day of incarceration is avoided at an average cost of $19 per day	Not clear if cost savings from avoided incarcerations were included in the cost estimation. Decision criteria employed was to compare ICER to the average daily cost of incarceration (i.e. if ICER is higher than costs, intervention is not cost-effective)
(McCollister, French, Prendergast, et al., 2003)	CEA of the Amity in-prison TC and Vista aftercare programs for criminal offenders in California	Number of days incarcerated over the one-year follow-up period	For the average offender, treatment reduced recidivism at a cost of $80 per incarceration day	ICER considers only incremental treatment cost and incremental days incarcerated during follow-up. Decision criteria are vague and do not compare benefits (recidivism and cost savings) and opportunity costs
(McCollister et al., 2004)	Extension of previous CEA of the Amity in-prison TC and Vista aftercare programs for criminal offenders in southern California	Number of days re-incarcerated over a 5-year follow-up	Estimated ICER of $65	ICER is compared to the average daily cost of incarceration in California ($72) to determine cost-effectiveness
(Daley et al., 2004)	Cost-effectiveness of four levels of substance abuse treatment programs for a sample of 831 offenders	Reductions in the likelihood of re-arrest within six months, one year and 18 months post-release. Recidivism includes re-arrest only	The benefits (from the perspective of the Connecticut Department of Correction, and measured in terms of the costs of avoided re-incarcerations) were from 1.8 to 5.7 times the cost of implementing the programs	Results are presented as a comparison between costs and cost savings. Opportunity costs are not assessed. Cost savings are not included in the costs of the intervention and are used instead as a measure of benefit (instead of the natural units of offending). There is a value of reducing crime itself, and then there might be potential cost savings that must be taken into account as well

(Shanahan et al., 2004)	CEA of an Australian adult drug court (ADC) program as an alternative to jail for criminal offenders addicted to illicit drugs	Time to the first offense and offending frequency per unit time	ADC was as cost-effective as were conventional sanctions in delaying the time to the first offense, and more cost-effective in reducing the frequency of offending for those outcome measures selected	Estimated ICERs were not compared to any benchmark. It is not clear how conclusions about cost-effectiveness are drawn
(Mansdotter et al., 2007)	CEA of a multi-component alcohol prevention programme targeting licenced premises in Stockholm	Savings due to prevented violence, and the health gains in terms of QALYs	The base case cost-saving ratio was 1:39. 236 gained QALYs for society as a whole. Cost-effective from the HC perspective	Analysis of crime is only based on the comparison of intervention costs and cost savings. Cost-effectiveness from HC perspective does not take into account where the intervention costs fall
(Barrett & Byford, 2012)	CEA and cost-offset analysis of the dangerous severe personality disorder (DSPD) programme	Serious reoffending	In the CEA, the intervention programme was not cost-effective for all values a decision maker may be willing to pay for preventing a serious offence. In the cost-offset analysis, expected costs of the programme were greater than the monetary value of the expected benefits	
(Muser et al., 2015)	CEA for the Paliperidone palmitate Research In Demonstrating Effectiveness (PRIDE) trial	Number of CJS events, number of incarceration events, number of incarcerations with a duration 42 days, and number of patients with at least one incarceration event	ICERs ranged from \$17,391 per CJS event avoided to \$77,731 per patient that avoided any incarceration for the paliperidone palmitate group compared with the oral antipsychotic group. ICER per CJS event avoided was \$24,409 for a 15 month period, which is consistent with what a community might be willing to pay	Community's willingness to pay to prevent incarceration used as a benchmark for cost-effectiveness. To inform decision an estimate of the CJS threshold should have been used

J. The Peterborough Social Impact Bond

According to Fox and Albertson (2011), the first pilot model based on SIB in the English and Welsh CJS was the prisoner resettlement project at Her Majesty's Prison Peterborough (Fox & Albertson, 2011). In 2010, 17 private investors (through Social Finance) committed £5 million to the Peterborough SIB to fund a series of rehabilitative interventions. The Ministry of Justice and the Big Lottery Fund agreed to pay an outcome payment of up to £8 million based on the performance over 8 years of three cohorts of 1,000 offenders each (Social Finance, 2017). The British Broadcasting Corporation suggested that a comparable rate of return on investment in a conventional bond-market was 7.5% per year[57] (BBC, 2010). This return on investment was agreed to be paid in the event of a reduction in the frequency of reconviction events of at least 10% in each of the cohorts, and/or a reduction of 7.5% in the combination of those cohorts that do not achieve the 10% reduction (Disley et al., 2015).

Reductions in reconviction frequency (calculated using data held on the Police National Computer, PNC) in the SIB cohort were compared to the number of reconviction events for a comparison group drawn from individuals released from other similar prisons who have similar characteristics (Cave et al., 2012). The frequency of reconviction events was selected as the outcome metric, rather than a binary measure of whether offenders were reconvicted or not. This had the objective of incentivising the service to continue to work with cohort members even if they were reconvicted (Disley et al., 2015).

In 2015, the programme was terminated ahead of schedule for unspecified reasons, with only two cohorts of 1000 short-sentenced male prisoners for a period of up to 12 months post-release. The results for the first cohort showed a 8.4% reduction in the frequency of reconviction events (Jolliffe & Hedderman, 2014). The reduction in reoffending for the second cohort was 9.7%. While both reductions were below the 10% target required to trigger an outcome payment for the single cohort, the weighted average of the two cohorts was above the minimum threshold of 7.5% and sufficient to trigger an outcome payment (Anders & Dorsett, 2017). As a result, investors in the Peterborough Social Impact Bond were repaid in full and received a single payment representing their initial capital plus an amount represented by a return of just over 3% per annum for the period of investment (Social Finance, 2017).

[57] In fact: 5m + 5m*0.075*8y = 8m

K. Regression outcomes for PNC data

Poisson regression models

	Model 1		Model 2		Model 3	
	Coeff.	95% CI	Coeff.	95% CI	Coeff.	95% CI
Intervention allocation:						
Brief Advice	-0.3554	[-0.7269, 0.016]	-0.4151	[-0.7662, -0.064]	-0.3208	[-0.6935, 0.0518]
Brief Lifestyle Counselling	-0.3044	[-0.7149, 0.106]	-0.2683	[-0.7543, 0.2176]	-0.2789	[-0.6915, 0.1335]
Number of cautions prior to screening	0.1765	[0.1367, 0.2163]			0.1686	[0.1253, 0.212]
Number of times in last 6 months been arrested or cautioned (at baseline)			0.1064	[0.0451, 0.1677]	0.0233	[-0.0331, 0.0798]
Constant	-0.2711	[-0.5649, 0.0226]	0.2280	[-0.0214, 0.4776]	-0.2855	[-0.5864, 0.0154]
AIC	1744		1965		1722	
BIC	1761		1982		1743	

Zero inflated negative binomial regression models

	Model 1		Model 2		Model 3	
	Coeff.	95% CI	Coeff.	95% CI	Coeff.	95% CI
Intervention allocation:						
Brief Advice	-0.3682	[-0.7242, -0.0123]	-0.3431	[-0.7268, 0.0405]	-0.3021	[-0.6566, 0.0523]
Brief Lifestyle Counselling	-0.4807	[-0.8504, -0.111]	-0.1642	[-0.5478, 0.2193]	-0.4553	[-0.8219, -0.0886]
Number of cautions prior to screening	0.1612	[0.1077, 0.2148]			0.1557	[0.105, 0.2065]
Number of times in last 6 months been arrested or cautioned (at baseline)			0.1280	[0.0224, 0.2336]	0.0257	[-0.0546, 0.1061]
Constant	0.0509	[-0.3478, 0.4497]	0.3541	[-0.0107, 0.719]	0.0619	[-0.3052, 0.4292]
Inflate						
Number of cautions prior to screening	-0.2782	[-0.581, 0.0245]			-0.1110	[-0.2926, 0.0705]
Number of times in last 6 months been arrested or cautioned (at baseline)			-1.4571	[-2.836, -0.0781]	-1.0541	[-2.0653, -0.043]
Constant	-0.6533	[-1.8018, 0.4952]	-0.5484	[-1.3749, 0.278]	-0.1153	[-0.9364, 0.7057]
Alpha [overdispersion]	1.2444	[0.7062, 2.1928]	1.7369	[1.1921, 2.5307]	1.1305	[0.7084, 1.8039]
AIC	1401		1415		1368	
BIC	1431		1444		1406	
zinb vs. zip	Pr>=chibar2 = 0.0000		Pr>=chibar2 = 0.0000		Pr>=chibar2 = 0.0000	
zinb vs. nb	Pr>z = 0.1580		Pr>z = 0.0622		Pr>z = 0.0362	

L. Criminal events (original, mean and minimum imputation scenarios)

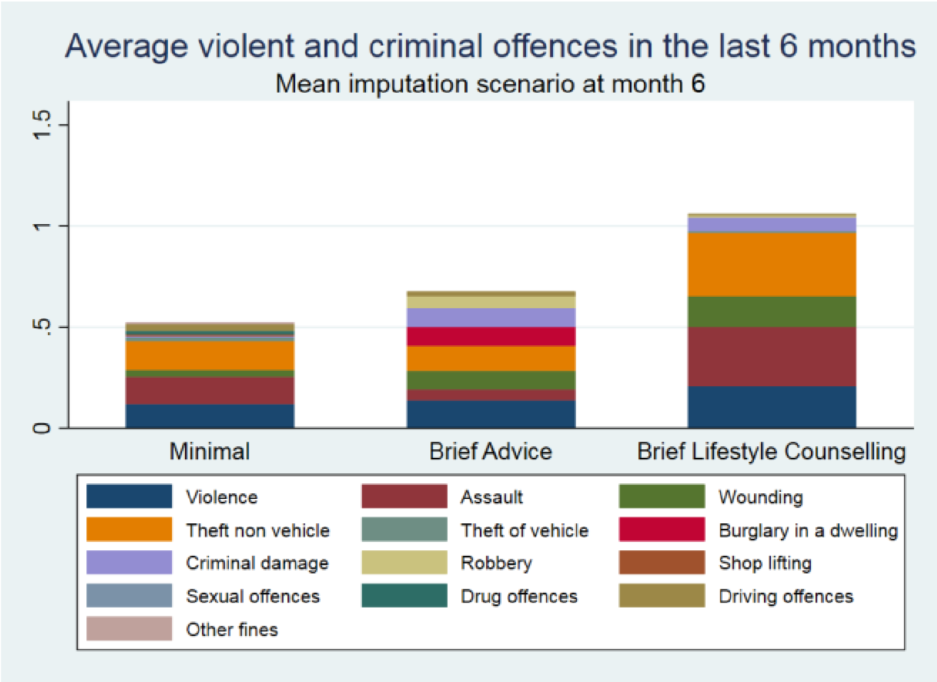
Criminal events at 6 months

	CIL			BA			BLC		
	Original	**Mean**	**Min**	**Original**	**Mean**	**Min**	**Original**	**Mean**	**Min**
Violent offences									
Violence	0.08 (0.58) $_{115}$	0.11 (0.63) $_{2}$	0.1 (0.59) $_{2}$	0.11 (0.52) $_{128}$	0.13 (0.54) $_{2}$	0.13 (0.53) $_{2}$	0.16 (0.74) $_{107}$	0.2 (0.79) $_{2}$	0.18 (0.74) $_{2}$
Assault	0.09 (0.42) $_{114}$	0.13 (0.47) $_{3}$	0.11 (0.43) $_{3}$	0.03 (0.17) $_{127}$	0.05 (0.22) $_{3}$	0.05 (0.22) $_{3}$	0.23 (1.22) $_{107}$	0.29 (1.29) $_{2}$	0.24 (1.21) $_{2}$
Wounding	0.01 (0.13) $_{115}$	0.03 (0.18) $_{2}$	0.03 (0.18) $_{2}$	0.06 (0.62) $_{129}$	0.09 (0.7) $_{1}$	0.06 (0.62) $_{1}$	0.1 (0.78) $_{107}$	0.15 (0.85) $_{2}$	0.11 (0.79) $_{2}$
Sexual offences	0 (0) $_{118}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{129}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{106}$	0 (0) $_{0}$	0 (0) $_{0}$
Criminal offences									
Theft (non-vehicles)	0.11 (0.93) $_{117}$	0.14 (0.96) $_{1}$	0.12 (0.93) $_{1}$	0.06 (0.51) $_{127}$	0.12 (0.7) $_{2}$	0.07 (0.52) $_{2}$	0.23 (1.98) $_{105}$	0.31 (2.12) $_{1}$	0.24 (1.97) $_{1}$
Vehicles theft	0 (0.09) $_{117}$	0.01 (0.12) $_{1}$	0.01 (0.12) $_{1}$	0 (0) $_{129}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0.09) $_{106}$	0 (0.09) $_{0}$	0 (0.09) $_{0}$
Burglary	0 (0) $_{118}$	0 (0) $_{0}$	0 (0) $_{0}$	0.09 (0.89) $_{129}$	0.09 (0.89) $_{0}$	0.09 (0.89) $_{0}$	0 (0) $_{106}$	0 (0) $_{0}$	0 (0) $_{0}$
Criminal damage	0 (0.09) $_{118}$	0 (0.09) $_{0}$	0 (0.09) $_{0}$	0.09 (0.88) $_{129}$	0.09 (0.88) $_{0}$	0.09 (0.88) $_{0}$	0.06 (0.5) $_{106}$	0.06 (0.5) $_{0}$	0.06 (0.5) $_{0}$
Robbery	0 (0) $_{118}$	0 (0) $_{0}$	0 (0) $_{0}$	0.04 (0.27) $_{128}$	0.05 (0.3) $_{1}$	0.05 (0.28) $_{1}$	0 (0.09) $_{106}$	0 (0.09) $_{0}$	0 (0.09) $_{0}$
Shop lifting	0 (0.09) $_{118}$	0 (0.09) $_{0}$	0 (0.09) $_{0}$	0 (0) $_{129}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{106}$	0 (0) $_{0}$	0 (0) $_{0}$
Drug offences	0.01 (0.12) $_{118}$	0.01 (0.12) $_{0}$	0.01 (0.12) $_{0}$	0 (0) $_{129}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{106}$	0 (0) $_{0}$	0 (0) $_{0}$
Driving offences	0.03 (0.22) $_{118}$	0.03 (0.22) $_{0}$	0.03 (0.22) $_{0}$	0.02 (0.15) $_{129}$	0.02 (0.15) $_{0}$	0.02 (0.15) $_{0}$	0 (0.09) $_{106}$	0 (0.09) $_{0}$	0 (0.09) $_{0}$
Other fines	0 (0.09) $_{118}$	0 (0.09) $_{0}$	0 (0.09) $_{0}$	0 (0) $_{129}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{106}$	0 (0) $_{0}$	0 (0) $_{0}$

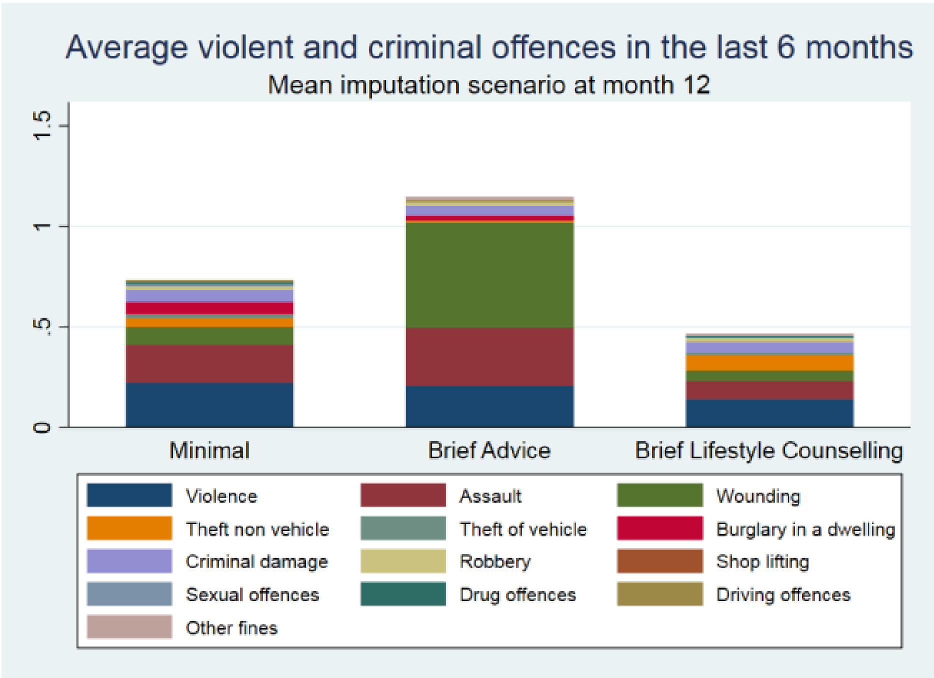
Criminal events at 12 months

	CIL			BA			BLC		
	Original	**Mean**	**Min**	**Original**	**Mean**	**Min**	**Original**	**Mean**	**Min**
Violent offences									
Violence	0.19 (0.77) $_{100}$	0.22 (0.8) $_{2}$	0.2 (0.77) $_{2}$	0.16 (0.73) $_{113}$	0.2 (0.78) $_{2}$	0.18 (0.73) $_{2}$	0.14 (0.71) $_{99}$	0.14 (0.71) $_{0}$	0.14 (0.71) $_{0}$
Assault	0.15 (0.71) $_{100}$	0.18 (0.76) $_{2}$	0.16 (0.71) $_{2}$	0.21 (1.88) $_{113}$	0.29 (1.96) $_{2}$	0.22 (1.87) $_{2}$	0.09 (0.38) $_{99}$	0.09 (0.38) $_{0}$	0.09 (0.38) $_{0}$
Wounding	0.06 (0.34) $_{100}$	0.08 (0.39) $_{2}$	0.07 (0.36) $_{2}$	0.17 (1.88) $_{113}$	0.52 (3.2) $_{2}$	0.19 (1.86) $_{2}$	0.05 (0.33) $_{99}$	0.05 (0.33) $_{0}$	0.05 (0.33) $_{0}$
Sexual offences	0 (0.09) $_{101}$	0 (0.09) $_{0}$	0 (0.09) $_{0}$	0 (0) $_{112}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{95}$	0 (0) $_{0}$	0 (0) $_{0}$
Criminal offences									
Theft (non-vehicles)	0.03 (0.22) $_{100}$	0.04 (0.26) $_{1}$	0.03 (0.24) $_{1}$	0 (0.09) $_{112}$	0 (0.09) $_{0}$	0 (0.09) $_{0}$	0.05 (0.27) $_{93}$	0.07 (0.31) $_{2}$	0.07 (0.3) $_{2}$
Vehicles theft	0.01 (0.1) $_{100}$	0.01 (0.14) $_{1}$	0.01 (0.14) $_{1}$	0 (0) $_{112}$	0 (0) $_{0}$	0 (0) $_{0}$	0.01 (0.1) $_{95}$	0.01 (0.1) $_{0}$	0.01 (0.1) $_{0}$
Burglary	0.03 (0.3) $_{100}$	0.05 (0.42) $_{1}$	0.03 (0.31) $_{1}$	0.02 (0.21) $_{112}$	0.02 (0.21) $_{0}$	0.02 (0.21) $_{0}$	0 (0) $_{95}$	0 (0) $_{0}$	0 (0) $_{0}$
Criminal damage	0.03 (0.3) $_{100}$	0.05 (0.42) $_{1}$	0.03 (0.31) $_{1}$	0.03 (0.23) $_{111}$	0.04 (0.26) $_{1}$	0.04 (0.24) $_{1}$	0.05 (0.22) $_{95}$	0.05 (0.22) $_{0}$	0.05 (0.22) $_{0}$
Robbery	0.01 (0.1) $_{100}$	0.01 (0.14) $_{1}$	0.01 (0.14) $_{1}$	0.01 (0.18) $_{112}$	0.01 (0.18) $_{0}$	0.01 (0.18) $_{0}$	0 (0) $_{93}$	0.02 (0.14) $_{2}$	0.02 (0.14) $_{2}$
Shop lifting	0 (0) $_{101}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{112}$	0 (0) $_{0}$	0 (0) $_{0}$	0 (0) $_{95}$	0 (0) $_{0}$	0 (0) $_{0}$
Drug offences	0 (0.09) $_{101}$	0 (0.09) $_{0}$	0 (0.09) $_{0}$	0 (0) $_{112}$	0 (0) $_{0}$	0 (0) $_{0}$	0.01 (0.1) $_{95}$	0.01 (0.1) $_{0}$	0.01 (0.1) $_{0}$
Driving offences	0 (0.09) $_{101}$	0 (0.09) $_{0}$	0 (0.09) $_{0}$	0 (0.09) $_{112}$	0 (0.09) $_{0}$	0 (0.09) $_{0}$	0 (0) $_{95}$	0 (0) $_{0}$	0 (0) $_{0}$
Other fines	0 (0) $_{101}$	0 (0) $_{0}$	0 (0) $_{0}$	0.01 (0.13) $_{112}$	0.01 (0.13) $_{0}$	0.01 (0.13) $_{0}$	0.01 (0.1) $_{95}$	0.01 (0.1) $_{0}$	0.01 (0.1) $_{0}$

Graph of the criminal events at 6 months (under mean imputation scenario)



Graph of the criminal events at 12 months (under mean imputation scenario)

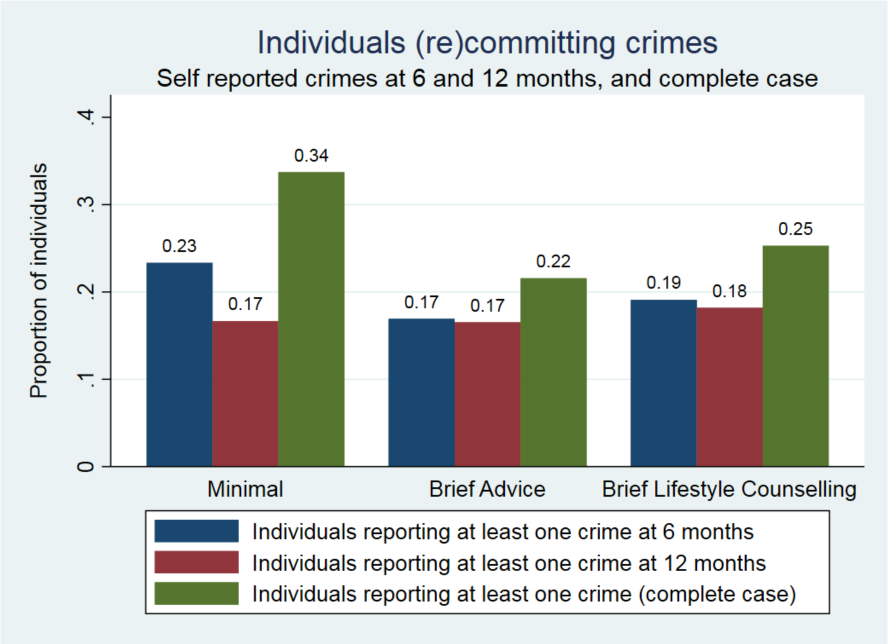


M. NB and ZINB regression models for self-reported data

	Negative binomial		Zero inflated negative binomial	
	Coeff.	95% CI	Coeff.	95% CI
Intervention allocation:				
Brief Advice	0.2136	[-0.7718, 1.1991]	0.4005	[-0.4408, 1.242]
Brief Lifestyle Counselling	0.3006	[-0.9002, 1.5015]	0.4265	[-0.436, 1.2891]
Number of times in last 6 months been arrested or cautioned (at baseline)	0.0264	[-0.1775, 0.2305]	-0.1134	[-0.3194, 0.0926]
Constant	0.1288	[-0.5599, 0.8176]	0.6984	[-0.0461, 1.443]
Inflate				
Number of times in last 6 months been arrested or cautioned (at baseline)			-0.8716	[-1.7517, 0.0084]
Constant			0.2512	[-0.7599, 1.2623]
Alpha	8.2288	[5.6319, 12.0232]	4.0983	[2.0821, 8.0669]
AIC	649.4357		645.7260	
BIC	667.3532		670.8105	
zinb vs. zip			Pr>=chibar2 = 0.0000	
zinb vs. nb			Pr>z = 0.1180	

N. Proportion of individuals self-reporting at least one crime after the intervention

In the figure below, the first (blue) and second (red) bars are based on self-reported data at six and twelve months, respectively; the third (green) bar is based on the complete cases.



From the figure, it appears that individuals treated with CIL tend to self-report a higher number of crimes both after six and twelve months. Complete cases (i.e. obtained by aggregating six- and twelve-month data, if both are non-missing) show the same pattern. Higher values in the complete case scenario are due to a reduction of the sample: 360 individuals reported information at six months, 316 at twelve months, but only 278 complete cases were available.

To estimate the appropriate odds ratios it was necessary to adjust the estimates for baseline differences. A logistic regression was performed using as the dependent variable the proportion of individuals self-reporting at least one crime in year following the intervention (complete cases only). For reasons of consistency, baseline adjustment was also based on self-reported data. Potential imbalances were adjusted using the proportions of individuals who reported that they were arrested or cautioned in the previous six months (presented in the first column of Table 5.3).

Compared to CIL, resulting odds ratios associated with individuals treated with BA and BLC were 0.5815 (95% CI: 0.2711, 1.2469) and 0.7138 (95% CI: 0.3257, 1.5643), respectively. Being allocated to BA or BLC is therefore associated with a (non-statistically significant) reduction in the risk of being committing another crime least once within one year after the intervention.

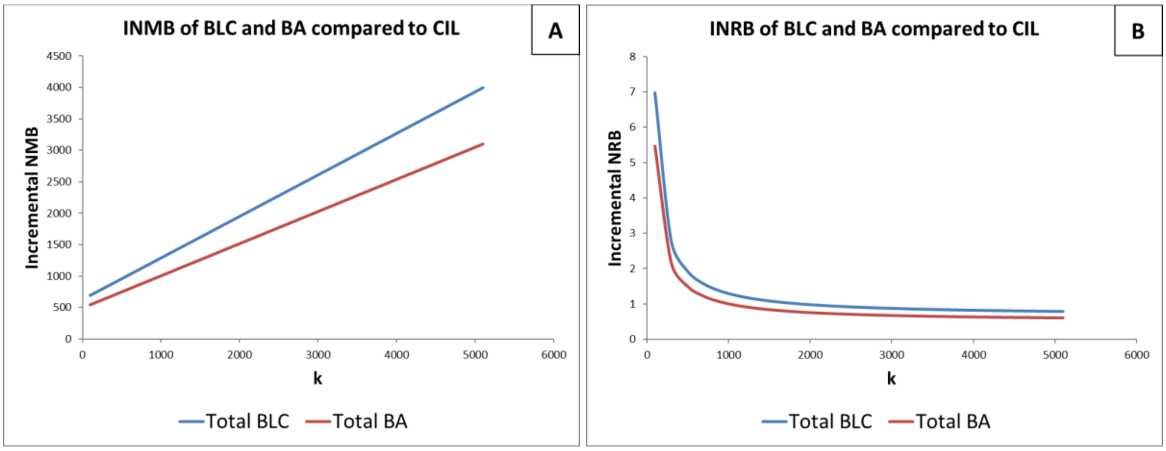
O. CEA of SIPS from the CJS perspective: further analyses

Exploring the interaction between k and cost components of SIPS

As explained in Section 5.9, the higher *k*, the greater INMBs associated with BA and BLC. This is because higher *k* means that the effectiveness (i.e. convictions avoided) of the programme is valued more in monetary terms. Nonetheless, when assessing the impact of alternative values of *k*, special caution must be taken if interpreting the results in natural terms, and not in monetary terms.

When comparing BLC and BA to CIL, for construction, conclusions are always consistent if using INMB or INRB. In fact, looking at the INB measured in reconviction events (INRB), BLC and BA provide higher benefits than CIL because they are not much more expensive and are associated with greater reductions in crime (more reconvictions avoided). Nevertheless, higher *k* means that CJS is less efficient in avoiding reconviction episodes. In other words, the CJS has a lower marginal productivity and must invest more resources in preventing one criminal episode. For this reason, the higher *k*, the lower reductions in convictions can be obtained via cost savings.

The trend of the INRB (increase in k, decrease in INRB) is therefore actually the opposite with respect to the INMB trend (increase in k, increase in INMB). For illustrative purposes, trends of INRB and INMB resulting from a univariate analysis are shown in Graphs A and B.



The opposite trends of INMB and INRB are explained by the fact that cost savings are higher than the incremental intervention costs. This situation is described in Equation 5.10.

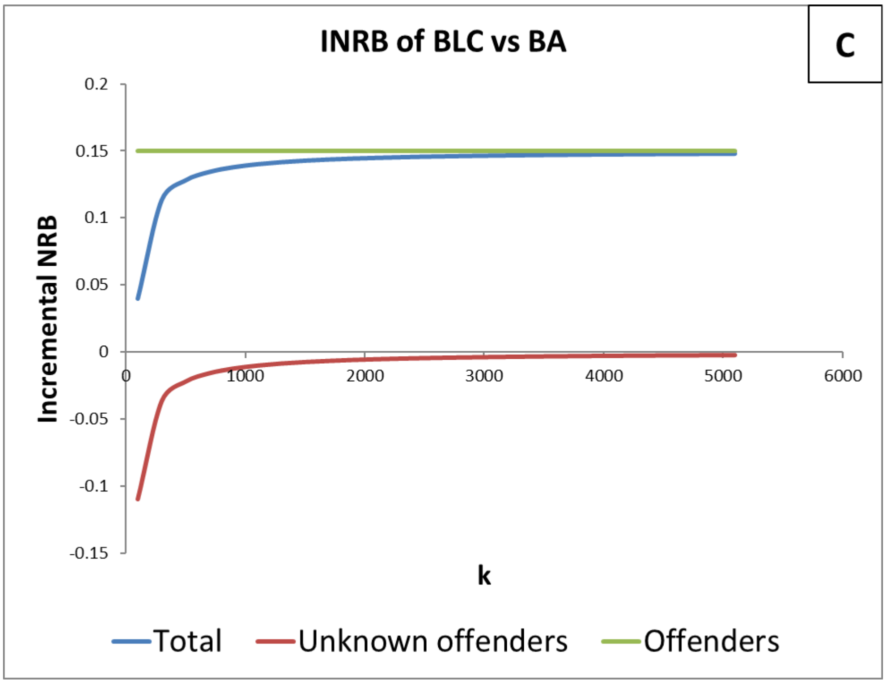
$$\Delta R * C_R > \Delta C_{int} \tag{5.10}$$

In such a scenario, for growing values of *k*, the intervention is associated with higher INMBs, but also to fewer reconvictions avoided. In fact, the costs saving component of the OCs (that can be expressed in further reconvictions averted) is greater than the costs of the intervention (i.e. the reductions in reconviction that cannot be obtained due to the resources spent to cover intervention costs).

The discrepancy in the trends of INRB and INMB disapperars when cost savings are lower than incremental intervention costs. In fact, if the condition in Equation 5.11 is fulfilled, for growing values of *k* the intervention is again associated with higher INMBs. Nevertheless, the intervention is associated also to more reconvictions averted, because the OCs (i.e. potential reductions in reconvictions that cannot be obtained due to the intervention costs) become smaller when compared to the DEs (reductions in reconvictions) of the intervention.

$$\Delta R * C_R < \Delta C_{int} \qquad (5.11)$$

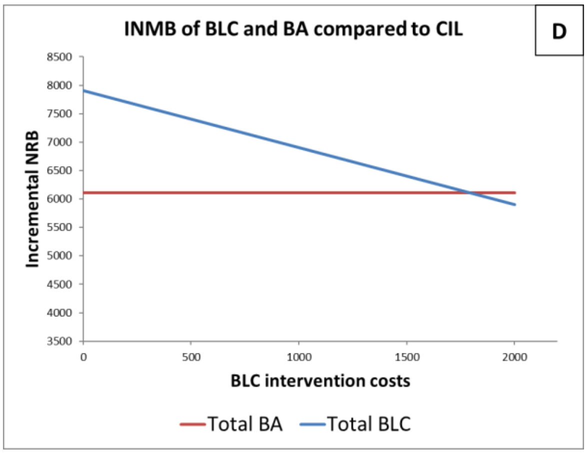
In the specific case of SIPS, Equation 5.11 is fullfilled if, for example, it is assumed that reconviction costs C_R are zero. Such a situation is illustrated in Graph C. The negative impact on unknow offenders gets smaller as *k* increases, and therefore total INRB increases together with *k*.

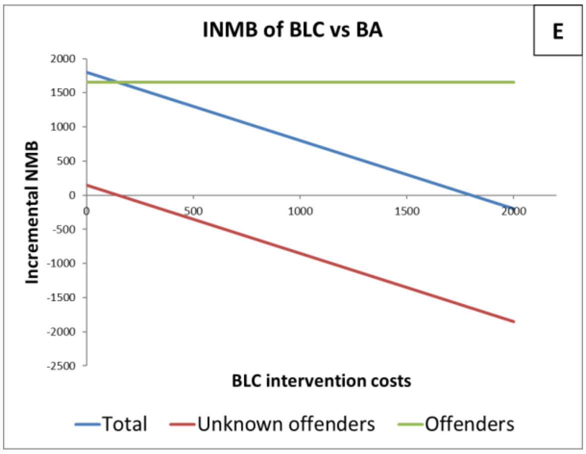


Scenario analysis: alternative intervention costs associated with BLC

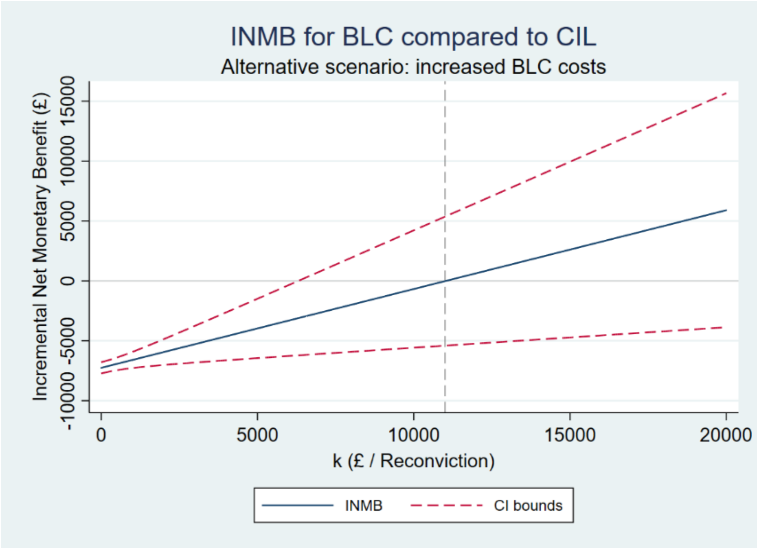
BA and BLC are the two more intense interventions, and cost £ 2.2 and £ 13.2, respectively. The minimal intervention, CIL, costs only £ 0.2. Intervention costs are very low compared to costs associated with criminal events, and differences in the costs of implementing alternative treatments are very small. Nevertheless, as discussed in Sub-section 5.8.1, intervention costs might have been actually underestimated. For this reason, alternative scenarios were explored, where intervention costs associated with BLC ranged from £ 0 (i.e. cost saving) to £ 2000 (i.e. more than 100 times more costly than what was estimated).

Graphs D and E show the INMBs associated with the variation in BLC costs. Even when the cost of implementing BLC is much higher (e.g. £ 2000), INMBs are always positive when compared to CIL. Nevertheless, ceteris paribus, £ 1700 represents the tipping point after which it would be better on average to choose BA rather than BLC.





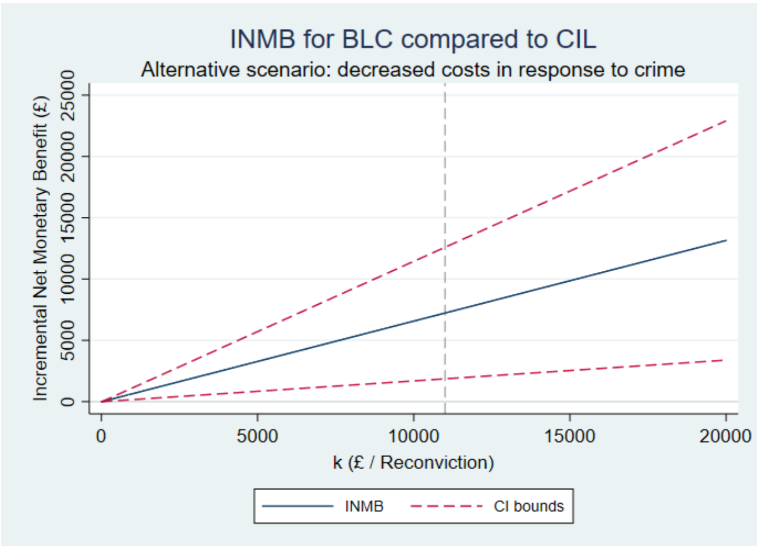
However, these are the results of a univariate sensitivity analysis and do not consider the uncertainty of other estimates. As illustrated in the figure below, if the uncertainty around the effectiveness of BLC is included, for k = 11,000 £/R, BLC should cost approximately £ 8,000 to be associated with zero incremental benefits on average.



Scenario analysis: alternative costs in response to crime

As pointed out in Sub-section 5.8.3, estimates of the costs in response to crime (i.e. CJS costs associated with each re-conviction episode) might be deemed inaccurate. These costs play an important role. If costs in response to crime were higher, INMBs associated with BA and BLC would increase because preventing one additional crime would help saving more resources. On the contrary, lower costs of recidivism would imply that cost savings are lower, and cost of the intervention would matter more.

However, in this specific case, being the costs of the intervention particularly low, INMBs of BA and BLC compared to CIL are likely to be positive even if costs in response to crime were lower. In fact, as shown in the figure below, for *k* = 11,000 £/R, BLC provides higher INMBs than CIL even if costs in response to crime are set to zero.



P. Equity weighting: threshold analysis

Results of the CEA from the HC perspective with a weight for victims' health = 1.1

Incremental Net weighted Health Benefit (INwHB)

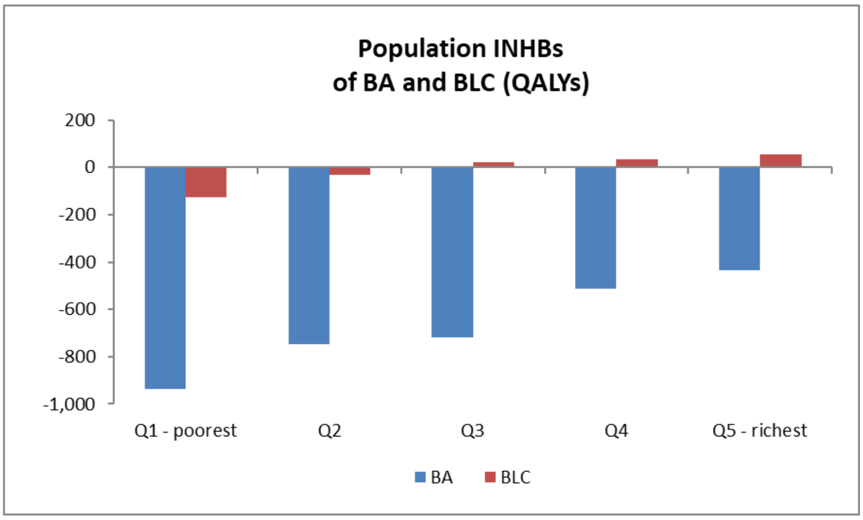
Group affected	'Traditional' estimate		Additional component		Total	
	DE	OC	DE	OC	DE	OC
Brief advice (BA)						
Offenders	-0.010				**-0.010**	
Victims			0.018		**0.018**	
Unknown victims				0.000		**0.000**
General population		0.029		-0.020		**0.009**
INwHB	**-0.039 QALYs**		**0.039 QALYs**		**0.000 QALYs**	
Brief Lifestyle Counselling (BLC)						
Offenders	-0.018				**-0.018**	
Victims			0.024		**0.024**	
Unknown victims				-0.001		**-0.001**
General population		-0.017		-0.026		**-0.044**
INwHB	**-0.001 QALYs**		**0.050 QALYs**		**0.050 QALYs**	

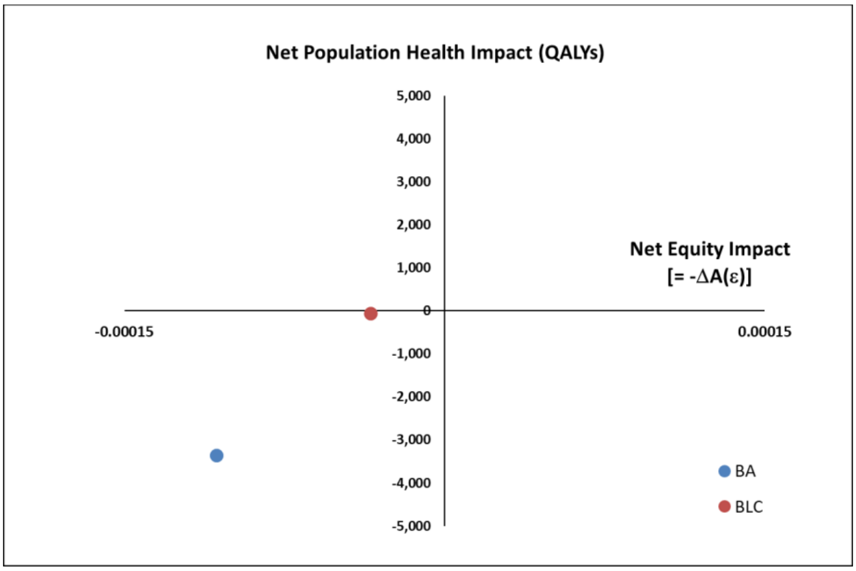
Incremental Net weighted Monetary Benefit (INwMB)

Group affected	'Traditional' estimate		Additional component		Total	
	DE	OC	DE	OC	DE	OC
Brief advice (BA)						
Offenders	-130				**-130**	
Victims			237		**237**	
Unknown victims				-5		**-5**
General population		374		-263		**111**
INwMB	**-504 £**		**505 £**		**1 £**	
Brief Lifestyle Counselling (BLC)						
Offenders	-234				**-234**	
Victims			307		**307**	
Unknown victims				-7		**-7**
General population		-227		-340		**-567**
INwMB	**-7 £**		**654 £**		**647 £**	

Q. Distributional analysis of the 'traditional' CEA

The following graphs illustrate the results of the DCEA without considering the effects on victims' health due to avoided crimes and the related additional cost savings.





Estimates associated with both BA and BLC fall into the 'lose-lose' quadrant of the health equity impact plane, being neither cost-effective nor likely to improve health equity insofar as they disproportionately benefit well-off groups. BA is the intervention that provides the greatest overall reduction in health and the worst impact on health equity.

Abbreviations

A&E: Accident and Emergency

AED: Admission and Emergency Department

AHW: Alcohol Health Worker

AIC: Akaike information criterion

APQ: Alcohol Problems Questionnaire

ASCOT: Adult Social Care Outcomes Toolkit

AUC: area under the curve

AUDIT: Alcohol Use Disorders Identification Test

BA: Brief advice

BCR: Benefit-Cost Ratio

BIC: Bayesian information criterion

BLC: Brief Lifestyle Counselling

CBA: Cost-Benefit Analysis

CCA: Cost-Consequence Analysis

CD-MCAR: MCAR conditional on baseline missing data

CE: Cost-Effectiveness

CEA: Cost-Effectiveness Analysis

CEAC: Cost-Effectiveness Acceptability Curve

CIL: Client Information Leaflet

CJ: Criminal Justice

CJS: Criminal Justice System

CMA: Cost-Minimisation Analysis

CUA: Cost-Utility Analysis

CV: Contingent Valuation

DALY: Disability-Adjusted Life Year

DCE: Discrete Choice Experiment

DCEA: Distributional Cost-Effectiveness Analysis

DE: Direct Effect

DHSC: Department of Health and Social Care

ECEA: Extended CEA

ED: Emergency Departments

EDEH: Equally Distributed Equivalent Health

FAE: Finished Admission Episode

FSM: Free School Meals

GDP: Gross Domestic Product

GLM: Generalized Linear Model

HC: Health Care

HES: Hospital Episode Statistics

HRQoL: Health-Related Quality of Life

HTA: Health Technology Assessment

IBA: Identification and Brief Advice

ICB: Inter-sectoral Costs and Benefits

ICC: Intracluster Correlation Coefficient

ICECAP: ICEpop CAPability

ICER: Incremental Cost-Effectiveness Ratio

IMD: Index of Multiple Deprivation

INB: Incremental Net Benefit

INCB: Incremental Net Consumption Benefit

INHB: Incremental Net Health Benefit

INMB: Incremental Net Monetary Benefit

INNB: Incremental Net Non-health Benefit

INRB: Incremental Net Reconviction Benefits

INwB: Incremental Net weighted Benefit

INwHB: Incremental Net weighted Health Benefit

INwMB: Incremental Net weighted Monetary Benefit

IRR: Incident Rate Ratios

ITT: Intention to Treat

LA: Local Authority

MAR: Missing at Random

MCAR: Missing Completely at Random

MCDA: Multi-Criteria Decision Analysis

MI: Multiple Imputation

MNAR: Missing not at Random

MUP: Minimum Unit Price

NB: Negative Binomial

NHS: National Health Service

NICE: National Institute for Health and Care Excellence

OASys: Offender Assessment System

OC: Opportunity Cost

OECD: Organisation for Economic Co-operation and Development

OM: Offender Manager

OR: Odds Ratio

PbR: Payments by Result

PC: Primary Care

PMM: Predictive Mean Matching

PNC: Police National Computer

QALY: Quality-Adjusted Life Year

RCR: Readiness to Change Ruler

RCT: Randomised Control Trial

RUM: Resource-Use Measurement

SALY: Safety-Adjusted Life Year

SAPM: Sheffield Alcohol Policy Model

SD: Standard deviations

SE: Standard Error

SES: Socio-economic Status

SIB: Social Impact Bonds

SIPS: Screening and Intervention Programme for Sensible drinking

SMD: Severe and Multiple Disadvantage

SROI: Social Return on Investment

SUR: Seemingly Unrelated Regressions

SWB: Subjective Wellbeing

TTO: time trade-off

vNM: von Neumann and Morgenstern

VOI: Value of Information

WELBY: Wellbeing Adjusted Life Year

WEMWBS: Warwick-Edinburgh Mental Wellbeing Scale

WHO: World Health Organization

WTP: Willingness-to-pay

ZINB: Zero Inflated Negative Binomial

Bibliography

Acheson, D. (1988). Public health in England. *London: Department of Health*.

Al-Janabi, H., et al. (2016). A Framework for Including Family Health Spillovers in Economic Evaluation. *Med Decis Making,* 36(2), 176-186.

Al-Janabi, H., et al. (2016). Measuring Health Spillovers for Economic Evaluation: A Case Study in Meningitis. *Health Econ,* 25(12), 1529-1544.

Ambrey, C. L., Fleming, C. M. and Manning, M. (2013). The life satisfaction approach to estimating the cost of crime: An individual's willingness-to-pay for crime reduction. *Australia: Griffith Business School*.

Anders, J. and Dorsett, R. (2017). *Peterborough social impact bond: final report on cohort 2 analysis*.

Angus, C., et al. (2015). *Modelling the impact of Minimum Unit Price and Identification and Brief Advice policies using the Sheffield Alcohol Policy Model Version 3*. Sheffield.

Angus, C., et al. (2014). Cost-effectiveness of a programme of screening and brief interventions for alcohol in primary care in Italy. *BMC Fam Pract,* 15, 26.

Arrow, K. J. (1950). A difficulty in the concept of social welfare. *The Journal of Political Economy*, 328-346.

Asaria, M., Griffin, S. and Cookson, R. (2016). Distributional Cost-Effectiveness Analysis: A Tutorial. *Med Decis Making,* 36(1), 8-19.

Asaria, M., et al. (2015). Distributional cost-effectiveness analysis of health care programmes--a methodological case study of the UK Bowel Cancer Screening Programme. *Health Econ,* 24(6), 742-754.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of economic theory,* 2(3), 244-263.

Atkinson, G., Healey, A. and Mourato, S. (2005). Valuing the costs of violent crime: a stated preference approach. *Oxford Economic Papers-New Series,* 57(4), 559-585.

Baji, P., et al. (2016). Comparative analysis of decision maker preferences for equity/efficiency attributes in reimbursement decisions in three European countries. *Eur J Health Econ,* 17(7), 791-799.

Bala, M. V., Zarkin, G. A. and Mauskopf, J. A. (2002). Conditions for the near equivalence of cost-effectiveness and cost-benefit analyses. *Value in Health,* 5(4), 338-346.

Barbosa, C., Godfrey, C. and Parrott, S. (2010). Methodological assessment of economic evaluations of alcohol treatment: what is missing? *Alcohol Alcohol,* 45(1), 53-63.

Barrett, B. and Byford, S. (2012). Costs and outcomes of an intervention programme for offenders with personality disorders. *Br J Psychiatry,* 200(4), 336-341.

BBC (2010). Private backers fund scheme to cut prisoner reoffending. [Online]. Available at: http://www.bbc.co.uk/news/uk-11254308.

Bhattacharya, A. (2017). Which cost of alcohol? What should we compare it against? *Addiction,* 112(4), 559-565.

Birch, S. and Donaldson, C. (2003). Valuing the benefits and costs of health care programmes: where's the 'extra' in extra-welfarism? *Soc Sci Med,* 56(5), 1121-1133.

Bishop, K. C. and Murphy, A. D. (2011). Estimating the Willingness to Pay to Avoid Violent Crime: A Dynamic Approach. *American Economic Review,* 101(3), 625-629.

Bland, J. M. (2004). Cluster randomised trials in the medical literature: two bibliometric surveys. *BMC Med Res Methodol,* 4, 21.

Bobinac, A., et al. (2012). Inquiry into the relationship between equity weights and the value of the QALY. *Value Health,* 15(8), 1119-1126.

Bojke, L., et al. (2017). How to Appropriately Extrapolate Costs and Utilities in Cost-Effectiveness Analysis. *Pharmacoeconomics,* 35(8), 767-776.

Bramley, G., et al. (2015). Hard Edges: Mapping severe and multiple disadvantage in England. *England: Lankelly Chase Foundation*.

Brand, S. and Price, R. (2000). The economic and social costs of crime. *Home Office Research Study,* 217.

Brazier, J. (2018). Extending the QALY: generating, selecting and testing items for a new generic measure of quality of life – preliminary results. *EQ Plenary Meeting*. Lisbon (Portugal).

Brazier, J., et al. (2007). *Measuring and valuing health benefits for economic evaluation*. Oxford University Press.

Brazier, J. and Tsuchiya, A. (2015). Improving Cross-Sector Comparisons: Going Beyond the Health-Related QALY. *Appl Health Econ Health Policy*.

Briggs, A. (2016). A View from the Bridge: Health Economic Evaluation - A Value-Based Framework? *Health Econ,* 25(12), 1499-1502.

Briggs, A. and O'Brien, B. J. (2001). The death of cost-minimization analysis? *Health Econ,* 10(2), 179-184.

Brookes, N., et al. (2013). *Unit Costs in Criminal Justice*. Personal Social Services Research Unit.

Broome, J. (1978). Trying to value a life. *Journal of Public Economics,* 9(1), 91-100.

Brouwer, W. (2018). The Inclusion of Spillover Effects in Economic Evaluations: Not an Optional Extra. *Pharmacoeconomics*.

Brouwer, W., et al. (2008). Welfarism vs. extra-welfarism. *J Health Econ,* 27(2), 325-338.

Buchanan, J. and Wordsworth, S. (2015). Welfarism versus extra-welfarism: can the choice of economic evaluation approach impact on the adoption decisions recommended by economic evaluation studies? *Pharmacoeconomics,* 33(6), 571-579.

Buck, D. and Gregory, S. (2013). *Improving the public's health: a resource for local authorities*. The King's Fund.

Burton, R., et al. (2017). A rapid evidence review of the effectiveness and cost-effectiveness of alcohol control policies: an English perspective. *Lancet,* 389(10078), 1558-1580.

Butler, S. M. (2018). How “Wrong Pockets” Hurt Health. Last updated. *JAMA Forum*. [Online]. Available at: https://newsatjama.jama.com/2018/08/22/jama-forum-how-wrong-pockets-hurt-health/.

Byford, S., et al. (2013). Cost-effectiveness of injectable opioid treatment v. oral methadone for chronic heroin addiction. *Br J Psychiatry,* 203(5), 341-349.

Cameron, A. C. and Miller, D. L. (2015). A practitioner’s guide to cluster-robust inference. *Journal of Human Resources,* 50(2), 317-372.

Cameron, A. C. and Trivedi, P. K. (2010). *Microeconometrics using stata*. Vol. 2. Stata press College Station, TX.

Cave, S., et al. (2012). *Peterborough social impact bond: An independent assessment*. Ministry of Justice Research Series 8/12.

Chalkidou, K., et al. (2008). Cost-effective public health guidance: asking questions from the decision-maker's viewpoint. *Health Econ,* 17(3), 441-448.

Chaplin, R., Flately, J. and Smith, K. (2011). Crime in England and Wales 2010/11', Home Office Statistical Bulletin 10/11. *London: Home office*.

Claxton, K., et al. (2019). Accounting for Timing when Assessing Health-Related Policies. *Journal of Benefit-Cost Analysis,* 9(4).

Claxton, K., et al. (2015). Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technol Assess,* 19(14).

Claxton, K., et al. (2011). Discounting and decision making in the economic evaluation of health-care technologies. *Health Econ,* 20(1), 2-15.

Claxton, K., Sculpher, M. and Culyer, A. (2007). *Mark versus Luke? Appropriate methods for the evaluation of public health interventions*. CHE Research Paper 31.

Claxton, K., et al. (2015). Causes for Concern: Is Nice Failing to Uphold Its Responsibilities to All Nhs Patients? *Health Economics,* 24(1), 1-7.

Claxton, K., et al. (2010). *Appropriate Perspectives for Health Care Decisions*. CHE Research Paper 54.

Coast, J. (2004). Is economic evaluation in touch with society's health values? *British Medical Journal,* 329(7476), 1233-1236.

Coast, J., Kinghorn, P. and Mitchell, P. (2015). The development of capability measures in health economics: opportunities, challenges and progress. *Patient,* 8(2), 119-126.

Coast, J., Smith, R. D. and Lorgelly, P. (2008). Welfarism, extra-welfarism and capability: the spread of ideas in health economics. *Soc Sci Med,* 67(7), 1190-1198.

Cohen, M. A. (2015). Willingness to Pay to Reduce White-Collar and Corporate Crime. *Journal of Benefit-Cost Analysis,* 6(2), 305-324.

Cohen, M. A., et al. (2004). Willingness-to-pay for crime control programs. *Criminology,* 42(1), 89-109.

Cookson, R. (2016). Equity-informative health technology assessment: a commentary on Ngalesoni, Ruhago, Mori, Robberstad & Norheim. *Social Science and Medicine*, 218-219.

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

Culyer, A. J. (2012). *The Humble Economist: Tony Culyer on Health, Health Care and Social Decision Making*. York and London.

Culyer, A. J. (2015). *Efficiency, equity and equality in health and health care*. Centre for Health Economics, University of York York.

Culyer, A. J. (2016). Cost-effectiveness thresholds in health care: a bookshelf guide to their meaning and use. *Health Economics, Policy and Law,* 11(4), 415-432.

Culyer, A. J. and Chalkidou, K. (2019). Economic Evaluation for Health Investments En Route to Universal Health Coverage: Cost-Benefit Analysis or Cost-Effectiveness Analysis? *Value Health,* 22(1), 99-103.

Curtis, L. (2013). *Unit Costs of Health and Social Care 2010*. Personal Social Services Research Unit.

Dahlgren, G. and Whitehead, M. (1991). *Policies and strategies to promote social equity in health*.

Dahlgren, G. and Whitehead, M. (1993). Tackling Inequalities in Health: What can we Learn from What has Been Tried? *In:* Fund, K. s. ed. *King's Fund International Seminar on Tackling Inequalities in Health*. Ditchley Park, Oxfordshire: , pp. 20-32.

Daley, M., et al. (2004). Cost-effectiveness of Connecticut's in-prison substance abuse treatment. *Journal of Offender Rehabilitation,* 39(3), 69-92.

Daniels, N. (2000). Accountability for reasonableness. *BMJ,* 321(7272), 1300-1301.

Daniels, N. and Sabin, J. (1998). The ethics of accountability in managed care reform. *Health Aff (Millwood),* 17(5), 50-64.

Daniels, N. and Sabin, J. E. (2008). Accountability for reasonableness: an update. *BMJ,* 337, a1850.

de Salazar, L., et al. (2007). *Guide to economic evaluation in health promotion*. Washington.

Deaton, A. and Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Soc Sci Med,* 210, 2-21.

Deidda, M., et al. (2019). A framework for conducting economic evaluations alongside natural experiments. *Soc Sci Med,* 220, 353-361.

Department for Communities Local Government (2015). *The English Index of Multiple Deprivation (IMD) 2015–Guidance*. Department for Communities and Local Government London, UK.

Department of Health (2010). *NHS reference costs*. [Online]. Available at: https://www.gov.uk/government/collections/nhs-reference-costs.

Department of Health (2012). *The New Public Health Role of Local Authorities*. Department of Health London.

Devlin, N. and Sussex, J. (2011). Incorporating multiple criteria in HTA. *Methods and processes. London: Office of Health Economics*.

Disley, E., et al. (2015). The payment by results Social Impact Bond pilot at HMP Peterborough: final process evaluation report. *Ministry of Justice (UK) Retrieved November 22, 2016 from www. justice. gov. uk/publications/research-and-analysis/moj*.

Dolan, P. (1997). Modeling valuations for EuroQol health states. *Medical care,* 35(11), 1095-1108.

Dolan, P. and Edlin, R. (2002). Is it really possible to build a bridge between cost-benefit analysis and cost-effectiveness analysis? *Journal of Health Economics,* 21(5), 827-843.

Dolan, P., et al. (2005). Estimating the intangible victim costs of violent crime. *British Journal of Criminology,* 45(6), 958-976.

Dolan, P. and Metcalfe, R. (2011). *Measuring subjective wellbeing for public policy*.

Dolan, P. and Peasgood, T. (2007). Estimating the economic and social costs of the fear of crime. *British Journal of Criminology,* 47(1), 121-132.

Dolan, P., et al. (2005). QALY maximisation and people's preferences: a methodological review of the literature. *Health Econ,* 14(2), 197-208.

Dolan, P. and Tsuchiya, A. (2009). The social welfare function and individual responsibility: some theoretical issues and empirical evidence. *J Health Econ,* 28(1), 210-220.

Dolan, P. and White, M. P. (2007). How Can Measures of Subjective Well-Being Be Used to Inform Public Policy? *Perspectives on Psychological Science,* 2(1), 71-85.

Donaldson, C. (1998). The (near)equivalence of cost-effectiveness and cost-benefit analyses - Fact or fallacy? *Pharmacoeconomics,* 13(4), 389-396.

Drost, R. M., et al. (2016). A Web-Based Computer-Tailored Alcohol Prevention Program for Adolescents: Cost-Effectiveness and Intersectoral Costs and Benefits. *J Med Internet Res,* 18(4), e93.

Drost, R. M., et al. (2013). Inter-sectoral costs and benefits of mental health prevention: towards a new classification scheme. *J Ment Health Policy Econ,* 16(4), 179-186.

Drost, R. M., et al. (2017). Conceptualizations of the Societal Perspective within Economic Evaluations: A Systematic Review. *Int J Technol Assess Health Care,* 33(2), 251-260.

Drummond, C., et al. (2014). The effectiveness of alcohol screening and brief intervention in emergency departments: a multicentre pragmatic cluster randomized controlled trial. *PLoS One,* 9(6), e99463.

Drummond, M., et al. (2015). *Methods for the economic evaluation of health care programmes (4th edition)*.

Drummond, M., et al. (2007). Assessing the challenges of applying standard methods of economic evaluation to public health programmes. *Public Health Research Consortium*.

Dubourg, R., Hamed, J. and Thorns, J. (2005). The economic and social costs of crime against individuals and households 2003/04. *Home Office Online Report,* 30/05.

Dukhanin, V., et al. (2018). Integrating social justice concerns into economic evaluation for healthcare and public health: A systematic review. *Soc Sci Med,* 198, 27-35.

Edejer, T.-T., et al. (2003). *Making choices in health: WHO guide to cost-effectiveness analysis*.

Edlin, R., Tsuchiya, A. and Dolan, P. (2012). Public preferences for responsibility versus public preferences for reducing inequalities. *Health Econ,* 21(12), 1416-1426.

Edwards, R. T., Charles, J. M. and Lloyd-Williams, H. (2013). Public health economics: a systematic review of guidance for the economic evaluation of public health interventions and discussion of key methodological issues. *BMC Public Health,* 13.

Erickson, D., Galloway, I. and Cytron, N. (2012). Routinizing the extraordinary. In: Francisco, F. R. B. o. S. ed. *Investing in what works for America's communities: essays on people, place, and purpose*. San Francisco (CA), pp. 377–406.

Faria, R., et al. (2014). A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *Pharmacoeconomics,* 32(12), 1157-1170.

Flatley, J. (2016). Overview of violent crime and sexual offences. *Office for National Statistics*.

Fox, C. and Albertson, K. (2011). Payment by results and social impact bonds in the criminal justice sector: New challenges for the concept of evidence-based policy? *Criminology & Criminal Justice,* 11(5), 395-413.

French, M. T. and Drummond, M. (2005). A research agenda for economic evaluation of substance abuse services. *J Subst Abuse Treat,* 29(2), 125-137.

Frew, E. (2016). Economic Evaluation of Childhood Obesity Interventions: Reflections and Suggestions. *Pharmacoeconomics,* 34(8), 733-740.

Frew, E. (2017). Aligning Health Economics Methods to Fit with the Changing World of Public Health. *Appl Health Econ Health Policy,* 15(3), 287-289.

Frew, E. and Breheny, K. (2019). Health economics methods for public health resource allocation: a qualitative interview study of decision makers from an English local authority. *Health Econ Policy Law*, 1-13.

Gafni, A. (2006). Economic evaluation of health-care programmes: Is CEA better than CBA? *Environmental & Resource Economics,* 34(3), 407-418.

Glover, D. and Henderson, J. (2010). Quantifying health impacts of government policies: a how-to guide to quantifying the health impacts of government policies. *Department of Health*.

Goranitis, I., et al. (2017). Maximizing Health or Sufficient Capability in Economic Evaluation? A Methodological Experiment of Treatment for Drug Addiction. *Med Decis Making,* 37(5), 498-511.

Greco, G., Lorgelly, P. and Yamabhai, I. (2016). Outcomes in Economic Evaluations of Public Health Interventions in Low- and Middle-Income Countries: Health, Capabilities and Subjective Wellbeing. *Health Econ,* 25 Suppl 1, 83-94.

Griffin, S., et al. (2019). Evaluation of Intervention Impact on Health Inequality for Resource Allocation. *Med Decis Making*, 272989X19829726.

Griffith, J. D., et al. (1999). A cost-effectiveness analysis of in-prison therapeutic community treatment and risk classification. *The Prison Journal,* 79(3), 352-368.

Grosse, S. D., Wordsworth, S. and Payne, K. (2008). Economic methods for valuing the outcomes of genetic testing: beyond cost-effectiveness analysis. *Genet Med,* 10(9), 648-654.

Hansard (2009). *NHS Direct* [Online]. Available at: https://publications.parliament.uk/pa/cm200809/cmhansrd/cm090203/text/90203w0029.htm and https://publications.parliament.uk/pa/cm200809/cmhansrd/cm090223/text/90223w0053.htm.

Hill, S. R., et al. (2017). Economic evaluations of alcohol prevention interventions: Is the evidence sufficient? A review of methodological challenges. *Health Policy,* 121(12), 1249-1262.

HM Treasury (2003). *The green book: appraisal and evaluation in central government*. London: HM Treasury.

Hoang, V. P., et al. (2016). A systematic review of modelling approaches in economic evaluations of health interventions for drug and alcohol problems. *BMC Health Serv Res,* 16, 127.

Holland, W. W. (2004). Health technology assessment and public health: a commentary. *Int J Technol Assess Health Care,* 20(1), 77-80.

Holmes, J., et al. (2014). Effects of minimum unit pricing for alcohol on different income and socioeconomic groups: a modelling study. *Lancet,* 383(9929), 1655-1664.

Hunter, R. M., et al. (2015). An educational review of the statistical issues in analysing utility data for cost-utility analysis. *Pharmacoeconomics,* 33(4), 355-366.

Hunter, R. M. and Hasan, S. (2013). Injectable heroin - effective treatment for opiate misusers, but is it cost-effective? *Br J Psychiatry,* 203(5), 325-326.

Hurley, J. (2000). An overview of the normative economics of the health sector. In: Culyer AJ, N. J. ed. *Handbook of Health Economics*. Vol. 1. Elsevier, pp. 55-118.

Imran Akhtar, M. G.-S., Kim Hutchings, Ah Mun Kuan, Paul Mills, Tim Phillips, Aileen Murphie (2011). The cost of a cohort of young offenders to the criminal justice system - Technical Paper. *National Audit Office,* (Ministry of Justice).

Institute of Alcohol Studies (2016). *Factsheet: Crime and social impacts of alcohol*.

Jolliffe, D. and Farrington, D. P. (2014). Self-Reported Offending: Reliability and Validity. In: Bruinsma, G. and Weisburd, D. eds. *Encyclopedia of Criminology and Criminal Justice*. New York, NY: Springer New York, pp. 4716-4723.

Jolliffe, D. and Hedderman, C. (2014). Peterborough Social impact bond: final report on cohort 1 analysis. *QinetiQ, University of Leicester*.

Jones, A. M., et al. (2013). *Applied health economics*. Routledge.

Jonsson, B. (2009). Ten arguments for a societal perspective in the economic evaluation of medical innovations. *Eur J Health Econ,* 10(4), 357-359.

Kaner, E., et al. (2013). Effectiveness of screening and brief alcohol intervention in primary care (SIPS trial): pragmatic cluster randomised controlled trial. *BMJ,* 346, e8501.

Karimi, M. and Brazier, J. (2016). Health, Health-Related Quality of Life, and Quality of Life: What is the Difference? *Pharmacoeconomics,* 34(7), 645-649.

Kattan, M. W. and Cowen, M. E. (2009). *Encyclopedia of medical decision making*. Thousand Oaks, Calif.: SAGE Publications.

Kelly, M. P., et al. (2005). *Economic appraisal of public health interventions*.

Kitchen, S., et al. (2013). *Evaluation of the free school meals pilot*. Impact Report. Ref: DFE-RR227. London: Department for Education.

Latimer, N., et al. (2009). *Prevention and early identification of alcohol use disorders in adults and young people: screening and brief interventions: cost effectiveness review*. Sheffield.

Leontaridi, R. (2003). Alcohol misuse: how much does it cost. *London: Cabinet Office*.

Little, R. J. and Rubin, D. B. (2014). *Statistical analysis with missing data*. Vol. 333. John Wiley & Sons.

Lomas, J. R. S., Martin, S. and Claxton, K. P. (2018). *Estimating the marginal productivity of the English National Health Service from 2003/04 to 2012/13*. CHE Research Paper 158.

Lorgelly, P. K., et al. (2010). Outcome measurement in economic evaluations of public health interventions: a role for the capability approach? *Int J Environ Res Public Health,* 7(5), 2274-2289.

Love-Koh, J., et al. (2015). The Social Distribution of Health: Estimating Quality-Adjusted Life Expectancy in England. *Value Health,* 18(5), 655-662.

Love-Koh, J., et al. (2016). Does public healthcare spending reduce inequality? Estimating the socioeconomic distribution of health gains and losses from marginal changes to NHS expenditure in England. *Winter 2016 Health Economists' Study Group Meeting*.

Love-Koh, J., et al. (2018). *[Draft report] Plotting interventions from nice public health guidelines on to the health equity impact plane*. YHEC, CHE York, NICE.

Man, L.-H. (2002). *Dealing with alcohol-related detainees in the custody suite*. Home Office Research, Development and Statistics Directorate.

Manca, A., Hawkins, N. and Sculpher, M. J. (2005). Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ,* 14(5), 487-496.

Manning, M., Fleming, C. M. and Ambrey, C. L. (2016). Life Satisfaction and Individual Willingness to Pay for Crime Reduction. *Regional Studies,* 50(12), 2024-2039.

Mansdotter, A. M., et al. (2007). A cost-effectiveness analysis of alcohol prevention targeting licensed premises. *Eur J Public Health,* 17(6), 618-623.

Marks, L., et al. (2015). The return of public health to local government in England: changing the parameters of the public health prioritization debate? *Public Health,* 129(9), 1194-1203.

Marmot, M., et al. (2010). *Fair society, healthy lives - The Marmot Review*.

Marsh, K., Chalfin, A. and Roman, J. K. (2008). What does cost-benefit analysis add to decision making? Evidence from the criminal justice literature. *Journal of Experimental Criminology,* 4(2), 117.

Marsh, K., et al. (2012). Estimating cost-effectiveness in public health: a summary of modelling and valuation methods. *Health Econ Rev,* 2(1), 17.

Mason, A. J., et al. (2018). A Bayesian framework for health economic evaluation in studies with missing data. *Health Econ,* 27(11), 1670-1683.

Mason, A. J., et al. (2017). Development of a practical approach to expert elicitation for randomised controlled trials with missing health outcomes: Application to the IMPROVE trial. *Clin Trials,* 14(4), 357-367.

Mayer, S., et al. (2017). Health-Related Resource-Use Measurement Instruments for Intersectoral Costs and Benefits in the Education and Criminal Justice Sectors. *Pharmacoeconomics,* 35(9), 895-908.

McCabe, C. (2018). Expanding the Scope of Costs and Benefits for Economic Evaluations in Health: Some Words of Caution. *Pharmacoeconomics*.

McCollister, K. E., et al. (2003). Post-release substance abuse treatment for criminal offenders: A cost-effectiveness analysis. *Journal of Quantitative Criminology,* 19(4), 389-407.

McCollister, K. E., et al. (2004). Long-term cost effectiveness of addiction treatment for criminal offenders. *Justice Quarterly,* 21(3), 659-679.

McCollister, K. E., et al. (2003). Is in-prison treatment enough? A cost-effectiveness analysis of prison-based treatment and aftercare services for substance-abusing offenders. *Law & Policy,* 25(1), 63-82.

McDougall, C., et al. (2003). The costs and benefits of sentencing: a systematic review. *The Annals of the American Academy of Political and Social Science,* 587(1), 160-177.

McIntosh, C. N. and Li, J. (2012). *An introduction to economic analysis in crime prevention: The why, how and so what*. National Crime Prevention Centre.

McKenna, C., et al. (2010). Budgetary policies and available actions: a generalisation of decision rules for allocation and research decisions. *J Health Econ,* 29(1), 170-181.

Meier, P. S., et al. (2016). Estimated Effects of Different Alcohol Taxation and Price Policies on Health Inequalities: A Mathematical Modelling Study. *PLoS Med,* 13(2), e1001963.

Microsoft (2010). *Excel 2010*. Version 14.0.

Ministry of Justice (2013). Transforming Rehabilitation: A Summary of Evidence on Reducing Reoffending.

Mitchell, P. M., et al. (2017). Applications of the Capability Approach in the Health Field: A Literature Review. *Soc Indic Res,* 133(1), 345-371.

Montgomery, A. A., Peters, T. J. and Little, P. (2003). Design, analysis and presentation of factorial randomised controlled trials. *BMC Med Res Methodol,* 3, 26.

Muser, E., et al. (2015). Cost effectiveness of paliperidone palmitate versus oral antipsychotics in patients with schizophrenia and a history of criminal justice involvement. *J Med Econ,* 18(8), 637-645.

Musgrave, R. A. (1959). Theory of public finance; a study in public economy.

MVH Group (1995). The measurement and valuation of health: Final report on the modelling of valuation tariffs. *Centre for Health Economics, University of York*.

National Offender Management Service (2008). *National rules for tiering cases and associated guidance, Probation Circular 08/2008*. London.

Neumann, P. J., et al. (2018). Future Directions for Cost-effectiveness Analyses in Health and Medicine. *Med Decis Making,* 38(7), 767-777.

Newbury-Birch, D., et al. (2014). Alcohol screening and brief interventions for offenders in the probation setting (SIPS Trial): a pragmatic multicentre cluster randomized controlled trial. *Alcohol Alcohol,* 49(5), 540-548.

NHS Digital (2012). *Provisional Monthly Hospital Episode Statistics for Admitted Patient Care, Outpatients and Accident and Emergency Data - April 2011 to March 2012*. Hospital Episode Statistics for Admitted Patient Care, Outpatient and Accident and Emergency Data.

NICE (2008). Social Value Judgements: Principles for the development of NICE guidance. *London: National Institute for Health and Care Excellence,* (Version 2).

NICE (2012). Methods for the development of NICE public health guidance. *NICE: London*.

NICE (2013). *Guide to the methods of technology appraisal*. National Institute for Health and Clinical Excellence (NICE) London, UK.

NICE (2015). Developing NICE Guidelines: The Manual. *NICE: London*.

Nicholls, J., et al. (2009). A guide to social return on investment. *London: Office of the Third Sector, The Cabinet Office*.

Norheim, O. (2012). *Atkinson's index applied to health. Can measures of economic inequality help us understand tradeoffs in health care priority setting*. Oxford University Press New York.

Ochalek, J., Lomas, J. and Claxton, K. (2018). Estimating health opportunity costs in low-income and middle-income countries: a novel approach and evidence from cross-country data. *BMJ Glob Health,* 3(6), e000964.

Oemar, M. and Oppe, M. (2013). EQ-5D-3DL User Guide: Basic information on how to use the EQ-5D-3L instrument. *EuroQol Group: Rotterdam*, 7.

Office for National Statistics (2017). *Deaths and Populations by sex, age and IMD decile, England & Wales, 2001-2015*.

Owen, L., et al. (2012). The cost-effectiveness of public health interventions. *J Public Health (Oxf),* 34(1), 37-45.

Palumbo, J. and Learmonth, I. (2014). *Report on the applicability of the Social Impact Bond model to reducing reoffending in Western Australia*. Social Ventures Australia, Australia.

Parkes, T., et al. (2010). Prison health needs assessment for alcohol problems. *Encyclopedia of Health Economics,* 1, 459-467.

Parliament of the United Kingdom (2010). *Equality Act, Ch. 15*. London: The Stationery Office.

Patsopoulos, N. A. (2011). A pragmatic view on pragmatic trials. *Dialogues Clin Neurosci,* 13(2), 217-224.

Paulden, M. and Claxton, K. (2012). Budget allocation and the revealed social rate of time preference for health. *Health Econ,* 21(5), 612-618.

Paulden, M., et al. (2014). Some Inconsistencies in NICE's Consideration of Social Values. *Pharmacoeconomics,* 32(11), 1043-1053.

Pearson, D. A. S., et al. (2016). Evaluation of the Citizenship Evidence-Based Probation Supervision Program Using a Stepped Wedge Cluster Randomized Controlled Trial. *Crime & Delinquency,* 62(7), 899-924.

Pepinsky, T. B. (2018). A Note on Listwise Deletion versus Multiple Imputation. *Political Analysis,* 26(4), 480-488.

Purshouse, R. C., et al. (2013). Modelling the cost-effectiveness of alcohol screening and brief interventions in primary care in England. *Alcohol Alcohol,* 48(2), 180-188.

Quartagno, M. and Carpenter, J. R. (2018). Multilevel Multiple Imputation in presence of interactions, non-linearities and random slopes. *49th Scientific meeting of the Italian Statistical Society*.

Raftery, J., et al. (2015). Clinical trial metadata: defining and extracting metadata on the design, conduct, results and costs of 125 randomised clinical trials funded by the National Institute for Health Research Health Technology Assessment programme. *Health Technol Assess,* 19(11), 1-138.

Ramsey, S. D., et al. (2015). Cost-effectiveness analysis alongside clinical trials II-An ISPOR Good Research Practices Task Force report. *Value Health,* 18(2), 161-172.

Rawls, J. (1971). *A theory of justice*. Cambridge, Harvard University Press.

Remme, M., Martinez-Alvarez, M. and Vassall, A. (2017). Cost-Effectiveness Thresholds in Global Health: Taking a Multisectoral Perspective. *Value Health,* 20(4), 699-704.

Remme, M., et al. (2014). Financing structural interventions: going beyond HIV-only value for money assessments. *Aids,* 28(3), 425-434.

Roberts, G. (2015). *Estimating the "wider societal impacts" of health conditions and treatments*. Policy Research Unit in Economic Evaluation of Health and Care Interventions. Universities of Sheffield & York.

Roberts, J. K. (2004). An introductory primer on multilevel and hierarchical linear modeling. *Learning Disabilities: A Contemporary Journal,* 2(1), 30-38.

Robson, M., et al. (2017). Eliciting the Level of Health Inequality Aversion in England. *Health Econ,* 26(10), 1328-1334.

Round, J. and Paulden, M. (2018). Incorporating equity in economic evaluations: a multi-attribute equity state approach. *Eur J Health Econ,* 19(4), 489-498.

Ryen, L. and Svensson, M. (2014). The Willingness to Pay for a Quality Adjusted Life Year: A Review of the Empirical Literature. *Health Econ*.

Sanders, G. D., et al. (2016). Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine. *JAMA,* 316(10), 1093-1103.

Sculpher, M. and Claxton, K. (2012). Real economics needs to reflect real decisions: a response to Johnson. *Pharmacoeconomics,* 30(2), 133-136.

Sculpher, M., Claxton, K. and Pearson, S. D. (2017). Developing a Value Framework: The Need to Reflect the Opportunity Costs of Funding Decisions. *Value Health,* 20(2), 234-239.

Sculpher, M., Griffin, S. and Walker, S. (2011-2019). *The economic evaluation of public health programmes with costs and effects falling outside the NHS and local authority public health budgets*. [Online]. Available at: http://phrc.lshtm.ac.uk/project_2011-2019_021.html.

Sculpher, M., et al. (2014). *Research to Inform Resource Allocation in Health and Social Care. Results of a Scoping Study*. EEPRU Research Report 017.

Sen, A. (1970). Impossibility of a Paretian Liberal. *Journal of Political Economy,* 78(1), 152-157.

Sen, A. (1993). Capability and well-being. In: Nussbaum, M. C. ed. *The quality of life*. Oxford: Clarendon Press.

Shah, K. K., Tsuchiya, A. and Wailoo, A. J. (2015). Valuing health at the end of life: a stated preference discrete choice experiment. *Soc Sci Med,* 124, 48-56.

Shanahan, M., et al. (2004). Cost-effectiveness analysis of the New South Wales adult drug court program. *Eval Rev,* 28(1), 3-27.

Shelter (2012). *Research: briefing immediate costs to government of loss of home*. [Online]. Available at: http://england.shelter.org.uk/__data/assets/pdf_file/0003/415596/Immediate_costs_to_government_of_losing_a_home.pdf.

Smith, K. and Foster, J. (2014). *Alcohol, health inequalities and the harm paradox: why some groups face greater problems despite consuming less alcohol*.

Snijders, T. A. B. and Bosker, R. J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. SAGE Publications.

Social Finance (2017). *World's 1st Social Impact Bond Shown to Cut Reoffending and to Make Impact Investors a Return*. http://www.payforsuccess.org/sites/default/files/resource-files/Final-press-release-PB-July-2017.pdf.

Squires, H., et al. (2016). A systematic literature review of the key challenges for developing the structure of public health economic models. *Int J Public Health,* 61(3), 289-298.

StataCorp (2015). *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP.

Stevens, K., Brazier, J. and Rowen, D. (2018). Estimating an exchange rate between the EQ-5D-3L and ASCOT. *Eur J Health Econ,* 19(5), 653-661.

Sugden, R. and Williams, A. (1978). *The principles of practical benefit-cost analysis*. Oxford University Press, Oxford.

Swaray, R. B., Bowles, R. and Pradiptyo, R. (2005). The application of economic analysis to criminal justice interventions: a review of the literature. *Criminal Justice Policy Review,* 16(2), 141-163.

Tariq, L., et al. (2009). Cost-effectiveness of an opportunistic screening programme and brief intervention for excessive alcohol use in primary care. *PLoS One,* 4(5), e5696.

Thokala, P., et al. (2018). Cost-Effectiveness Thresholds: the Past, the Present and the Future. *Pharmacoeconomics,* 36(5), 509-522.

Tobin, J. (1970). Limiting Domain of Inequality. *Journal of Law & Economics,* 13(2), 263-277.

Torrance, G. W. and Feeny, D. (1989). Utilities and quality-adjusted life years. *Int J Technol Assess Health Care,* 5(4), 559-575.

Trueman, P. and Anokye, N. K. (2013). Applying economic evaluation to public health interventions: the case of interventions to promote physical activity. *J Public Health (Oxf),* 35(1), 32-39.

Tsuchiya, A. and Williams, A. (2001). Welfare economics and economic evaluation. In: Press, O. U. ed. *Drummond MF. and McGuire A. (eds) Economic evaluation in health care: merging theory and practice*

Verguet, S., Kim, J. J. and Jamison, D. T. (2016). Extended Cost-Effectiveness Analysis for Health Policy Assessment: A Tutorial. *Pharmacoeconomics,* 34(9), 913-923.

Vining, A. and Weimer, D. L. (2010). An assessment of important issues concerning the application of benefit-cost analysis to social policy. *Journal of Benefit-Cost Analysis,* 1(1), 1-40.

Walker, S., et al. (2019). Striving for a societal perspective: a framework for economic evaluations when costs and effects fall on multiple sectors and decision makers. *Working paper*.

Wanless, D. (2004). Securing good health for the whole population. *London: HM Treasury,* 4, 2003-2004.

Weatherly, H., Cookson, R. and Drummond, M. (2014). Economic evaluation of public health interventions. In: *Encyclopedia of Health Economics*. Vol. 1. Elsevier, pp. 217-223.

Weatherly, H., et al. (2009). Methods for assessing the cost-effectiveness of public health interventions: Key challenges and recommendations. *Health Policy,* 93(2–3), 85-92.

Weatherly, H., Faria, R. and Van Den Berg, B. (2014). Valuing informal care for economic evaluation.

Weinstein, M. C. (1990). Principles of cost-effective resource allocation in health care organizations. *Int J Technol Assess Health Care,* 6(1), 93-103.

White, I. R., Royston, P. and Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine,* 30(4), 377-399.

Whitehurst, D. G. and Bryan, S. (2013). Trial-based clinical and economic analyses: the unhelpful quest for conformity. *Trials,* 14, 421.

WHO (2014). *Global status report on alcohol and health*. World Health Organization.

WHO (2015). *The case for investing in public health - Strengthening public health services and capacity: A key pillar of the Europe regional health policy framework, Health 2020*.

Wickramasekera, N., et al. (2015). Cost of crime: A systematic review. *Journal of Criminal Justice,* 43(3), 218-228.

Wildman, J., et al. (2016). Economic evaluation of integrated new technologies for health and social care: Suggestions for policy makers, users and evaluators. *Soc Sci Med,* 169, 141-148.

Williams, A. (1981). Welfare economics and health status measurement. In: (eds.), J. v. d. G. a. M. P. ed. *Health, Economics, and Health Economics*. pp. 271–281.

Wonderling, D. (2011). *Introduction to health economics*. McGraw-Hill Education (UK).

Woods, B., et al. (2016). Country-Level Cost-Effectiveness Thresholds: Initial Estimates and the Need for Further Research. *Value Health,* 19(8), 929-935.

Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association,* 57(298), 348-&.