White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/167647/

Version: Published Version

Monograph:
Griffin, Susan, Fusco, Francesco, Naidoo, Bhash et al. (2020) Does Health Technology Assessment guidance give adequate consideration to decisions about less costly and less effective alternatives? Working Paper. CHE Research Paper . Centre for Health Economics, University of York , York, UK.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

UNIVERSITY of York

RESEARCH

CHE

Centre For Health Economics

Does Health Technology Assessment Guidance Give Adequate Consideration to Decisions About Less Costly and Less Effective Alternatives?

Susan Griffin, Francesco Fusco, Bhash Naidoo, Matthew Taylor, Simon Walker

CHE Research Paper 175

Does Health Technology Assessment guidance give adequate consideration to decisions about less costly and less effective alternatives?

[a]Susan Griffin
[ab]Francesco Fusco
[c]Bhash Naidoo
[d]Matthew Taylor
[a]Simon Walker

[a]Centre for Health Economics, University of York, York, UK
[b]Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
[c]Centre for Guidelines, National Institute for Health and Care Excellence, London, UK
[d]York Health Economics Consortium, University of York, York, UK

November 2020

Background to series

CHE Discussion Papers (DPs) began publication in 1983 as a means of making current research material more widely available to health economists and other potential users. So as to speed up the dissemination process, papers were originally published by CHE and distributed by post to a worldwide readership.

The CHE Research Paper series takes over that function and provides access to current research output via web-based publication, although hard copy will continue to be available (but subject to charge).

Acknowledgements

Financial support for this study was provided entirely by a contract with the National Institute for Health and Care Excellence (NICE). The funding agreement ensured the authors' independence in designing the study, interpreting the data, and writing and publishing the report. The following authors were employed by NICE while the research was undertaken: Bhash Naidoo. The views expressed in this article are those of the authors and not necessarily those of NICE (www.nice.org.uk).

The authors have no conflicts to declare.

No ethical approval was needed.

Further copies

Only the latest electronic copy of our reports should be cited. Copies of this paper are freely available to download from the CHE website www.york.ac.uk/che/publications/. Access to downloaded material is provided on the understanding that it is intended for personal use. Copies of downloaded papers may be distributed to third parties subject to the proviso that the CHE publication source is properly acknowledged and that such distribution is not subject to any payment.

Printed copies are available on request at a charge of £5.00 per copy. Please contact the CHE Publications Office, email che-pub@york.ac.uk for further details.

Centre for Health Economics
Alcuin College
University of York
York,
YO10 5DD, UK
www.york.ac.uk/che

© Susan Griffin, Francesco Fusco, Bhash Naidoo, Matthew Taylor, Simon Walker

Abstract:

Cost-effectiveness analysis (CEA) plays a key role informing decision-making in healthcare and, consequently, the interpretation of its results is discussed in formal guidance from health technology assessment (HTA) organisations. A body of research indicates different willingness to pay for more effective interventions than willingness to accept less effective interventions, which some suggest supports application of different cost-effectiveness thresholds depending on whether an intervention is considered more or less effective than the comparator. We review the theoretical basis for the use of differential thresholds within HTA organisations, and question whether they are compatible with coherent decisions and social values. The National Institute for Health and Care Excellence (NICE) is one such organisation, providing recommendations on which healthcare interventions to adopt in the United Kingdom. NICE guidance describes the decision rules it employs, including comparing CEA results to a cost-effectiveness threshold that defines the boundaries beyond which an intervention is no longer considered to provide value for money. Our review of NICE guidance finds that it describes a common threshold range for all alternatives, in line with the theoretical basis for a supply-side threshold. However, we also find that the guidance focuses on the application of the threshold as a decision rule for more effective and more expensive treatments, with less guidance provided on less effective and less expensive treatments. We make suggestions for how HTA organisations can better support application of decision rules to interventions that are less effective and less expensive.

Key points:

- Guidance on comparing cost-effectiveness results to decision thresholds is often applicable only to decisions about more costly and more effective alternatives
- The need to refer to a decision threshold indicates that both health losses and gains will follow regardless of whether a more or less effective alternative is recommended
- Prominent guidance on how population health is improved by releasing resources could improve consistency in decision making

Keywords: Cost-effectiveness analysis, willingness to pay, willingness to accept, opportunity cost, cost-effectiveness thresholds, priority setting

1. Introduction

In many jurisdictions, the decision about whether to endorse an intervention for use in the healthcare system is informed by an appraisal process that includes economic evaluation to estimate the costs and health consequences from introducing alternative health interventions. The cost-effectiveness evidence aims to guide investment to interventions that provide value for money (where the benefits outweigh the costs). Although the majority of appraisals inform decisions about whether to recommend new and more costly interventions, there is potential to add value by informing decisions that recommend less effective but also less costly interventions that free up resources for other purposes.

This paper considers whether formal Health Technology Appraisal (HTA) and guideline processes work for both new and existing interventions. We examine how well decision-making committees are supported in making consistent recommendations concerning the adoption of more versus less effective interventions. We use as an exemplar the National Institute for Health and Care Excellence (NICE), which makes recommendations intended to direct health spending in the United Kingdom (UK) [1].

There are examples of decision-making committees applying different rules to existing drugs versus new drugs. This may imply that they consider the trade-off between increased cost for increased effect on a different basis to reduced cost for reduced effect [2]. In 2016 pirfenidone was part of NHS practice following a recommendation by NICE in 2013 to treat idiopathic pulmonary fibrosis sufferers with 50%-80% forced vital capacity (a patient population with moderate disease). During reappraisal the ICER was found to range between £25,700 or £28,900 per quality adjusted life year (QALY) gained compared to best supportive care in patients with 50%-80% forced vital capacity, and was similar for patients with 80%-90% forced vital capacity (a mild patient population) [3]. Although the ICERs in mild and moderate patients were comparable, the appraisal committee chose to maintain the previous recommendation in the moderate group without extending it to the mild subgroup. In the appeals that followed, it was noted that:

> *"The Committee chose not to reverse the recommendation for the 50-80% group …. because it felt that this would be unfair"[4]*
>
> *"The committee took a view that it was not cost effective in mild or moderate disease. The committee did not find the drug cost effective in the whole population, but it chose not to take away a treatment from patients with moderate disease because the NHS currently offers pirfenidone, so the NHS would have to disinvest and that is something that the committee chose not to do." [5]*

The issue of 'fairness' represents a value judgement that removing access to existing interventions warrants different consideration to providing access to previously unavailable interventions. This paper considers the role of such a value judgement in the methods and processes of HTA. Section two outlines the principles of cost-effectiveness analysis. Section three provides an overview of the documentation that guides NICE decision-making committees in making recommendations. Section four discusses how cost-effectiveness analysis informs recommendations in the NICE decision-making context, and considers the implications of applying different values to health losses compared to health gains.

2. Losses, gains and the cost-effectiveness plane

Cost-effectiveness evidence includes the difference between an intervention and a comparator in terms of incremental costs and incremental effects. Decision making committees who determine recommendations are typically presented with cost-effectiveness results in terms of the incremental cost-effectiveness ratio (ICER), which is the difference in mean incremental cost divided by the difference in mean incremental health gains, usually in terms of quality adjusted life years (QALYs). The incremental costs and effects do not show the extent of gains and losses to each individual, but rather the mean impact in the group of patients that would be subject to the recommendation [6]. The range of experiences from which those averages are calculated can be comprised of individual gains and losses.

This information can be depicted as a point on the incremental cost-effectiveness plane (Figure 1). This plane includes four quadrants, with incremental cost shown on the y-axis and incremental effects on the x-axis, with the origin representing the comparator intervention. The value of interventions that appear in the North-West (more costly, less effective) and South-East (more effective, less costly) quadrants is apparent without recourse to any cost-effectiveness threshold. When interventions appear in the North-East or South-West quadrants, there is a trade-off between costs and effects, and a cost per QALY threshold can define the accepted rate of exchange [6].

The recommendation of an intervention imposing greater costs on the healthcare sector than existing service provision requires, in the context of a fully allocated fixed budget, that service providers reduce provision of existing activities to release the required resources [6]. It represents an investment to generate health gains among individuals treated with that intervention, at the expense of disinvestment in other activities that could have produced health benefits for other individuals. The recommendation of an intervention that is less costly compared with existing service provision releases resources that can be invested in other health interventions. This recommendation can be reframed as the recommendation to cease a specific activity that is more costly and more effective. The term disinvestment is used to describe a decision to discontinue funding for an intervention currently provided within the healthcare system [7].

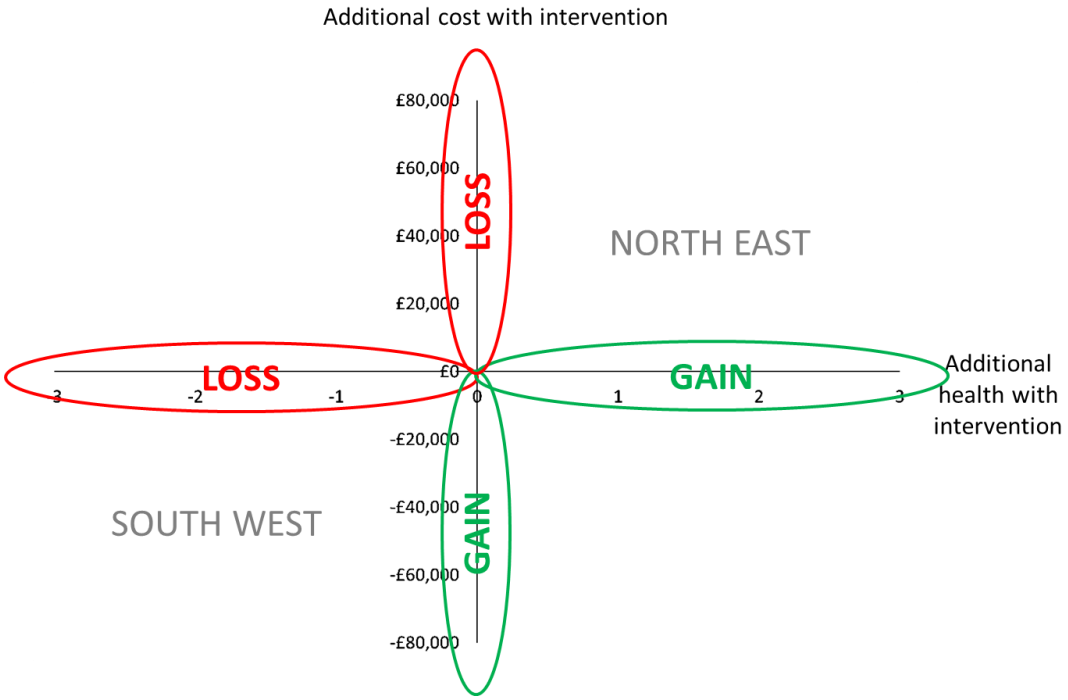

Figure 1. The cost-effectiveness plane

2.1 Framing the decision problem

With a common cost per QALY threshold in the North-East and South-West quadrants, the framing in terms of which is labelled the 'intervention' and which is labelled the 'comparator' is inconsequential. However, if different decision rules (cost-effectiveness thresholds), apply in these quadrants the designation of 'intervention' and 'comparator' is important.

When pirfenidone was chosen through the topic selection process, best supportive care (BSC) was identified as a relevant comparator. BSC was represented at the origin of the incremental cost-effectiveness plane, and the incremental costs and effects for pirfenidone versus BSC were located in the North-East quadrant of the plane (ICER £24,000 per QALY gained) [8]. NICE does not define a maximum acceptable ICER, but does indicate that below £20,000 per QALY gained, the intervention would be regarded as cost-effective [9, 10]. A threshold of £20,000 per QALY indicates that discontinuing funding from existing activities would result in health losses at a rate of one QALY per £20,000. Transferring those funds to pirfenidone would provide 0.83 QALYs per £20,000, resulting in a loss of 0.17 QALYs per £20,000 spent.

If, in the review process, BSC is appraised as the intervention and with pirfenidone labelled as a comparator, the economic evaluation would estimate an ICER for BSC compared with pirfenidone. As BSC is less costly and less effective than pirfenidone, it would fall in the South-West quadrant (ICER £24,000 saved per QALY lost). The principles of cost-effectiveness analysis indicate that less costly and less effective interventions should be recommended if the ICER is *greater* than the decision threshold, in this case £20,000 per QALY. This would ensure that technologies are recommended if they save more than £20,000 per QALY lost. For every £24,000 no longer spent on pirfenidone, there would be one fewer QALY spread among idiopathic pulmonary fibrosis sufferers, but 1.2 QALYs more across all patients served by the NHS, resulting in a gain of 0.2 QALYs per £24,000 spent. The principle is that the reduction in costs from recommending BSC would translate into more QALYs through investment in other activities than would continued investment in pirfenidone.

The notion of being less or more effective is relative, so how is the presentation determined? Identifying one intervention as current practice can provide a reference point from which to judge relative effect. Labelling interventions as current, 'new' or 'old' is problematic when current practice is composed of multiple interventions. For example, pirfenidone was represented in the North-East quadrant even when it was reappraised and represented at least a component of current NHS practice. In the presence of more than two alternatives, options are commonly ranked by their mean effects, or costs. Then, the least effective, or expensive, option is represented on the origin of the cost-effectiveness plane. In such cases only the North East quadrant of the cost-effectiveness plane is used.

2.2 Investment and disinvestment due to NICE recommendations

NICE makes decisions that affect the availability and delivery of health interventions for future individuals, without altering the entitlement of patients who are currently in receipt of an intervention ('endowed' with treatment). 'Loss' of health is in terms of comparing possible future resource allocations, and no patients are subject to interventions that reduce their quality adjusted life expectancy. In the case of pirfenidone, a decision not to recommend it would not affect the ability to continue its use in patients already receiving the drug, but would inhibit its use in future patients with the same indication.

Where NICE compares mutually exclusive alternatives, the recommendation of a new intervention can result in disinvestment of current care, unless the new intervention is adjunct or sequential to usual care. Recommending an intervention under the multiple technology appraisal process implies the discontinuation of the alternatives, at least for future patients. Under the single technology appraisal process, a recommended intervention becomes an option alongside existing alternatives, albeit one that patients can oblige the NHS to fund. Under the guidelines process, recommendations do not carry any legal mandate. Overall, a recommendation to use an intervention cannot necessarily be interpreted as an instruction to cease the availability of the alternatives.
In the appraisal process, the activities to be displaced following the recommendation of more costly interventions are unspecified. Returning to Figure 1, in the North-East quadrant the positive incremental costs represent an expected loss of health for unidentified groups, and incremental effects represent a gain in expected health among those eligible for treatment with the recommended intervention. In the South-West quadrant, the negative incremental costs represent a gain in expected health benefits for unidentified groups, and negative incremental effects represent a loss in expected health benefits among those eligible for treatment with the recommended intervention.

3. The foundations of the NICE guidance process

The decision-making committees in charge of recommendations comprise a multidisciplinary mixture of experts not limited to health economists. NICE documentation, including methods and process guides, act as an important reference point for understanding economic evaluation. Social value judgements are an inevitable feature of the NICE guidance process, and NICE has produced a document summarising the principles adopted when making decisions about cost-effectiveness. We reviewed the NICE manual (used for developing NICE guidelines), the guide to the methods of technology appraisal (which informs the technology appraisal process), and NICE Social Value Judgements document for references on the recommendation of less effective and less costly alternatives [9-11].

3.1 NICE process and methods guidance

The manual for developing NICE guidelines clearly states that guidance committees "*should be encouraged to consider recommendations that increase effectiveness at an acceptable level of increased cost or; are less effective than current practice, but free up sufficient resources that can be re-invested in public sector care or services to increase the welfare of the population receiving care*". [10]

The guide to the methods for technology appraisal states that NICE must take account of the overall resources available to the NHS, and highlights that decisions on the cost-effectiveness of an intervention must include judgements on the implications for healthcare programmes for other patient groups that may be displaced because of the new intervention.

The NICE process and methods guides stress the importance of consistency between appraisals, including in the judgements committees make regarding the cost-effective use of NHS resources. Committees are guided not to give particular priority to interventions that are part of current practice.

- *"In any situation where 'current practice', compared with an alternative approach, generates an ICER above a level that would normally be considered cost-effective, the case for continuing to invest in it should be carefully considered, based on similar levels of evidence and considerations that would apply to an investment decision."* [10]

Despite these explicit acknowledgements of recommendations of less costly and less effective interventions, our review found that throughout the guidance documents the descriptions of how the ICER is used to determine cost-effectiveness is mostly framed in terms of cost per QALY gained (the North-East quadrant of figure 1).

3.2 The NICE cost per QALY threshold

The decision-making context for NICE recommendations is the allocation of public resources, the extent of which are determined outside of NICE control. The NICE manual instructs committees to consider whether an intervention is consuming more resources than its value is contributing based on NICE's cost per QALY threshold. The NICE technology appraisal guidance states that given the fixed budget of the NHS, the appropriate maximum acceptable ICER is determined by the opportunity cost of programmes displaced by new, more costly technologies. NICE does not define a maximum acceptable ICER, but indicates a range between £20,000 and £30,000 per QALY gained.

The nuance of how the application of a decision rule to compare an ICER against a threshold differs in the South-West quadrant is not explicitly acknowledged within the NICE methods and process guides. In the South-West quadrant the threshold is a representation of the minimum acceptable ICER (the minimum amount of cost savings required to generate more value than the loss of health benefits compared to the more effective alternative). Guiding committee members to recommend interventions with ICERs greater than this threshold may reinforce the intention to free up resources where they can be re-invested to increase welfare.

3.3 NICE Social Value Judgements

What principles underlie the use of cost-effectiveness evidence by the HTA body? NICE subscribes to four moral principles (respect for autonomy; non-maleficence; beneficence; distributive justice) and the accompanying guidance notes that treatments that have adverse consequences should provide benefits that balance these harms. NICE Social Value Judgements highlights four (out of eight) principles in the chapter on evidence-based decision-making. Principles 2 to 4 relate to cost-effectiveness. Principle two states that committees must consider cost-effectiveness in making recommendations. Principle three notes that decisions should not only be based on relative costs and effects but also reflect the need to distribute health resources fairly within society. Principle four describes how committees should interpret ICERs based on cost per QALY. It notes that NICE should explain its reasons when it decides that an intervention with an ICER below £20,000 per QALY gained is not cost-effective, and when it decides that an intervention with an ICER above £20,000-£30,000 per QALY gained is cost-effective. Principle five describes how NICE reconciles respect for autonomy and individual choice against the interests of users of the NHS as a whole. It states that the expectation of individual NHS users to receive treatments to which their condition will respond does not impose a requirement on committees to recommend interventions that are not cost-effective.

Broadly, the social value judgements make clear that fairness should be assessed with reference to a balance between benefits and harms, and in terms of impacts on all NHS users as a whole. Given that NICE informs the allocation of a fixed pot of NHS resources, the NICE cost per QALY threshold does not require committees to decide how much an additional QALY is worth, but simply how many QALYs the NHS can produce with the resources demanded by alternative interventions. Consequently, NICE Social Value Judgements do not, and do not need to, include comments on the societal value to attach to an additional QALY.

4. Theoretical foundations for differential treatment of losses and gains

Here, we briefly review how previous researchers have investigated whether people treat losses differently to gains and what are the potential implications for health technology assessment [12].

In cost-effectiveness analysis and in NICE guidance, the decision rule variably refers to a cost per QALY threshold, opportunity costs, and a maximum acceptable ICER. The derivation and interpretation of cost per QALY thresholds and the determination of a maximum acceptable ICER can differ markedly according to the decision-making context. Decision-making committee members may benefit from greater support to understand these terms and why some interpretations should not have bearing on their recommendations. The basis for NICE's cost per QALY threshold is discussed in detail in the literature, and we briefly review the economic concepts they reference in support of how this threshold should be interpreted in the NICE decision-making context.

4.1 Opportunity cost

The decision to invest in a particular good removes an opportunity to use the resources invested to purchase other desirable goods. When NICE appraises interventions from a health sector perspective, opportunity cost is in terms of the restricted set of alternative investments that fall within the remit of the health sector. There has been a question as to how well the NICE cost per QALY threshold reflects opportunity cost [13, 14]. Inconsistency can arise if the basis for the NICE cost per QALY threshold departs from the opportunity cost determined by the fixed NHS budget.

4.2 Maximum prices, willingness to pay and maximum acceptable ICERs

For an individual, the willingness to pay (WTP) for healthcare refers to the maximum that individual could pay without reducing their utility (overall satisfaction). It represents the maximum price at which the individual would buy the healthcare intervention. Willingness to accept (WTA) refers to the amount of compensation the individual would require to maintain their utility following the loss of a good (it represents the minimum price at which an individual would sell the good). The appropriateness of using WTA instead of WTP depends on whether individuals currently have the good in question or have a legal entitlement to it [15].

For society, the WTP for healthcare represents the maximum that society would be willing to forgo in non-healthcare or private consumption to obtain that healthcare [16, 17]. For example, the societal WTP for health could be used to inform decisions that determine the size of the NHS budget. Although the terms WTP and cost-effectiveness threshold has been used interchangeably in the past, differences arise if the size of the budget does not correspond with societal WTP for health gains [18]. The cost per QALY threshold should represent the opportunity costs of the decision, which are the effects on patients' health due to a marginal change in expenditure given a fixed budget. The maximum acceptable ICER indicates the cost per QALY gained from those programmes that would be discontinued to release funds for other activities.

4.3. Differential treatment of losses and gains

Many studies find that WTA exceeds WTP for the same good, but there is a lack of consensus as to whether the difference is meaningful, and the magnitude of the difference [19-28]. Competing explanations for why the discrepancy might reflect a true difference in value include loss aversion and the endowment effect, status quo bias, and the substitution effect [29, 30]. Section 4.2 showed that the notion of WTP should not apply within the NICE decision-making context. However, this discrepancy between valuing health losses versus health gains could potentially influence committee members in forming recommendations.

4.3.1 Differential treatment of losses and gains from the demand side

For decision-making contexts where the budget is not fixed, one might value health gains in terms of WTP and health losses in terms of WTA. Recalling the comparison of pirfenidone (P) to BSC, recommendation of the more costly intervention pirfenidone leads to a loss of health from alternative health improving interventions. This means that additional costs ($C_P - C_{BSC}$) should be valued using WTA, and compared with the health gains ($Q_P - Q_{BSC}$) valued using WTP. If we say that WTA is simply WTP multiplied by the ratio of WTA to WTP (WTA:WTP), we can weight health losses by this ratio in calculating the ICER.[1] The decision rule for a more costly intervention such as pirfenidone can be written as in Equation (1).

$$(1) \qquad \frac{(C_P - C_{BSC})\frac{WTA}{WTP}}{(Q_P - Q_{BSC})} < maximum\ ICER$$

We rearrange the decision rule in Equation (1) to show the ICER for pirfenidone versus BSC against a maximum ICER or cost per QALY threshold weighted by WTA:WTP. Equation (2) applies to recommendations for interventions depicted in the North-East quadrant of the incremental cost-effectiveness plane:

$$(2) \qquad \frac{(C_P - C_{BSC})}{(Q_P - Q_{BSC})} < maximum\ ICER \div \frac{WTA}{WTP}$$

Recommendation of less costly interventions leads to expected health gains through the release of resources to provide other health improving interventions. This means that cost savings should be valued using WTP, and compared with health losses valued using WTA. Equation (3) shows the decision rule for BSC, where the health losses are weighted using the ratio of WTA:WTP. Equation (4) rearranges this in terms of an intervention that is depicted in the South-West quadrant of the incremental cost-effectiveness plane:

$$(3) \qquad \frac{(C_{BSC} - C_P)}{(Q_{BSC} - Q_P)\frac{WTA}{WTP}} > maximum\ ICER$$

$$(4) \qquad \frac{(C_{BSC} - C_P)}{(Q_{BSC} - Q_P)} > maximum\ ICER.\frac{WTA}{WTP}$$

Equations (2) and (4) reveal how a discrepancy between WTA and WTP leads to the use of different thresholds in the North-East and South-West quadrants. The maximum acceptable ICER is multiplied by the ratio of WTP:WTA in the North-East quadrant, and the ratio of WTA:WTP in the South-West quadrant.

4.3.2 The implications of kinked thresholds

Figure 2 compares a kinked decision threshold (line A0B), assuming a WTA to WTP ratio of 2:1, with a single threshold, assuming a maximum acceptable ICER of £20,000 per QALY (line C0D). Based on these assumptions, equation (4) implies that the decision threshold would be to recommend technologies with ICERs above £40,000 saved per QALY lost in the South-West quadrant. Equation (2) indicates that the decision rule would be to recommend technologies with ICERs under £10,000 per QALY gained in the North-East quadrant. More effective and more costly technologies with ICERs

[1] This equally applies if the cost-effectiveness results are presented using the net monetary benefit (NMB) or net health benefit (NHB). The calculation of the NMB and NHB relies on the conversion of health benefits into monetary terms, or costs into health benefit equivalent, i.e. on the rate of exchange between costs and health outcomes.6. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes: Oxford university press; 2015.

in the range £11,000 to £20,000 would improve net health, but would not be recommended. Less effective and less costly technologies with ICERs in the range £20,000 to £39,000 saved per QALY lost would improve net health, but would not be recommended. A technology that was recommended with an ICER below £10,000 per QALY would have to be re-appraised at an ICER above £40,000 per QALY before the recommendation would be reversed.

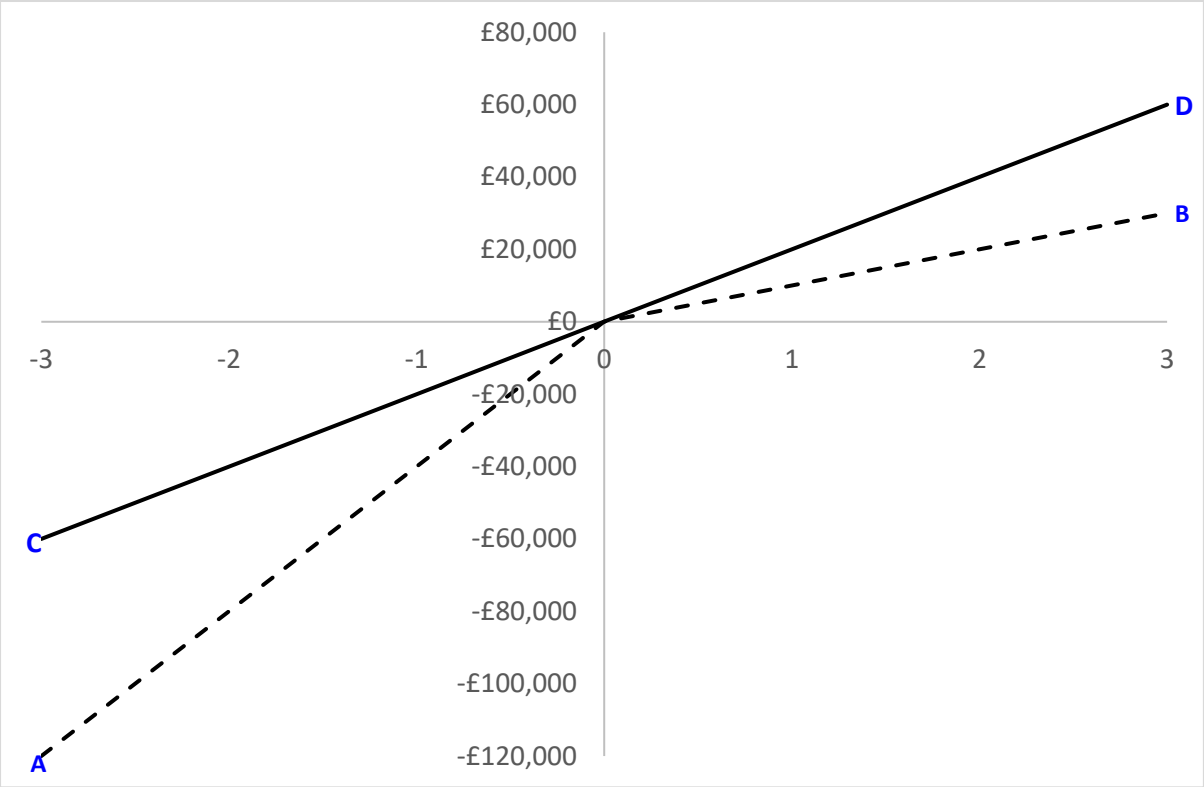

Figure 2. Kinked threshold

A kinked threshold leads to a *status quo* bias, where the current treatment is compared with a more permissive threshold (e.g. £40,000 per QALY) than the one used for new programmes (e.g. £10,000 per QALY). While the current patients will receive the already funded treatment, this will not be necessarily true for future patients who could benefit as much or even more from new treatments. This introduces an inconsistency in how health is valued between current and future patients [13, 31].

An additional issue with a kinked threshold would be the time ordering of the decision. In this respect, applying different values to increased costs compared with reduced costs means that the decision will differ depending on the chronological order in which treatments are appraised. This has implications for NICE processes and review of previous recommendations. Following a decision to approve a costlier and more effective intervention like pirfenidone for use in the NHS, it becomes the new current practice. In light of new evidence, and without the arrival of new comparators, the review compares the less effective and less costly BSC against this new current practice. Unless new evidence suggests that pirfenidone is less expensive than BSC, the threshold that the ICER for BSC compared with pirfenidone would have to exceed would be higher than that applied in the initial appraisal. In general, this could create sclerosis in the system in favour of the comparator selected to represent the origin of the cost-effectiveness plane.

4.4 Differential treatment of losses and gains from the supply side

Existing interventions are displaced or contracted to release resources when a more costly intervention is recommended. This set of displaced interventions differ to the programmes that would be introduced or expanded when funds are released following the recommendation of less expensive intervention [32]. Diminishing marginal returns indicates that as you expand budgets, the amount of health gained from each additional pound spend will fall. This could be represented by a kinked threshold.

Lomas et al. attempted to estimate the average change in health observed consequent to decisions that accompany either the contraction or the expansion of a healthcare budget [33]. For the NHS, it was empirically determined that expanding expenditure by 3% might provide one additional QALY per £13,464 invested, but that contracting expenditure by 2% would remove interventions that were generating one QALY per £12,047 invested [33]. Based on these estimates, the curvature in the supply side of the threshold would be relatively small. Eckerman and Pekarsky took an alternative view that compared the optimal programme for expansion to the optimal programme for contraction [34]. The kink in the supply side of the threshold proposed by Eckerman and Pekarsky is substantially larger than the one observed in the empirical study conducted by Lomas et al [35]. Eckerman and Pekarsky assume that resources obtained from recommending less costly and less effective alternatives would be reallocated to the most cost-effective funded technology, while the funds for more costly alternatives are obtained by disinvesting from the least cost-effective technology. This assumption does not appear to be reflected in the observed data, nor reflect the HTA process where NICE does not direct the resource reallocation required to support its recommendations.

5. Discussion

This paper summarises the foundations of using cost-effectiveness evidence in health technology assessment and examined the consequences of differential treatment of losses and gains in healthcare.

Whilst decision-making in healthcare is not merely concerned with cost-effectiveness, formal HTA processes employing cost-effectiveness thresholds to define value for money still play an important role in resource allocation. However, much of the process and application of HTA tools has focussed on decision rules in the case of more effective and more expensive technologies, and describing the social value judgments that mitigate these rules [10, 11]. On the contrary, the decision rules in evaluating programmes able to free up resources – albeit at the cost of reduced effectiveness – are described in less detail. Guidance committees are asked to make consistent judgements across interventions. There is clear evidence for biases in individual decision-making favouring identified patients over the unidentified patients who bear the opportunity cost if additional resources are required for an intervention. Similarly, individuals ascribe a greater value to a loss compared with a gain of an equivalent amount.

Based on the current NICE manuals, the same threshold should be applied regardless of whether the ICER is presented in the North-East or the South-West quadrant of the incremental cost-effectiveness plane. However, the appraisal of currently funded technologies, such as the reappraisal of pirfenidone, have suggested otherwise. By better elucidating how cost-effectiveness analysis justifies the adoption of a less effective technology which would release *sufficient* resources to be reinvested in programmes that offset the health loss, existing NICE processes and manuals could potentially do more to prevent committee members inadvertently applying psychological bias in forming recommendations.

Differential treatment of losses and gains could provide a possible explanation for incongruence in the decision rule applied to the North-East and South-West quadrants. Although WTA the loss of a currently provided service might exceed WTP for the same service at the individual level, there is limited evidence as to whether this discrepancy manifests at a societal level [12, 20, 22]. However, the decision context and the cost-effectiveness threshold applied in HTA bodies such as NICE is not the societal WTP for healthcare (although this societal value may inform the budgets within which they operate). Where decision bodies such as NICE do not control the size of the healthcare budget, any discrepancy between societal WTA and WTP for health should not affect their decisions. HTA processes operating in settings without fixed healthcare budgets may issue recommendations that control how the resources are deployed between health and other sectors.

Limited attention has been paid to the fact that health gains and losses occur as a consequence of all NICE recommendations. For example, we found little in the NICE manuals and social value judgements to make explicit the fact that disinvestment from existing, health improving activities will occur because of recommendations for more costly interventions.

The existing literature provides extensive support for why a demand-side interpretation of the cost-effectiveness threshold is irrelevant in the decision-making context of a body such as NICE. This supports the use of a single decision-making threshold. Some have explored whether there is support for a different decision rule for North-East versus South-West quadrants from a supply side interpretation of the cost-effectiveness threshold. The cost of generating a unit of healthcare through the alternative activities forgone because of any recommendation will differ between cost increasing versus cost saving activities. The empirical evidence on the diminishing marginal returns of health expenditure suggests that if there is any discrepancy it might be negligible. For example,

the impact of pirfenidone on the total healthcare expenditure in 2016 was 0.014% [36, 37], which is much smaller than the budget expansion (3%) or contraction (2%) required to produce a kink of around £1,500 in the threshold [33].

A consequence of adopting a different decision rule for assessing a more costly intervention compared with ceasing to provide the same intervention is that the temporal ordering of decisions affects the recommendations. Allowing for a kinked threshold would lead to differential consequences for future patients, which implies a different weighting of current patients over future patients [31]. When HTA compares the average change in cost divided by the average change in health provided by a new programme against a threshold, it is in pursuit of improving society-wide health. Applying a kinked threshold that goes beyond any difference in opportunity costs for cost saving as opposed to cost increasing technologies, appears in contrast with making coherent decisions and, therefore, it would pose important normative and ethical questions.

The deployment of a health programme will inevitably result in health losses, either directly or indirectly. When a more effective and more costly intervention is appraised, this loss is indirectly captured by the threshold, which represents the alternative use of the forgone resources. The health loss due to a less effective and less expensive intervention, on the other hand, is directly captured by the reduction in population health. The NICE documentation shed light on the importance of balancing these health losses, ultimately improving population health, setting the cost-effectiveness boundaries for the North-East technologies. Future guidelines have the opportunity to review the current process and, thereby, to make more prominent that the assessment of South-West technologies does not differ from the North-East one since they both impose health losses.

References

1. National Institute for Health and Care Excellence. Mission statement. 2019 [cited 2019 18/06/19]; Available from: https://www.nice.org.uk/about
2. Garner S, Littlejohns P. Disinvestment from low value clinical interventions: NICEly done? *BMJ* 2011;343:d4519.
3. National Institute for Health and Care Excellence. Final appraisal determination Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837] 2016 [cited 2019 28/06/19]; Available from: https://www.nice.org.uk/guidance/ta504/documents/final-appraisal-determination-document-2
4. National Institute for Health and Care Excellence. Single technology appraisal appeal hearing. Advice on pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837] - appeal decision letter 1. 2016 [cited 2019 28/06/19]; Available from: https://www.nice.org.uk/guidance/ta504/documents/appeal-decision
5. National Institute for Health and Care Excellence. Single technology appraisal appeal hearing. Advice on pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837] - appeal decision letter 2. 2016 [cited 2019 28/06/19]; Available from: https://www.nice.org.uk/guidance/ta504/documents/appeal-decision-2
6. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes: Oxford university press; 2015.
7. National Institute for Health and Care Excellence. NICE Do not Do prompts.2019.
8. National Institute For Health and Care Excellence. Pirfenidone for treating idiopathic pulmonary fibrosis. [TA282]: National Institute for Health and Clinical Excellence.; 2013.
9. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 4 April 2013.
10. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual: National Institute for Health and Care Excellence; 2014 31 October 2014.
11. National Institute for Health and Care Excellence. Social Value Judgements: Principles for the Development of NICE Guidance: National Institute for Health and Care Excellence, 2008 July 2008.
12. Tunçel T, Hammitt JK. A new meta-analysis on the WTP/WTA disparity. *Journal of Environmental Economics and Management* 2014;68(1):175-87.
13. Claxton K, Sculpher M, Palmer S, Culyer AJ. Causes for concern: is NICE failing to uphold its responsibilities to all NHS patients? *Health Economics* 2015;24(1):1-7.
14. Dillon A. Carrying NICE over the threshold. NICE; 2015.
15. Carson RT. Contingent Valuation: A User's Guide. *Environ Sci Technol*. 2000;34:1413r8.
16. Sugden R, Williams A. The principles of practical cost-benefit analysis. Oxford University Press Catalogue. 1978.
17. Arrow KJ. A difficulty in the concept of social welfare. *Journal of Political Economy*. 1950;58(4):328-46.
18. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics* 2008;26(9):733-44.
19. Sugden R. Anomalies and stated preference techniques: a framework for a discussion of coping strategies. *Environmental and Resource Economics* 2005;32(1):1-12.
20. Horowitz JK, McConnell KE. A review of WTA/WTP studies. *Journal of environmental economics and Management* 2002;44(3):426-47.
21. Shogren JF, Shin SY, Hayes DJ, Kliebenstein JB. Resolving differences in willingness to pay and willingness to accept. *The American Economic Review* 1994:255-70.
22. Whynes DK, Sach TH. WTP and WTA: do people think differently? *Social Science & Medicine.* 2007;65(5):946-57.

23. Morrison GC. Willingness to pay and willingness to accept: some evidence of an endowment effect. *Applied Economics* 1997;29(4):411-7.
24. Martín-Fernández J, Ariza-Cardiel G, Peña-Longobardo LM, Polentinos-Castro E, Oliva-Moreno J, Gil-Lacruz AI, et al. "Gaining or losing": The importance of the perspective in primary care health services valuation. *PloS One*. 2017;12(12):e0188969.
25. O'brien BJ, Goeree R, Gafni A, Torrance GW, Pauly MV, Erder H, et al. Assessing the value of a new pharmaceutical: a feasibility study of contingent valuation in managed care. *Medical Care.* 1998:370-84.
26. Grutters JP, Kessels AG, Dirksen CD, Van Helvoort-Postulart D, Anteunis LJ, Joore MA. Willingness to accept versus willingness to pay in a discrete choice experiment. *Value in Health*. 2008;11(7):1110-9.
27. Knetsch JL, Sinden JA. Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. *The Quarterly Journal of Economics.* 1984;99(3):507-21.
28. O'Brien BJ, Gertsen K, Willan AR, Faulkner A. Is there a kink in consumers' threshold value for cost-effectiveness in health care? *Health Economics* 2002;11(2):175-80.
29. Kahneman D, Knetsch JL, Thaler RH. Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*. 1991;5(1):193-206.
30. Hanemann WM. Willingness to Pay and Willingness to Accept: How Much Can They Differ? *The American Economic Review*. 1991;81(3):635-47.
31. Dowie J. Why cost-effectiveness should trump (clinical) effectiveness: the ethical economics of the South West quadrant. *Health Economics*. 2004;13(5):453-9.
32. Culyer AJ. Cost-effectiveness thresholds in health care: a bookshelf guide to their meaning and use. *Health Economics, Policy and Law*. 2016;11(4):415-32.
33. Lomas J, Claxton K, Martin S, Soares M. Resolving the "Cost-Effective but Unaffordable" Paradox: Estimating the Health Opportunity Costs of Nonmarginal Budget Impacts. *Value Health*. 2018 Mar;21(3):266-75.
34. Eckermann S, Pekarsky B. Can the Real Opportunity Cost Stand Up: Displaced Services, the Straw Man Outside the Room. *Pharmacoeconomics* 2014 April 01;32(4):319-25.
35. Paulden M, McCabe C, Karnon J. Achieving Allocative Efficiency in Healthcare: Nice in Theory, not so NICE in Practice? *Pharmacoeconomics* 2014 April 01;32(4):315-8.
36. Harker R. NHS Funding and Expenditure 2017.
37. National Institute for Health and Care Excellence. Single Technology Appraisal. Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837] Committee Papers.