Health Economics, Policy and Law (2022), **17**, 444–470
doi:10.1017/S1744133122000032

HEALTH ECONOMICS, POLICY and LAW

ARTICLE

The normative grounds for NICE decision-making: a narrative cross-disciplinary review of empirical studies

Victoria Charlton

Department of Global Health & Social Medicine, King's College London, London, UK
Corresponding author. Email: victoria.charlton@kcl.ac.uk

(Received 15 February 2021; revised 10 January 2022; accepted 15 February 2022; first published online 16 March 2022)

Abstract

The National Institute for Health and Care Excellence (NICE) is the UK's primary health care priority-setter, responsible for advising the National Health Service on its adoption of health technologies. The normative basis for NICE's advice has long been the subject of public and academic interest, but the existing literature does not include any comprehensive summary of the factors observed to have substantively shaped NICE's recommendations. The current review addresses this gap by bringing together 29 studies that have explored NICE decision-making from different disciplinary perspectives, using a range of quantitative and qualitative methods. It finds that although cost-effectiveness has historically played a central role in NICE decision-making, 10 other factors (uncertainty, budget impact, clinical need, innovation, rarity, age, cause of disease, wider societal impacts, stakeholder influence and process factors) are also demonstrably influential and interact with one another in ways that are not well understood. The review also highlights an over-representation in the literature of appraisals conducted prior to 2009, according to methods that have since been superseded. It suggests that this may present a misleading view of the importance of allocative efficiency to NICE's current approach and illustrates the need for further up-to-date research into the normative grounds for NICE's decisions.

Key words: Health care priority setting; health technology assessment; National Institute for Health and Care Excellence; value judgements; resource allocation; literature review

1. Introduction

In any contemporary health system, the population's demand for health care is likely to exceed the system's capacity to provide it. Policymakers must therefore decide which interventions to fund and which to exclude from coverage (Ubel, 2000; Fleck, 2002; Alexander *et al.*, 2004; Scheunemann and White, 2011). Given the implications of such decisions for those who contribute to and benefit from national health systems, centralised processes are often put in place to ensure that they are seen to be made fairly (Kenny and Joffres, 2008). In the UK, these fall under the remit of the National Institute for Health and Care Excellence (NICE), a public body which has come to be seen as a world-leader in health care priority-setting (Smith, 2004; Timmons *et al.*, 2016; Schaefer and Schlander, 2018; Littlejohns *et al.*, 2019; Catchpole and Barrett, 2020).

In making recommendations to the National Health Service (NHS), NICE is legally required to 'have regard to the broad balance between the benefits and costs' of the technologies that it considers (NICE, 2013). This aligns with the NHS's mandate to provide 'best value for taxpayers' money' by maximising the amount of health that can be delivered from its budget (Department of Health, 2015). A technology's cost-effectiveness – understood as the number

of quality-adjusted life-years (QALYs) delivered per unit cost, compared with existing treatment – is therefore NICE's key substantive consideration. However, maximising health is not the only relevant objective: the NHS also promises to employ resources in a way that is 'fair and sustainable' and fulfils 'a wider social duty to promote equality' (Department of Health, 2015). NICE therefore takes the position that its advice should also 'take into account other factors' which might on occasion justify the recommendation of technologies that likely displace more QALYs than they deliver (NICE, 2008*b*, 2020*b*).

Given the real-world implications of NICE's advice and the organisation's reputation as an authority on health care priority-setting, the normative grounds for its recommendations have long been the subject of public and academic interest. However, absent from the current literature is any comprehensive summary of what factors have been substantively employed by NICE's independent appraisal committees in practice. This review aims to address this gap by bringing together studies that have empirically examined NICE decision-making from a range of disciplinary perspectives, using a variety of quantitative and qualitative methods. In doing so, it provides a foundation for further in-depth research into the grounds for NICE decision-making and the basis on which NICE's evolving approach can be ethically justified. It is hoped that this contribution will prove particularly timely at present given NICE's recent review of its processes and methods.[1]

2. Methods

2.1 Approach

The study takes the form of a narrative literature review (Grant and Booth, 2009; O'Connor and Sargeant, 2015; Paré *et al*., 2015), a methodology well-suited to the aim of obtaining insight into a specific research question from studies that are broad in scope, use a wide range of different methods and are informed by an array of disciplinary perspectives. In synthesising evidence from these studies, narrative summary (Dixon-Woods *et al*., 2005) has been adopted as the most appropriate way of integrating diverse quantitative and qualitative findings into a readily digestible review suitable for a multi-disciplinary audience.

2.2. Data sources and searches

A comprehensive literature search was conducted across five databases: ProQuest, PubMed, Scopus, Web of Knowledge and Lexis Library. These were selected to provide coverage across multiple disciplines including medicine (PubMed), social science and humanities (ProQuest and Scopus), biomedical and natural science (Web of Knowledge) and law (Lexis). Coverage was tested by searching for five articles known in advance to be relevant to the review, deriving from health economics (Devlin and Parkin, 2004, Dakin *et al*., 2006), health policy (Mauskopf *et al*., 2013), sociology (Milewa and Barry, 2005) and ethics (Charlton and Rid, 2019). Each of these articles was successfully identified through at least two of the five databases.

The main search (search A) was carried out in October 2019 and used two sets of terms, linked by an AND operator. The first set comprised of NICE's current and past institutional titles,

[1]The recent review of NICE's process and methods for health technology evaluation was initiated in July 2019. Topics of focus for the methods review included the 'modifiers' considered in decision-making (i.e. value-based criteria for exceeding the standard cost-effectiveness threshold such as disease severity, health inequality, rarity and innovation), treatment of uncertainty, types of evidence, health-related quality of life, technology-specific issues (such as the consideration of histology-independent cancer treatments) and discounting. Topics of focus for the process review included overall simplification of the health technology evaluation process, the process for consideration of highly specialised technologies and the negotiation and operation of managed and commercial access agreements. NICE consulted twice on a suite of proposed changes which are now reflected in a new programme manual, which was implement in February 2022 (NICE, 2020*a*, 2022).

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

including common misspellings. The second set included various terms used to describe factors considered during NICE decision-making, identified through key articulations of NICE's approach (Rawlins and Culyer, 2004; NICE, 2005, 2008*b*, 2020*b*) and experience gained from conducting related reviews (Charlton & Rid, 2019; Charlton, 2020). Several broad terms were also included to increase search sensitivity (e.g. 'other factors', 'criteria', 'equity'). During the initial search, it became evident that a small number of articles use the acronym 'NICE' without further definition. A supplementary search (search B) was therefore carried out in November 2019 to identify additional articles whose titles contain only this acronym (see Appendix 1).

Studies were retained if they met the following inclusion criteria:

- an academic article, published in a peer-reviewed journal;
- presents empirical data relating to NICE technology appraisal and
- includes findings that describe what factors appear to substantively influence NICE decision-making, as conducted by its appraisal committees.

No limits were specified regarding date or language.

A parallel search of the grey literature was also conducted which adapted the above strategy for use in Google Scholar. In addition, research articles published by NICE's decision support unit and the UK Office for Health Economics were manually screened for inclusion.

The review was updated in November 2020 according to the same protocol.

2.3. Study selection

Searches A and B together identified 5419 articles, which were compiled in Endnote X9 for desktop. Exact duplicates were automatically removed, with further duplicates removed manually, leaving 2881 potentially eligible articles. These were categorised as either eligible or ineligible based on title, with articles only excluded if the reviewer was confident that they would not satisfy the inclusion criteria. This eliminated a further 2359 articles. The reviewer read the abstract of the remaining 522 articles and categorised them in the same way. Articles not containing an abstract were automatically retained. After this process, 108 potentially eligible articles remained and the full text of each was read to determine final inclusion. This left 20 articles which were deemed to meet the inclusion criteria; however, two of these duplicated data already presented in other articles and were therefore excluded. The references of each of the remaining 18 articles were then hand searched, identifying a further five eligible articles. The review of the grey literature did not yield any further results, giving 23 eligible articles in total.

On updating the search in November 2020, an additional six articles were identified. In total, therefore, 29 articles were included in the review.

2.4. Data extraction and analysis

Following study selection, full texts of the included articles were re-read and key data extracted to an Excel spreadsheet. This recorded: (1) basic bibliographic information (title, author, journal, year); (2) study aim, scope and date range; (3) study methods, (4) a narrative summary of the main findings and (5) the factors observed by the study to have influenced NICE decision-making. Once data extraction was complete, articles identifying particular factors (e.g. cost-effectiveness, uncertainty and innovation) were collated and read for a third time to facilitate narrative synthesis.

3. Results

The search identified 29 eligible studies, details of which are provided in Table 1.

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

Table 1. Included studies

	Study	Title	Method[a]	Aim	Scope
1	Raftery (2001)	NICE: faster access to modern treatments? Analysis of guidance on health technologies	Quantitative: descriptive analysis of retrospective documentary data	To provide an overview of NICE's appraisal decisions up to March 2001	All appraisals, 1999–2001
2	Devlin and Parkin (2004)	Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis	Quantitative: binary choice model using logistic regression analysis	To analyse and visually present a hierarchical set of decision-factors considered in NICE technology assessment	All appraisals, 1999–2002
3	Rawlins and Culyer (2004)[b]	National Institute for Clinical Excellence and its value judgments	Qualitative: purposively selected case studies	To provide an overview of NICE's approach	Pre-2004
4	Milewa and Barry (2005)	Health policy and the politics of evidence	Qualitative: semi-structured interviews and non-participant observation of appraisal committee meetings	To explore whether NICE decision-making is primarily based on evidence and quantitative data, or whether more subtle, less transparent characteristics of context and interaction are also evident in the shaping of decisions	Four appraisals, 2003–2004
5	Milewa (2006)	Health technology adoption and the politics of governance in the UK	Qualitative: semi-structured interviews and non-participant observation of appraisal committee meetings	To understand the tenor and orientation of NICE's deliberations about the adoption of health technologies	Four appraisals, 2003–2004
6	Dakin *et al*. (2006)	'Yes', 'No' or 'Yes, but'? Multinomial modelling of NICE decision-making	Quantitative: multinomial model using logistic regression analysis	To understand the factors influencing NICE decisions and their relative importance	All appraisals, 1999–2003
7	Raftery (2006)	Review of NICE's recommendations, 1999–2005	Quantitative: descriptive analysis of retrospective documentary data	To provide an overview of NICE's appraisal decisions during the first five years of its existence	All appraisals, 1999–2005
8	Bryan *et al*. (2007)	Seeing the NICE side of cost-effectiveness analysis: a qualitative investigation of the use of CEA in NICE technology appraisals	Qualitative: semi-structured interviews, non-participant observation and documentary review	To explore how cost-effectiveness information is used during NICE technology appraisal and to establish how its impact might be increased	Seven appraisals, 2003–2004

(*Continued*)

Table 1. (*Continued.*)

	Study	Title	Method[a]	Aim	Scope
9	Williams *et al.* (2007)	How should cost-effectiveness analysis be used in health technology coverage decisions? Evidence from the National Institute for Health and Clinical Excellence approach	Qualitative: semi-structured interviews, non-participant observation and documentary review	To explore the influence and use of economic evaluation in the decision-making process	Seven appraisals, 2003–2004
10	Tappenden *et al.* (2007)	A stated preference binary choice experiment to explore NICE decision making	Quantitative: stated preference binary choice experiment using logistic regression analysis	To explore whether NICE takes account of concerns other than just incremental cost effectiveness in commissioning health care services	n/a
11	Mason and Drummond (2009)	Public funding of new cancer drugs: is NICE getting nastier?	Quantitative: descriptive analysis of retrospective documentary data	To determine whether NICE is rejecting a higher proportion of cancer drugs and whether the reasons for restricting technologies have changed	Cancer drug appraisals, 2000–2008
12	Clement *et al.* (2009)	Using effectiveness and cost-effectiveness to make drug coverage decisions: a comparison of Britain, Australia, and Canada	Quantitative: descriptive analysis of retrospective documentary data and comparative case studies	To describe how clinical and cost-effectiveness evidence is used in coverage decisions both within and across jurisdictions (UK, Australia and Canada)	Drug appraisals, 2001–2008
13	Rawlins *et al.* (2010)[b]	Pharmacoeconomics: NICE's approach to decision-making	Qualitative: purposively selected case studies	To describe NICE's approach to making scientific and social value judgements	Pre-2009
14	Chalkidou (2012)[b]	Evidence and values: paying for end-of-life drugs in the British NHS	Qualitative: descriptive policy analysis	To consider the triggers for NICE's EOL guidance, the challenges NICE faces in implementing it and the policy's implications for the future role of NICE in the NHS	EOL appraisals, 2009–2011
15	Mauskopf *et al.* (2013)	Drug reimbursement decisions by the National Institute for Health and Clinical Excellence: have they impacted the National Health Service budget?	Quantitative: statistical analysis of retrospective data using descriptive and multivariate logistic analyses	To estimate the correlation between a technology's budget impact and the degree of restrictions on reimbursement recommended by NICE	Drug appraisals, 2001–2011

16	Shah *et al.* (2013)	NICE's social value judgements about equity in health and health care	Qualitative: review of purposively selected policy documents and case studies	To describe the substantive SVJs about equity in health and health care that NICE uses to guide its decision making and to compare NICE policy on these SVJs with NICE practice	Pre-2013
17	Cerri *et al.* (2014)	Decision making by NICE: examining the influences of evidence, process and context	Quantitative: multinomial model using logistic regression analysis	To examine the impact of evidence, process and context factors on NICE decision-making between 2004 and 2009	Adult drug appraisals, 2004–2009
18	Dakin *et al.* (2015)	The influence of cost-effectiveness and other factors on NICE decisions	Quantitative: binary choice model using logistic regression analysis	To investigate the influence of cost-effectiveness and other factors on NICE decisions and whether NICE's decision-making has changed over time	All appraisals, 1999–2011
19	Griffiths *et al.* (2015)	Acceptance of health technology assessment submissions with incremental cost-effectiveness ratios above the cost-effectiveness threshold	Quantitative: descriptive analysis of retrospective documentary data	To identify the key rationale provided by four different HTA agencies when their decisions went against the assumed ICER thresholds, and evaluate any differences between disease areas	Single technology appraisals, 2000–2014
20	Nicod and Kanavos (2016)	Scientific and social value judgements for orphan drugs in health technology assessment	Qualitative: thematic analysis of documents	To explore how aspects of an orphan treatment's value not captured by routine health technology assessment (HTA) may influence HTA processes in different settings	10 orphan drug appraisals, 2006–2012
21	Calnan *et al.* (2017)	Still elegantly muddling through? NICE and uncertainty in decision making about the rationing of expensive medicines in England	Qualitative: semi-structured interviews, non-participant observation and documentary review	To explore the various ways in which different forms of uncertainty are perceived and tackled within NICE single technology appraisals	Three appraisals, 2012–2014
22	De Folter *et al.* (2018)[b]	Decision-components of NICE's technology appraisals assessment framework	Quantitative: automated text analysis of appraisal documents	To analyse and visually present a hierarchical set of decision-factors considered in NICE technology assessment	All appraisals, 2007–2016
23	Charlton and Rid (2019)	Innovation as a value in health care priority-setting: the UK experience	Mixed: quantitative content analysis and thematic analysis of documents	To explore innovation's role as a social value in NICE technology appraisal	Drug appraisals, 1999–2018[c]

(*Continued*)

Table 1. (*Continued.*)

	Study	Title	Method[a]	Aim	Scope
24	Schaefer and Schlander (2018)	Is the National Institute for Health and Care Excellence (NICE) in England more 'innovation friendly' than the Federal Joint Committee (G-BA) in Germany?	Quantitative: descriptive analysis of retrospective data using multivariate linear analysis	To explore whether, and how, different methodological choices are associated with different health technology assessment outcomes in Germany and the UK	All appraisals, 2011–2015
25	Chang (2020)	On cost effectiveness analysis and fairness: normalizing control of and resistance to NICE technology appraisals	Qualitative: critical discourse analysis of case study	To examine to what extent the discourse of the QALY-based CEA constituted and was constituted by power struggles derived from conflicting intentions among the NICE and other stakeholders	Kidney cancer appraisals, 2008–2009
26	Kieslich (2020)	Paradigms in operation: explaining pharmaceutical benefit assessment outcomes in England and Germany	Qualitative: comparative case study using qualitative content analysis	To understand the role played by policy paradigms in influencing the outcomes of HTA processes	10 appraisals, 2011–2012
27	Wood and Hughes (2020)	The new and non-transparent Cancer Drugs Fund	Quantitative: descriptive analysis of retrospective data	To explore the operation of the new Cancer Drugs Fund since its establishment in April 2016	CDF appraisals, 2016–2018
28	Yuasa *et al.* (2021)	Investigation of factors considered by health technology assessment agencies in eight countries	Quantitative: descriptive and correspondence analysis of retrospective data	To understand differences in HTA decision processes and criteria across eight western countries	Cancer and hepatitis drugs, 2012–2019
29	Kleinhout-Vliek *et al.* (2020)[a]	Around the tables – contextual factors in health care coverage decisions across Western Europe	Qualitative: group interviews and workshops and qualitative analysis of purposively selected documents	To understand the use of contextual factors in HTA in Belgium, England, Germany, and the Netherlands	One appraisal, 2017

[a]Where possible, descriptions of the methods used have been taken from the original article.
[b]List of authors includes an individual who was an employee of NICE or held another formal role with the organisation (e.g. Chair, Vice-Chair, Appraisal Committee Chair) at the time of writing.
[c]This analysis excluded non-drug appraisals, terminated appraisals and appraisals that have since been updated or withdrawn and are therefore no longer publicly available.

The included studies were published between 2001 and 2020 and cover NICE appraisals conducted between 1999 and 2019. Across these studies, 11 distinct factors were each observed by at least two studies to have substantively influenced NICE decision-making. These are indexed in Appendix 2.

During the period spanned by this review, NICE's formal approach has evolved considerably and differences in how some factors have been treated over time may reflect these changes. To aid interpretation of the results, key changes in NICE's processes and methods are summarised for reference purposes in Table 2. It should also be noted that aspects of NICE's approach have recently undergone further revision as part of a major update of NICE's processes and methods (see previous footnote). The implications of this update for future research are briefly considered in the discussion.

The following sections summarise the evidence relating to each of the 11 identified factors, starting with NICE's primary substantive consideration: cost-effectiveness.

3.1. Influence of cost-effectiveness on NICE decision-making

Unsurprisingly, research has shown cost-effectiveness to exert considerable influence on NICE decision-making. However, studies also illustrate that the role played by economic evaluation varies for different appraisal committee members and that a technology's incremental cost-effectiveness ratio (its ICER, or cost per QALY) is far from determinative of appraisal outcome.

The earliest study of NICE decision-making, conducted by Raftery in 2001, emphasises the importance of clinical- rather than cost-effectiveness. Of the 19 recommendations by then issued by NICE, all cited clinical benefit as a reason for the technology's adoption; in contrast, only half of the completed appraisals (11/22) reported an ICER, with committees often finding it 'very difficult' or 'impossible' to estimate cost-effectiveness (Raftery, 2001). Raftery concludes from this evidence that economics has a 'lesser role' to play in NICE decision-making than evidence of clinical benefit (*ibid.*). Other work from the same period offers a more nuanced interpretation. Drawing on documentary analysis and interviews relating to seven appraisals conducted during 2002 and 2003, Williams *et al.* (2007) propose that appraisal committee members draw on cost-effectiveness analysis in two distinct ways: either as a general structure for considering and discussing key issues (the 'framework approach'), or as factor to be considered only once clinical value has been demonstrated (the 'ordinal approach'). Under the latter approach, calculation of a technology's ICER may sometimes be unnecessary: in the words of one committee member, 'If it doesn't get through the clinical effectiveness hurdle then I'm not that interested in the economics' (*ibid.*). Other evidence supports the hypothesis that appraisal committees are often able to reach a decision without calculating a technology's ICER, with studies indicating that between 5 and 48% of NICE's decisions were made in this way in the years up to 2011 (Devlin and Parkin, 2004; Dakin *et al.*, 2006; Cerri *et al.*, 2014; Dakin *et al.*, 2015) (Table 3).

When ICERs *are* calculated as part of the decision-making process, quantitative research has repeatedly demonstrated a strong correlation with decision outcome. Dakin *et al.*'s study of appraisals completed by December 2011 – the largest and most recent retrospective analysis of all NICE decisions – estimates that technologies costing £40,000/QALY have a 50% chance of recommendation, compared with 75% at £27,000/QALY and 25% at £52,000/QALY (Dakin *et al.*, 2015). Other studies similarly demonstrate a clear correlation between a technology's estimated cost-effectiveness and its likelihood of recommendation, with the average ICER of recommended technologies found to be substantially lower than that of technologies that are rejected, and the ICERs of technologies in which recommendation is restricted to a subgroup of patients typically coming somewhere between the two (Dakin *et al.*, 2006; Cerri *et al.*, 2014; Griffiths *et al.*, 2015; Schaefer and Schlander, 2018).

It therefore seems reasonable to conclude from the available evidence that cost-effectiveness has historically played a major, if not always essential, role in NICE decision-making.

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

Table 2. Evolution of NICE methods, key events 1999–2021

Date	Event
April 1999	Establishment of NICE
December 1999	*Publication of interim methods guide (first edition)* *Establishes that appraisal will involve consideration of:* • clinical effectiveness; • *cost effectiveness and* • *wider NHS implications*.
June 2001	Publication of methods guide (second edition) *Provides further technical guidance on the conduct of technology appraisal and confirms NICE's general preference for cost-effectiveness or cost–utility analysis.*
April 2004	Publication of methods guide (third edition) *Sets out NICE's use of ICERs and specifies a range of £20,000–30,000/QALY, within which recommendations are likely to make explicit reference to the following factors:* • *The degree of uncertainty surrounding the calculation of ICERs;* • *The innovative nature of the technology;* • *The particular features of the condition and population receiving the technology and* • *Wider societal costs and benefits*.
December 2005	Publication of Social Value Judgements: Principles for the development of NICE guidance (first edition) *Describes the social value judgements that should generally be incorporated into NICE guidance. Comprises thirteen principles which stipulate, amongst other things, that:* • *NICE guidance should not support the use of interventions for which evidence of clinical effectiveness is either absent or too weak for reasonable conclusions to be reached;* • *Economic consideration must be taken into account and cost–utility analysis is therefore necessary, but should not be considered the sole basis for decisions on cost-effectiveness and* • *NICE guidance should explain, explicitly, the committee's reasons for recommending a technology whose ICER exceeds £20,000–30,000/QALY.*
September 2006	Introduction of single technology appraisal (STA) *Establishes a new process for the appraisal of single technologies, as distinct from the multiple technology appraisal (MTA) process that was previously exclusively used. Intended to reduce delays in initiating appraisal and accelerate the appraisal process*.
June 2008	Publication of methods guide (fourth edition) *Further specifies how committees should respond to several factors:* • *Committees should exercise more caution above an ICER of £20,000/QALY when there is less certainty about the ICER.* • *Above an ICER of £20,000/QALY, committees should take into account a technology's innovative nature if the innovation adds substantial demonstrable and distinctive benefits which may not be fully reflected in the technology's ICER.* • *Budget impact should not determine the committee's decision but, in general, the committee will want to be increasingly certain of the cost effectiveness of a technology as the impact of its adoption on NHS resources increases*.
July 2008	Publication of Social Value Judgements: Principles for the development of NICE guidance (second edition) *Updates NICE's social value judgements, consolidating and somewhat modifying the principles set out in the first edition*.
July 2009	Publication of addendum on appraising life-extending, end of life treatments *Supplementary advice specifying that committees should consider the impact of giving greater weight to QALYs achieved in the later stages of terminal disease, when:* • *The treatment is indicated for patients with a short life expectancy, normally less than 24 months;*

(*Continued*)

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

Table 2. (*Continued.*)

Date	Event
	• *There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment, and;* • *The treatment is licensed or otherwise indicated for small patient populations.*
July 2011	Publication of clarification on discounting *Stipulates that where a technology's treatment effects are both substantial in restoring health and sustained over a very long period (normally at least 30 years), a discount rate of 1.5% for health effects and 3.5% for costs should be used for sensitivity analysis. Standard policy remains to discount both costs and health effects at 3.5%.*
April 2013	Publication of methods guide (fifth edition) *Formally incorporates the end-of-life rules and clarification on discounting into NICE's methods guide.*
May 2013	Establishment of the highly specialised technologies (HST) programme *Dedicated programme for the appraisal of ultra-orphan drugs for very rare diseases. Interim methods guide sets out six criteria for appraisal:* • *Nature of the condition* • *Impact of the new technology* • *Cost to the NHS and personal social services* • *Value for money* • *Impact of the technology beyond direct health benefits* • *Impact of the technology on the service delivery* *The guidance does not stipulate a cost-effectiveness threshold and does not require that decisions be based on calculation of a technology's ICER.*
April 2016	Publication of addendum on amendments to NICE methods *Sets out proposed changes to NICE technology appraisal in support of arrangements for the management of the new cancer drugs fund (CDF). Stipulates that when uncertainty is too great to recommend a drug for routine use, the committee can consider a recommendation for use within the CDF if uncertainty can be addressed through additional data collection and there is plausible potential for satisfying the criteria for routine use, taking into account the application of the end-of-life rules.* *Also includes amendment of the end-of-life rules, such that:* • *small patient population size is no longer a criterion for application, and* • *the maximum additional weight given to QALYs achieved at the end of life is specified as 1.7, implying a cost-effectiveness threshold of approximately £50,000/QALY.*
April 2017	Introduction of fast-track appraisal process *Process amendment to allow for accelerated and less resource-intensive appraisal of technologies for which the ICER is expected to be less than £20,000/QALY.*
	Introduction of budget impact test *Process amendment whereby technologies with an expected annual budget impact exceeding £20 million may be subject to an extended appraisal timeline to allow for commercial negotiation and phased adoption if the technology is recommended.*
	Publication of updated methods for HST programme *Updates the criteria for the appraisal of HSTs and introduces a QALY weighting formula to be used in making judgements about the acceptability of technologies whose ICER exceeds £100,000/QALY. Under these rules, HSTs will generally be recommended where the ICER is < £100,000/QALY and may be recommended at an ICER of up to £300,000/QALY if the incremental QALY gain (i.e. health benefit) is sufficiently high.*
February 2022	Implementation of updated methods and processes *To put into effect changes identified through a review of methods and processes initiated in July 2019.*

However, it does not follow that cost-effectiveness is the principal determinant of *all* NICE decisions, or that it is the *only* factor taken into consideration in most cases. Work by Dakin *et al.* and others implies a cost-effectiveness threshold somewhat higher than the £20,000–30,000/QALY

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

range generally suggested by NICE policy, indicating committees' willingness to recommend seemingly cost-ineffective technologies when this is justified by other considerations (Tappenden *et al.*, 2007; Rawlins *et al.*, 2010; Dakin *et al.*, 2015; Griffiths *et al.*, 2015) (Table 3). We now turn to these other considerations.

3.2. Other factors shown to substantively influence NICE decision-making

3.2.1. Uncertainty

Studies from the mid-2000s suggest that appraisal committees have long been sensitive to uncertainty about a technology's expected impacts and have historically shown a reluctance to fully recommend technologies that pose a substantial risk to the NHS. More recent work illustrates the continued ubiquity of uncertainty to NICE decision-making but is unable to address emerging questions about how changes to NICE's approach have modified committees' response to risk.

In 2005, Tappenden *et al.* conducted a binary choice experiment involving 37 past and present appraisal committee members to explore their preferences in deciding which technologies to recommend. This found that members were 69% less likely to recommend a technology when uncertainty about its cost-effectiveness was 'high' compared with when it was 'low' – an effect that was particularly pronounced when the hypothetical technology's ICER exceeded £25,000/QALY (Tappenden *et al.*, 2007).[2] This finding suggests that committees at the time were minded to follow NICE's advice, first issued in 2004, that when considering technologies at or beyond the cost-effectiveness threshold, special consideration should be given to 'the degree of uncertainty surrounding the calculation of ICERs' (NICE, 2004; Rawlins and Culyer, 2004). Further evidence of early committees' sensitivity to risk is provided by Raftery (2006), who found that around two-thirds of the rejections issued by NICE up to 2005 were on the grounds of insufficient evidence, and by Devlin and Parkin (2004), who observed that even technologies with relatively low ICERs were, on occasion, rejected where uncertainty was high. The role played by clinical vs financial risk in these decisions, and the extent to which the 'ordinal' approach might be understood as a way of assessing these different types of risk in turn, are outstanding questions.

Later studies have indicated a greater willingness by committees to recommend technologies about which there is significant uncertainty. An analysis by Clement *et al.* (2009) found that nearly half of decisions (46%) made up to December 2008 were subject to 'considerable' uncertainty about cost-effectiveness, but that committees nevertheless chose to recommend the technology in 87% of cases. The authors suggest that this apparent tolerance of uncertainty may reflect an approach in which appraisal committees seek to mitigate risk by restricting recommendations to patient subgroups, rather than fully rejecting technologies for which the evidence base is relatively weak (*ibid.*). Further evidence for this hypothesis is provided by Cerri *et al.* (2014), who found that technologies recommended for routine use between 2004 and 2009 were supported by substantially more robust clinical evidence – in terms of the number of randomised clinical trials conducted, their size, duration and design, and the size of the observed effect – than technologies that were recommended for restricted use, or those that were rejected. More recently still, qualitative research by Kieslich (2020), based on appraisals conducted in 2011 and 2012, has demonstrated NICE's willingness to rely on anecdotal evidence from clinical experts where 'gold standard' clinical trial evidence is lacking.

An apparently consistent feature throughout NICE's work has been the continual need to acknowledge and respond to uncertainty. An automated text analysis of appraisal documents published between 2007 and 2016 observed that terms relating to uncertainty arose in association with almost all of the 125 'decision factors' found to feature in committee discussions, demonstrating its pervasiveness across nearly all aspects of NICE decision-making (de Folter *et al.*, 2018).[3] More in-depth research by Calnan *et al.* (2017) confirms this finding, with the authors

[2]The original study did not provide participants with any further specification of 'high' and 'low' uncertainty.

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

Table 3. Summary of quantitative studies

Study	Method	Scope	No. of appraisals *(decisions)*[a] in scope	Appraisal/decision outcome	No. of appraisals/ decisions in which ICER not used for decision-making (%)	Implied threshold (£/QALY)	Influence of other factors
Raftery (2001)	Retrospective descriptive analysis of appraisal documentation	All appraisals up to Mar-01	22 appraisals *(decisions not given)*	19/22 (86%) recommended; 3/22 (14%) rejected	11/22 appraisals (50%)	30,000	'In only half the topics did the NICE guidance cite cost per QALY […], suggesting that economics had a lesser role than evidence of **clinical benefits**'.
Devlin and Parkin (2004)	Binary choice model using logistic regression analysis	All appraisals up to May-02	39 appraisals *(51 decisions)*	36/51 (71%) recommended; 15/51 (29%) rejected	18/51 decisions (35%)	>30,000	'NICE decisions are well explained by the cost-effectiveness evidence, with the effect of **uncertainty** and of the **burden of disease** explaining the rejection of some technologies with a relatively low ICER and the acceptance of some with a relatively high ICER'.
Dakin *et al*. (2006)	Multinomial model using logistic regression analysis	All appraisals up to Dec-03	73 appraisals *(94 decisions)*	82/94 (87%) recommended; 12/94 (13%) rejected *Of those recommended:* • *20/94 (21%) routine use* • *62/94 (66%) restricted use*	48% decisions[b]	Not given	'Results suggest that interventions supported by **more randomised trials** are more likely to be recommended and endorsed for routine use. Higher cost-effectiveness ratios increased the likelihood of interventions being rejected rather than recommended for restricted use but did not significantly affect the decision between routine and restricted use. **Pharmaceuticals, interventions appraised early in the NICE programme and those with more systematic reviews** were also less likely to be rejected, while **patient group submissions** made a recommendation for routine rather than restricted use more likely'.
Raftery (2006)	Retrospective descriptive analysis of appraisal documentation	All appraisals up to Apr-05	86 appraisals *(117 decisions)*	95/117 (81%) recommended; 22/117 (19%) rejected *Of those recommended:* • *27/117 (23%) routine use*	Not given	Not given	'The highest cost per QALY that NICE has accepted is an estimated £39 000 (range £35,000–43,000) for riluzole to treat motor neurone disease. […] Although NICE does not officially prioritise interventions that save lives over those that improve quality of life, its treatment of some topics suggests the **rule of rescue**, or prioritising **life saving therapies**, may play a part. With cancer drugs such as imatinib and trastuzumab,

(*Continued*)

Table 3. (*Continued.*)

Study	Method	Scope	No. of appraisals *(decisions)*[a] in scope	Appraisal/decision outcome	No. of appraisals/ decisions in which ICER not used for decision-making (%)	Implied threshold (£/QALY)	Influence of other factors
				• *68/117 (58%) restricted use*			which **extend life expectancy**, NICE accepted relatively poor cost effectiveness. However, the acceptance of riluzole was based on considerations of **quality of life** rather than on mortality'.
Tappenden *et al.* (2007)	Stated preference binary choice experiment	n/a – survey of 37 committee members					'[…] increases in the incremental cost-effectiveness ratio and **economic uncertainty**, and the **availability of other therapies** was associated with statistically significant reductions in the odds of a positive recommendation ($p < 0.01$). The transition from a very low to a comparatively high level of **baseline HR-QOL** [health-related quality of life] was also associated with a statistically significant reduction in the odds of a positive recommendation'.
Clement *et al.* (2009)	Descriptive analysis of retrospective documentary data and comparative case studies	All appraisals involving drugs, up to Dec-08	97 appraisals *(199 decisions)*	174/199 (87%) recommended; 25/199 (13%) rejected	Not given	Not given	'The data suggest that the 3 agencies make recommendations that are consistent with evidence on **effectiveness** and cost effectiveness but that **other factors** are often important. […] Significant **uncertainty around clinical effectiveness**, typically resulting from inadequate study design or the use of inappropriate comparators and unvalidated surrogate end points, was identified as a key issue in coverage decisions'.
Mason and Drummond (2009)	Retrospective descriptive analysis of appraisal documentation	All appraisals of cancer drugs, up to Oct-08	38 appraisals *(56 decisions)*	47/56 (84%) recommended; 9/56 (16%) rejected *Of those recommended:* • 31/56 (55%) routine use • *16/56 (29%) restricted use*	Not given	Not given	'The higher rejection rate for cancer drugs [after 2006] is partly explained by the **new appraisal process** [i.e. single technology appraisal], but the principal reason for the observed change is the shift from an **absence of evidence on cost-effectiveness** to **evidence of an absence of value-for-money**'.

Mauskopf *et al.* (2013)	Statistical analysis of retrospective data using descriptive and multivariate logistic analyses	Appraisals involving drugs, Jan-01 to Mar-11	97 appraisals *(144 decisions)*	112/144 (78%) recommended; 32/144 (22%) rejected *Of those recommended:* • *51/144 (36%) routine use* • *61/144 (42%) restricted use*	Not given	Not given	'After controlling for clinical effectiveness and cost-effectiveness, the degree of reimbursement restriction recommended by NICE remains significantly correlated with the PBI [**potential budget impact**], despite that fact that the NICE decision process does not consider budget impact'.
Cerri *et al.* (2014)	Multinomial model using logistic regression analysis	Adult drug appraisals only, 2004-2009	65 appraisals *(118 decisions)*	101/118 (86%) recommended; 17/118 (14%) rejected *Of those recommended:* • *32/118 (27%) routine use* • *69/118 (58%) restricted use*	95% decisions[c]	Not given	'The multinomial model showed significant associations ($p < 0.10$) between NICE outcome and four variables: (i) **demonstration of statistical superiority of the primary endpoint** in clinical trials by the appraised technology; (ii) the incremental cost-effectiveness ratio (ICER); (iii) the **number of pharmaceuticals appraised within the same appraisal**; and (iv) the **appraisal year**'.
Dakin *et al.* (2015)	Binary choice model using logistic regression analysis	All appraisals up to Dec-11	229 appraisals *(763 decisions)*	Not given	161/763 decisions (21%)[d]	>30,000[e]	'Our analyses demonstrate that cost-effectiveness is the principal determinant of most NICE decisions and that the probability of rejection increases significantly with increasing ICER. […] The single factor other than cost-effectiveness that emerged from our analyses as exerting a significant effect on decisions is the **type of disease** that the technology is intended to prevent, diagnose or treat. NICE rejections were significantly less likely for cancer and musculoskeletal disease but more likely for respiratory disease'.

(*Continued*)

Table 3. (*Continued.*)

Study	Method	Scope	No. of appraisals *(decisions)*[a] in scope	Appraisal/decision outcome	No. of appraisals/ decisions in which ICER not used for decision-making (%)	Implied threshold (£/QALY)	Influence of other factors
Griffiths *et al.* (2015)	Retrospective descriptive analysis of appraisal documentation	All appraisals for which ICERs were reported, up to May-14	114 appraisals *(decisions not given)*	87/114 (76%) recommended; 27/114 (24%) rejected *Of those recommended:* • *66/114 (58%) routine use* • *21/114 (18%) restricted use*	Excluded from analysis	Not given	'NICE recommended the highest proportion of submissions with ICERs higher than the threshold (34% accepted without restrictions; 20% with restrictions) […] Reasons for accepting submissions reporting ICERs above the threshold included **high clinical benefit** over the standard of care, and addressing an **unmet therapeutic need**'.
Shaefer and Schlander (2019)	Descriptive analysis of retrospective data using multivariate linear analysis	All single technology appraisals, Sep-06 to Apr-15	88 appraisals *(125 decisions)*	99/125 (79%) recommended; 26/125 (21%) rejected	Not given	Not given	'Appraisals for technologies with unspecified ICERs were primarily based on **clinical effectiveness** as well as **uncertainty** in the estimated ICER (range)'.
Yuasa *et al.* (2020)	Retrospective descriptive analysis of appraisal documentation	Appraisals of cancer and hepatitis C drugs, 2012 to Aug-19	25 appraisals *(decisions not given)*	Not given	None	Not given	Factors observed to play a role in NICE appraisals were: **clinical benefit, uncertainty, disease severity, rare/ orphan status, disadvantaged population, unmet needs, issues with alternative treatment, complex pathways of treatment, stakeholder persuasion, indirect benefits, innovation and fear of contagion.**

[a]Multiple technology appraisals comprise multiple decisions (about different technologies and/or indications) within the same appraisal. Hence, the number of distinct decisions analysed in many studies exceeds the number of appraisals included.
[b]In the original paper (Table 2), this figure represents the total proportion of decisions for which a cost–utility analysis was present. Appraisals were considered to have a cost–utility analysis if they gave an estimate of the cost per QALY gained.
[c]As in Dakin *et al.* (2006), this figure represents the total proportion of decisions for which a cost–utility analysis was present. However, the authors do not specify on what basis a cost–utility analysis was considered to have been performed.
[d]As per Figure 1 of the original paper: 'Seventy decisions were "no" as a result of clinical evidence […] Sixty-three decisions were "yes" on clinical grounds (e.g. because all alternative technologies were contraindicated or not tolerated), while 28 decisions were "no" on clinical grounds (e.g. because treatment was "clinically inappropriate" in that patient group)'.
[e]The original paper states: 'We estimate that, in practice, the ICER at which the probability switches from more-likely-to accept to more-likely-to-reject is between £39,000 and 44,000: well above the stated £20,000–30,000 range' (Dakin et al., 2015).

observing that committees' difficulties in dealing with different types of uncertainty across three 2012–2014 appraisals rendered straightforward decision-making 'problematic'. Calnan *et al.* identify several pragmatic strategies adopted by committees in trying to address such difficulties. These include explicit attempts to measure uncertainty and focus attention on areas about which there can be more confidence, as well as implicit approaches based, for example, on 'gut feeling' and the collective bypassing of certain uncertainties in order to reach a decision (employing the 'fudge factor'). According to one committee member, 'if there feels like there's a lot of unresolved uncertainty, then we're more conservative in our estimate of what we think the ICER is going to be', suggesting that considerations about uncertainty and cost-effectiveness interact in ways that are difficult to unravel and may exaggerate the apparent role of cost-effectiveness (*ibid.*). NICE's increasing use of 'managed access' – an arrangement in which a technology's recommendation is made conditional on additional data collection – adds a layer of complexity to this relationship between uncertainty and cost-effectiveness and raises further unanswered questions about appraisal committees' response to different types of risk.

3.2.2. Budget impact

NICE has long been clear that it considers affordability to be a concern primarily for politicians rather than itself (Timmons *et al.*, 2016) and, since 2008, it has stated as policy that a technology's potential budget impact 'does not determine' whether or not it will be recommended (NICE, 2008*a*, 2013). Nevertheless, this policy also advises appraisal committees that they should be 'increasingly certain' of a technology's ICER as its impact on NHS resources increases, and evidence suggests that committees have tended to follow this advice (NICE, 2008*a*, 2013, 2017*b*).

In their retrospective analysis of decisions made up to December 2003, Dakin *et al.* (2006) found budget impact to be secondary only to cost-effectiveness and clinical uncertainty in its ability to predict decision outcome, with the total potential cost to the NHS observed to be significantly higher for technologies eventually recommended for restricted use than for those recommended for routine use. A similar effect has also been observed by other studies (Mauskopf *et al.*, 2013; Cerri *et al.*, 2014), suggesting that appraisal committees may use restricted recommendations as a way of reducing total cost when a technology's potential impact on NHS resources is high. According to an analysis by Mauskopf *et al.* (2013), after controlling for clinical- and cost-effectiveness, the average potential budget impact for drugs appraised up to April 2011 ranged from £20.3 million for fully recommended drugs, to £49.8 million for drugs that were recommended with restrictions, to £71.1 million for drugs that were wholly rejected.

Although these quantitative studies strongly suggest that budget impact plays a substantive role in NICE decision-making, they are not able to provide any insight into where the normative basis for this role lies. For example, it is unclear whether committees' prudence in recommending technologies with high budget impact primarily reflects a concern for affordability (i.e. net cost to the NHS), or for the risk that such technologies pose to the system (a function of both net cost and uncertainty about their likely effects). Similarly, it is unclear whether committees' willingness to fully recommend technologies with a relatively low budget impact reflects a concern with affordability/risk or is evidence of an allocative preference for small population size in itself (i.e. rarity). The current evidence base is also unable to provide any insight into the impact on NICE decision-making of the 'budget impact test' introduced in 2017: a measure intended to identify technologies whose high net cost might necessitate further commercial negotiation and, in some cases, delayed adoption, however cost-effective they may be (Charlton *et al.*, 2017; NICE, 2018).

[3]The original text is not explicit about how 'decision factors' are defined but describes them as considerations that are 'involved in the decision making by NICE's advisory committees' (de Folter *et al.*, 2018). The presence of a decision factor in appraisal documentation does not necessarily imply that it influenced the decision of the appraisal committee.

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

3.2.3. Clinical need

Research has explored the role of several considerations associated with the clinical need addressed by an appraised technology, including disease severity, life expectancy, the availability of alternative treatments, baseline quality-of-life and therapeutic area. However, NICE's advice to its appraisal committees regarding these types of consideration has evolved substantially over time, contributing to a mixed and incomplete picture of this factor's influence on NICE decision-making.

In its 1999 directions from the Government, NICE was advised that its recommendations should have regard to 'the degree of clinical need of the patients with the condition under consideration' (NICE, 1999). Early studies highlight several appraisals in which committees demonstrably followed this advice, recommending technologies that addressed significant clinical need despite relatively high ICERs.[4] Appraisal committee members' willingness to prioritise technologies based on related considerations is also illustrated by the binary choice experiment conducted by Tappenden *et al.* in 2007, which found that members were more likely to recommend hypothetical technologies when baseline health-related quality-of-life was low and alternative treatment options were unavailable.

Quantitative studies from this period, however, provide little evidence for NICE's systematic prioritisation of technologies based on clinical need. Three large retrospective studies that specifically explored the relationship between the availability of alternative treatments and decision outcome failed to find any significant association between the two (Devlin and Parkin, 2004; Dakin *et al.*, 2006; Cerri *et al.*, 2014). Dakin *et al.*'s more up-to-date examination of decisions reached prior to 2012 identified a correlation between outcome and therapeutic area, but this relationship does not straightforwardly map onto clinical need; while indication for cancer was associated with increased odds of recommendation, treatments for musculoskeletal disease received even more favourable treatment (Dakin *et al.*, 2015).

In 2009, NICE's policy that its committees should give general consideration to clinical need was supplemented by the more specific advice that, when appraising potentially life-extending treatments for terminal diseases, committees should consider giving greater weight to QALYs achieved at the end of life (NICE, 2009). In effect, this increased the cost-effectiveness threshold for technologies meeting the 'end-of-life' criteria to £50,000/QALY: a figure that was formalised in NICE's methods in 2016 (NICE, 2016; Charlton, 2020). Also in 2016, NICE became responsible for the operation of the new Cancer Drugs Fund (CDF), an instrument that enables patients to access cancer drugs that have failed to meet NICE's cost-effectiveness requirements (NICE, 2016).

Two studies have examined in detail the events that preceded the introduction of the end-of-life criteria, concluding that they came about in large part because of NICE's inability to recommend an emerging cohort of expensive oncology drugs under its standard methods (Chalkidou, 2012; Chang, 2020). Support for this version of events is provided by Mason and Drummond (2009), whose research shows an increase in the rejection rate for cancer drugs in the years leading up to the change, from 11% between 1999 and June 2006, to 26% between June 2006 and October 2008. However, evidence on the actual effect of the end-of-life criteria is limited. In evaluating a subset of appraisals completed between January 2009 and December 2011, Dakin *et al.* found that technologies assessed under the new criteria were 3.4 times more likely to be recommended than those that were not. However, the overall rate of cancer drug recommendation actually fell during this period compared with appraisals conducted prior to 2009,

[4]These include the 2001 recommendation of riluzole (Rilutek), a quality-of-life improving treatment for patients with motor neurone disease; the recommendation of 'first-in-class' oncology drugs imatinib (Glivec) and trastuzumab (Herceptin) in 2002; the recommendation of omalizumab (Xolair) for use in patients with severe asthma in 2007 and the 2008 recommendation of pemetrexed (Alimta) for the treatment of malignant pleural mesothelioma, a terminal condition caused primarily by exposure to asbestos (Raftery, 2001, 2006; Rawlins and Culyer, 2004; Rawlins *et al.*, 2010; Shah *et al.*, 2013; Nicod and Kanavos, 2016).

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

suggesting that the end-of-life criteria may simply have formalised something that appraisal committees were already considering (Dakin *et al.*, 2015). More recent work does however demonstrate the regularity with which the criteria are now used to facilitate the recommendation of cancer drugs. According to Wood and Hughes, between April 2016 and March 2018, around half of all routine recommendations of cancer drugs (32/70, 46%) and a third of recommendations made through the CDF (14/42, 33%) applied the enhanced threshold permitted by the end-of-life criteria (Wood and Hughes, 2020).

Evidence suggests that the new CDF has also played a significant role in facilitating the recommendation of cancer drugs. According to NICE's own figures, since the CDF's introduction in April 2016, the rate of approval for cancer drugs has increased from 59 to 74%: a relative increase of 25% (NICE, 2020*c*). However, it is unclear how appraisal committees exercise judgement in their application of either the end-of-life criteria or in their recommendation of drugs as part of the CDF, or the extent to which these formal instruments represent the totality of committees' concern with clinical need. Given the increasingly dominant position of cancer drugs in NICE's programme of work (NICE, 2020*c*), such questions represent a potentially significant line of future research.

3.2.4. Innovation

Since 2008, NICE has advised its appraisal committees to take special account of a technology's 'innovative nature' in deciding whether it warrants recommendation beyond the usual cost-effectiveness threshold (NICE, 2008*a*, 2013) and research indicates that consideration of a technology's innovativeness does regularly enter into NICE decision-making. Questions remain, however, about how appraisal committees define innovation and the extent to which their concern for this factor overlaps with – and is potentially derived from – concern for other factors such as uncertainty, clinical need and rarity.

The first study to consider the influence of innovation was Dakin *et al.*'s retrospective study of appraisals completed by December 2011, which classed technologies as either innovative or non-innovative based on the time since their commercial launch, the drug class to which they belonged and whether or not they were pharmaceuticals (Dakin *et al.*, 2015).[5] This found 'innovation' to be one of the several factors weakly correlated with appraisal outcome, but highlighted the challenge of exploring the influence of factors that are undefined by NICE and difficult to measure empirically (*ibid.*).

Stronger evidence for the influence of innovation is provided by Charlton and Rid (2019), who establish through both qualitative and quantitative analysis of appraisal documentation that considerations about a technology's 'innovativeness' played a meaningful role in almost half of the drug appraisals completed between 2000 and mid-2018 (151/320, 47%). In 26/320 instances (8%), this role extended to innovation being explicitly invoked by committees – alongside other factors – to support a technology's recommendation beyond £20,000/QALY (*ibid.*). This study also identifies significant inconsistencies in how committees define and value innovation and highlights a substantial increase in committees' concern for this factor since 2008 (*ibid.*). This suggests that committees' consideration of innovation, unlike that of clinical need, has been prompted in large part by NICE's advice to do so, which was issued that year. The influence of innovation has also been highlighted by other recent studies (Kieslich, 2020; Yuasa *et al.*, 2021) and by de Folter *et al.*'s automated text analysis, which found that references to innovation were made in around 80% of all appraisals published between January 2007 and December 2018 (de Folter *et al.*, 2018). Additional work is needed to further explore appraisal committees' understanding of this concept, its relationship with other normative considerations and the actual substantive role that it plays in decision-making.

[5]Any molecule launched within 2 years of appraisal and in an ATC4 class that was created within 5 years of the appraisal. Non-pharmaceutical interventions were classed as non-innovative.

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

3.2.5. Rarity

NICE has historically advised its appraisal committees that they should 'evaluate drugs to treat rare conditions […] in the same way as any other treatment' (NICE, 2008*b*, 2020*b*). However, since 2013, 'ultra-orphan' drugs for very rare diseases have been systematically prioritised through the operation of NICE's highly specialised technologies (HST) programme (NICE, 2017*a*), raising questions about the role played by rarity in NICE's current approach and the normative basis for NICE's prioritisation of ultra-orphan drugs.

Prior to 2013, two quantitative studies which specifically explored the relationship between a technology's 'orphan' status and its likelihood of recommendation failed to find any statistically significant association (Cerri *et al.*, 2014; Dakin *et al.*, 2015). However, even prior to the establishment of the HST programme, qualitative evidence indicates that considerations about a condition's rarity have occasionally proved influential. In their study of 10 orphan drugs appraised by NICE between 2006 and 2012, Nicod and Kanavos (2016) identify five cases in which they consider rarity to have acted as a 'pivotal factor' in decision-making. In four of these five cases, concern for rarity was a function of the end-of-life criteria, which at the time required that the technology be indicated for a small patient population (NICE, 2009). However, this and other studies identify several other cases – all cancer drugs – in which rarity appears to have acted as a standalone basis for special treatment.[6] Indeed, de Folter *et al.*'s analysis suggests that consideration of a condition's rarity may be a fairly regular aspect of committee discussions, featuring in around 20% of appraisals (de Folter *et al.*, 2018).

Notably, the HST programme had not, at the time of the review, been the subject of any published empirical research, despite having been in operation for over 8 years.[7] The substantive role played by rarity in recent NICE decision-making (both in relation to HSTs and other technologies), the normative basis for this role, and rarity's relationship with other factors such as uncertainty, budget impact, clinical need, innovation and age, are therefore matters about which significant questions remain.

3.2.6. Age

NICE's methods do not formally vary based on the age of those who will benefit from a technology's adoption and, given that age is a protected characteristic under the 2010 Equality Act, it is not clear that it would be legal for them to do so (Government, 2010). However, since 2011 an amendment to NICE's formal methods has allowed technologies that offer very substantial health benefits over a period of at least 30 years to be assessed using a lower than usual discount rate.[8] This generally has the effect of lowering the ICER, typically of technologies indicated for severely ill young people.

The two studies to have quantitatively explored the relationship between patient age and decision outcome did not find any statistically significant correlation (Tappenden *et al.*, 2007; Dakin *et al.*, 2015). However, an article published by three senior NICE members in 2010 identified paediatric indication as one of the six 'special circumstances' in which committees might be willing to exceed the usual cost-effectiveness threshold (Rawlins *et al.*, 2010). Rawlins *et al.* highlight

[6]These include pemetrexed for the treatment of malignant pleural mesothelioma (Alimta, appraised in 2008), sunitinib for the treatment of renal cell carcinoma (Sutent, appraised in 2009), mifamurtide for the treatment of osteosarcoma (Mepact, appraised in 2011) and lenvatinib and sorafenib for the treatment of differentiated thyroid cancer (Lenvima/Nexavar, appraised in 2018) (Shah *et al.*, 2013; Nicod and Kanavos, 2016; Chang, 2020; Wood and Hughes, 2020).

[7]One paper meeting this description has since been published by the current author (Charlton, 2022).

[8]According to the 2013 NICE methods guide: 'In cases when treatment restores people who would otherwise die or have a very severely impaired life to full or near full health, and when this is sustained over a very long period (normally at least 30 years), cost-effectiveness analyses are very sensitive to the discount rate used. In this circumstance, analyses that use a non-reference-case discount rate for costs and outcomes may be considered. A discount rate of 1.5% for costs and benefits may be considered by the Appraisal Committee if it is highly likely that, on the basis of the evidence presented, the long-term health benefits are likely to be achieved'. The usual discount rate used for both costs and benefits is 3.5% (NICE, 2013).

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

two specific cases[9] in which such circumstances had been recognised, explaining that NICE 'understands that society would generally favour "the benefit of the doubt" being afforded to sick children' (*ibid.*). Quotes from committee members interviewed during other studies indicate a similar tendency, with members stating that they are inclined to 'give […] more weight', or be 'softer at the edges', when considering the value of paediatric treatments (Bryan *et al.*, 2007; Calnan *et al.*, 2017). More recently, de Folter *et al.*'s analysis found that consideration of 'children' entered into committee discussions in around 20% of the appraisals completed between 2007 and 2016 (de Folter *et al.*, 2018). The application of the special discounting rules introduced in 2011, however, has not been the subject of any empirical research and, given the current legal landscape, appraisal committees may be reluctant to acknowledge any consideration of patient age in their decision-making. As such, further in-depth qualitative research would likely be needed to ascertain what role – if any – considerations of patient age play in NICE technology appraisal.

3.2.7. Cause of disease

As in the case of age, consideration of cause of disease is not formally incorporated into NICE's approach. Specifically, NICE's current principles prohibit it from 'alter[ing] its normal approach because a condition may have been caused by the person's behaviour' (NICE, 2020*b*). Nevertheless, there is evidence to indicate that cause of disease has been considered where fault can be attributed to a third party. Three studies highlight the 2006 case of pemetrexed (Alimta) for the treatment of malignant pleural mesothelioma (MPM), which was recommended beyond the usual cost-effectiveness threshold partly due to the well-established link between MPM and occupational exposure to asbestos (Rawlins *et al.*, 2010; Chalkidou, 2012; Shah *et al.*, 2013).[10] Two studies also cite the 2002 case of imatinib, which was recommended at an ICER of £49,000/QALY for patients in the blast phase of chronic myeloid leukaemia because these patients would have been offered the drug earlier in disease progression were it not for 'failings in the healthcare system' (Rawlins and Culyer, 2004; Chalkidou, 2012). No further cases are identified in the current literature. However, given the small number of appraisals covered by the type of in-depth qualitative studies capable of identifying such occasional considerations, further research could feasibly identify additional cases.

3.2.8. Wider societal impacts

Although NICE's general approach is to take a relatively narrow 'health-only' perspective in assessing the costs and benefits of a technology's adoption (NICE, 2013), wider societal impacts can be taken into account on an exceptional basis and several studies identify instances in which committees have chosen to do so. Charlton and Rid (2019), for example, highlight the 2015 case of ledipasvir–sofosbuvir, in which the committee appears to have given weight to the 'improved earning capacity' of treated vs untreated hepatitis C patients.[11] Nicod and Kanavos (2016) similarly identify patients' 'ability to contribute to society' as a key consideration in NICE's 2011 appraisal of mifamurtide, a drug for the treatment of osteosarcoma (bone cancer) in children and young people. The societal value of equality appears to be a particularly common

[9]These were somatotropin to treat growth hormone deficiency and chronic subcutaneous insulin infusion for the treatment of type 1 diabetes in childhood.

[10]According to Shah *et al.*, when this aspect of decision-making was subjected to scrutiny at appeal, the appraisal committee argued that 'matters relating to the cause of the disease had been noted […] but were not actually factors in recommendation' (Shah *et al.*, 2013). However, according to Rawlins *et al.* (2010), the committee 'considered that […] there was a combination of factors' influencing its decision, including 'arguably, a corporate responsibility to provide treatment for an occupational hazard that at the time workers were exposed to asbestos was unrecognized'. Given that (co-author) Andrew Stevens was Chair of the committee responsible for the appraisal, this has been treated here as a reliable interpretation of events.

[11]Although NICE's methods allow for the exceptional consideration of wider societal impacts, the consideration of a technology's effects on economic productivity has been specifically prohibited since 2013 (NICE, 2013). The appraisal committee therefore appears to have contravened policy in this case.

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

consideration for appraisal committees, with de Folter *et al.*'s analysis indicating that equality is discussed in all appraisals and other studies highlighting specific cases in which consideration of socioeconomic or other forms of disadvantage have played a role in the decision to recommend a particular technology (Rawlins *et al.*, 2010; Yuasa *et al.*, 2021). Given NICE's ambition to promote health equality (NICE, 2008*b*, 2020*b*), further exploration of appraisal committees' understanding of this aim and its relationship with NICE's general approach appears warranted.

3.2.9. Stakeholder influence

Another of the six factors identified by Rawlins *et al.* (2010) as occasionally justifying a technology's recommendation beyond the usual threshold is 'stakeholder persuasion', with patients and their advocates playing 'an important role in shaping the views of NICE's advisory committees'. Quantitative evidence on the impact of stakeholder input is mixed (Dakin *et al.*, 2006; Cerri *et al.*, 2014; Dakin *et al.*, 2015; Yuasa *et al.*, 2021), but several qualitative studies have demonstrated stakeholders' ability to influence decision-making (Milewa and Barry, 2005; Milewa, 2006; Chang, 2020; Kieslich, 2020). According to Milewa and Barry (2005), this effect is mediated through four main strategies: (i) stakeholders' production of 'new' evidence; (ii) their accentuation of evidence that might not otherwise be considered relevant; (iii) alliance building across the stakeholder group and (iv) direct lobbying of NICE and the government. A case for which there is strong evidence of the latter is that of interferon beta, which, according to one committee member interviewed as part of Milewa and Barry's study, was 'one of the least cost-effective drugs there is' but was nevertheless recommended due to 'the pressure that was put on the government' by patient advocacy groups, 'back[ed] up' by treating clinicians (*ibid.*).[12] Stakeholder influence has also been shown to be influential in shaping NICE policy; specifically the end-of-life criteria (Chalkidou 2012; Chang, 2020).

3.2.10. Process factors

Finally, studies noted a range of process factors that appear to be correlated with appraisal outcome. Dakin *et al.* (2006) and Cerri *et al.* (2014) both identify an apparent association between the year of appraisal and its likelihood of recommendation, with later appraisals (up to 2009) more likely to result in either restriction or rejection than those conducted very early in NICE's lifetime. Dakin *et al.* (2006) also found that until 2004, pharmaceuticals were less likely to be rejected than other types of intervention (such as medical devices), although this finding was not replicated in a later study (Dakin *et al.*, 2015). Additionally, Cerri *et al.* (2014) found that an increase in the number of technologies considered within the same appraisal increased the odds of a restriction relative to a recommendation, indicating that committees may attempt to 'pick a winner' in such situations rather than fully recommending several similar technologies. More recent evidence on these types of process factors – and the drivers behind these trends – is lacking.

4. Discussion

This review identifies and provides insight into 11 factors that have been observed by multiple studies to play a substantive role in NICE decision-making. NICE's consideration of some of these factors – such as cost-effectiveness, uncertainty and clinical need – is well known and shaped in part by NICE's formal methods. However, the role played by other factors – including innovation, budget impact, rarity, age and cause of disease – is less well established and potentially in tension with the approach that NICE publicly articulates (NICE, 2013, 2020*b*). Across each of these factors, questions remain about the relationships that exist between them and their role in an evolving NICE policy landscape.

[12]This view is supported elsewhere in the literature (Crinson, 2004).

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

Unsurprisingly, evidence suggests that concern for cost-effectiveness is central to NICE decision-making. However, there are reasons to believe that the available literature may overestimate the importance of allocative efficiency to NICE's approach while failing to fully recognise the influence of other normative considerations. Several authors have pointed out the emergence in recent years of various 'decision rules' or 'modifiers' that codify exceptions to NICE's usual decision-making criteria, such that technologies are recommended that likely displace more QALYs from the NHS than they deliver (O'Mahony and Paulden, 2014; Paulden *et al*., 2014; Paulden, 2017; Charlton, 2020). Examples include the end-of-life criteria, the selective use of differential discount rates and the exceptional treatment of HSTs (Table 2). Much of the available literature, however, is based on appraisals conducted early in NICE's life, before these decision-rules were introduced: of the 29 articles included in the review, 15 (52%) are based entirely on appraisals completed before the introduction of the end-of-life criteria in 2009, with several others also drawing substantially on appraisals completed in the first decade of NICE's work (Table 1). This overrepresentation of early appraisals is made more pronounced by the increasing scale of NICE's core programme, which, as of 1 November 2021, has produced 731 appraisals, 563 of which (77%) have been completed since 2009 (NICE, 2020*c*). Where studies have included appraisals that span this period, the results are often not interpreted in the context of NICE's changing processes and methods, making it difficult to isolate the effects of such changes.

The review therefore identifies a considerable need to better understand the current (as opposed to the historical) grounds for NICE's decisions. Some outstanding research questions relating to individual factors have been highlighted in the previous section. However, a future programme of research might also focus on the following broad areas of investigation.

First, there is a need to unravel the roles played by different normative considerations within individual appraisals in order to better understand the ethical judgements that drive NICE decision-making. For example, in recommending a highly innovative but uncertain treatment for a rare and debilitating disease, an appraisal committee may cite each of these factors in explaining its decision. But while a committee driven primarily by concern for efficiency may justify its recommendation with reference to the future health benefits likely to be gained from supporting innovation and making research into rare diseases commercially viable, a committee driven by concern for equality might give greater weight to the importance of addressing current unmet clinical need. To date, little research has been conducted with the ability to identify such distinctions, explore their normative basis and evaluate the moral (and perhaps social and political) rationales that appraisal committees draw on in justifying their decisions. A similar knowledge gap also exists at the policy level, with further research needed to explore the normative basis and justification for, for instance, the prioritisation of cancer technologies and ultra-orphan drugs through the CDF and HST programme.

A second area of focus concerns the impact of recent policy changes on NICE decision-making. As previously highlighted, several major amendments to NICE's processes and methods have not yet been the subject of empirical research, making it difficult to establish their effects on the NHS and its users. If, as some have argued, such changes have increased NICE's tolerance for allocative inefficiency, then additional ethically oriented questions arise about the extent to which this evolving approach is morally coherent and consistent with NICE's stated principles (NICE, 2020*b*). Answering such questions will likely require further in-depth exploration of the normative factors embedded in NICE policy and the discretionary judgements made by appraisal committees across a range of recent cases.

A third area of potential focus relates to NICE's evolving approach to uncertainty and risk. In recent years, NICE has increasingly used 'managed access' as a way of mitigating risk to the NHS while providing patients with accelerated access to technologies whose benefits remain uncertain. This suggests that neither the 'framework approach' (in which cost-effectiveness analysis provides a general structure for considering and discussing key issues) nor the 'ordinal approach' (in which the dual hurdles of clinical- and cost-effectiveness are considered in turn) remain suitable ways of

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

conceptualising the evolving role of economic analysis in NICE appraisal (Williams *et al.*, 2007). Rather, it is plausible that such analysis might today be better characterised as contributing to a process of risk assessment, in which the potential impacts of a technology's adoption are identified primarily to determine how they might be managed and (where necessary) mitigated. Further research exploring how appraisal committees think about both clinical and financial risk, and how they balance these risks against cost-effectiveness and other normative factors, would be of significant value in understanding NICE's evolving approach. Also of value would be research that explores the quality of the data collected through managed access agreements and NICE's response when technologies made available through such arrangements are found to be either more or less effective than anticipated: an outcome with challenging ethical and political implications.

A significant obstacle to answering these and other questions is the increasing number of appraisals for which confidential commercial arrangements, often made in the context of managed access, mean that key factors in decision-making – including the technology's ICER – are not publicly reported. Another necessary consideration for those with an interest in NICE's work is the recent update to its processes and methods, implemented in February 2022 (NICE, 2020*a*). While the former change acts as a potential barrier to the conduct of robust empirical research, the latter highlights the need for such research to continue if the grounds for health care priority-setting in the UK are to remain transparent and well understood.

5. Study limitations

In considering what can be learnt from this study and how it can inform future research, it is necessary to acknowledge its limitations. The review was conducted by a single researcher and did not involve any formal quality assurance due to the difficulty in applying consistent evaluation criteria across widely differing methodologies. For this reason, it has been described as a narrative review, despite the systematic approach generally adopted. This systematic approach extended to the selection of search terms, which were deliberately broad and identified in part from previous work. However, the wide range of terms used to describe the subject of interest (i.e. factors considered during NICE technology appraisal) and their often non-specific nature ('other factors', 'criteria', 'judgements' and so on) make it possible that other relevant articles were missed. This risk was mitigated by conducting supplementary hand-searches of all included articles.

It should also be noted that the findings draw heavily on the results of several large retrospective analyses of NICE decisions, which are designed to demonstrate correlation rather than causation. As such, they are unable to conclusively prove that any given factor has influenced NICE decision-making.

6. Conclusions

In conclusion, this review demonstrates that though NICE decision-making has historically been strongly influenced by concern for cost-effectiveness, this is by no means the only consideration. Many other factors have also been observed to play a substantive role in decision-making, interacting with each other in ways that appear complex and are yet to be fully understood.

The review also highlights an over-representation in the literature of appraisals conducted early on in NICE's life, under methods that have since been superseded, offering a potentially misleading view of the importance of allocative efficiency to NICE's current approach. NICE's recent update to its processes and methods represents the next stage in the evolution of this approach and will likely further reduce the relevance of much of the existing literature. Given the consequence of NICE's advice on the way that resources are allocated within the NHS, and the organisation's status as a global authority on health care priority-setting, further research

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

that provides an empirical basis for scrutiny of NICE's approach, now and in the future, should be considered a priority.

Acknowledgements. My thanks go to Dr Courtney Davis and Dr Annette Rid, for their input to the design and preparation of this article and their support to the wider project of which it forms part. Thanks also to Professor John Abraham, who was kind enough to provide feedback on an earlier version of this article. This research was funded by the Wellcome Trust (Grant number 203351/Z/16/Z). For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Conflict of interest. The author declares none.

References

Alexander G, Werner R and Ubel P (2004) The costs of denying scarcity. *Archives of Internal Medicine* **164**, 593–596.

Bryan S, Williams I and McIver S (2007) Seeing the NICE side of cost-effectiveness analysis: a qualitative investigation of the use of CEA in NICE technology appraisals. *Health Economics* **16**, 179–193.

Calnan M, Hashem F and Brown P (2017) Still elegantly muddling through? NICE and uncertainty in decision making about the rationing of expensive medicines in England. *International Journal of Health Services* **47**, 571–594.

Catchpole P and Barrett V (2020) Keeping pace with pharmaceutical innovation: the importance of the NICE methods review. *PharmacoEconomics* **38**, 901–903.

Cerri K, Knapp M and Fernandez JL (2014) Decision making by NICE: examining the influences of evidence, process and context. *Health Economics, Policy and Law* **9**, 119–141.

Chalkidou K (2012) Evidence and values: paying for end-of-life drugs in the British NHS. *Health Economics, Policy and Law* **7**, 393–409.

Chang LC (2020) On cost effectiveness analysis and fairness: normalizing control of and resistance to NICE technology appraisals. *Financial Accountability and Management* **37**, 37–53.

Charlton V (2020) NICE and fair? Health technology assessment policy under the UK's National Institute for Health and Care Excellence, 1999–2018. *Health Care Analysis* **28**, 193–227.

Charlton V (2022) Does NICE apply the rule of rescue in its approach to highly specialised technologies?. *Journal of Medical Ethics* **48**, 118–125.

Charlton V and Rid A (2019) Innovation as a value in healthcare priority-setting: the UK experience. *Social Justice Research* **32**, 208–238.

Charlton V, Littlejohns P, Kieslich K, Mitchell P, Rumbold B, Weale A, Wilson J and Rid A (2017) Cost effective but unaffordable: an emerging challenge for health systems. *BMJ* **356**, j1402.

Clement F, Harris A, Li J, Yong K, Lee K and Manns B (2009) Using effectiveness and cost-effectiveness to make drug coverage decisions: a comparison of Britain, Australia, and Canada. *JAMA* **302**, 1437–1443.

Crinson I (2004) The politics of regulation within the 'modernized' NHS: the case of beta interferon and the 'cost-effective' treatment of multiple sclerosis. *Critical Social Policy* **24**, 30–49.

Dakin H, Devlin N and Odeyemi I (2006) 'Yes', 'No' or 'Yes, but'? Multinomial modelling of NICE decision-making. *Health Policy* **77**, 352–367.

Dakin H, Devlin N, Feng Y, Rice N, O'Neill P and Parkin D (2015) The influence of cost-effectiveness and other factors on NICE decisions. *Health Economics* **24**, 1256–1271.

de Folter J, Trusheim M, Jonsson P and Garner S (2018) Decision-components of NICE's technology appraisals assessment framework. *International Journal of Technology Assessment in Health Care* **34**, 163–171.

Department of Health (2015) *The NHS Constitution for England*. London: Department of Health.

Devlin N and Parkin D (2004) Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Economics* **13**, 437–452.

Dixon-Woods M, Agarwal S, Jones D, Young B and Sutton A (2005) Synthesising qualitative and quantitative evidence: a review of possible methods. *Journal of Health Services Research and Policy* **10**, 45–53.

Fleck L (2002) Rationing: don't give up. *Hastings Center Report* **32**, 35–36.

Government HM (2010) *Equality Act 2010*. London: The Stationary Office.

Grant M and Booth A (2009) A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal* **26**, 91–108.

Griffiths E, Hendrich J, Stoddart S and Walsh S (2015) Acceptance of health technology assessment submissions with incremental cost-effectiveness ratios above the cost-effectiveness threshold. *Clinicoeconomics and Outcomes Research* **7**, 463–476.

Kenny N and Joffres C (2008) An ethical analysis of international health priority-setting. *Health Care Analysis* **16**, 145–160.

Kieslich K (2020) Paradigms in operation: explaining pharmaceutical benefit assessment outcomes in England and Germany. *Health Economics, Policy and Law* **15**, 370–385.

https://doi.org/10.1017/S1744133122000032 Published online by Cambridge University Press

Kleinhout-Vliek T, de Bont A, Boysen M, Perleth M, van der Veen R, Zwaap J and Boer B (2020) Around the tables - Contextual factors in healthcare coverage decisions across Western Europe. *International Journal of Health Policy and Management* **9**(9), 390–402. doi: 10.15171/ijhpm.2019.145

Littlejohns P, Chalkidou K, Culyer A, Weale A, Rid A, Kieslich K, Coultas C, Max C, Manthorpe J, Rumbold B, Charlton V, Roberts H, Faden R, Wilson J, Krubiner C, Mitchell P, Wester G, Whitty J and Knight S (2019) National Institute for Health and Care Excellence, social values and healthcare priority setting. *Journal of the Royal Society of Medicine* **112**, 173–179.

Mason A and Drummond M (2009) Public funding of new cancer drugs: is NICE getting nastier? *European Journal of Cancer* **45**, 1188–1192.

Mauskopf J, Chirila C, Birt J, Boye K and Bowman L (2013) Drug reimbursement recommendations by the National Institute for Health and Clinical Excellence: have they impacted the National Health Service budget? *Health Policy* **110**, 49–59.

Milewa T (2006) Health technology adoption and the politics of governance in the UK. *Social Science and Medicine* **63**, 3102–3112.

Milewa T and Barry C (2005) Health policy and the politics of evidence. *Social Policy and Administration* **39**, 498–512.

NICE (1999) *Appraisal of New and Existing Technologies: Interim Guidance for Manufacturers and Sponsors*. London: NICE.

NICE (2004) *Guide to the Methods of Technology Appraisal*. London: NICE.

NICE (2005) *Social Value Judgements: Principles for the Development of NICE Guidance*. London: NICE.

NICE (2008*a*) *Guide to the Methods of Technology Appraisal*. London: NICE.

NICE (2008*b*) *Social Value Judgements: Principles for the Development of NICE Guidance*. London: NICE.

NICE (2009) *Appraising Life-Extending, End of Life Treatments*. London: NICE.

NICE (2013) *Guide to the Methods of Technology Appraisal*. London: NICE.

NICE (2016) *PMG19 Addendum A – Final Amendments to the NICE Technology Appraisal Processes and Methods Guides to Support the Proposed new Cancer Drugs Fund Arrangements*. London: NICE.

NICE (2017*a*) *Interim Process and Methods of the Highly Specialised Technologies Programme: Updated to Reflect 2017 Changes*. London: NICE.

NICE (2017*b*) *Procedure for Varying the Funding Requirement to Take Account of Net Budget Impact*. London: NICE.

NICE (2018) *Guide to the Processes of Technology Appraisal*. London: NICE.

NICE (2020*a*) Changes we're making to health technology evaluation (online). Available at https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance/changes-to-health-technology-evaluation (accessed 10 November 2021).

NICE (2020*b*) The principles that guide the development of NICE guidance and standards (online). Available at https://www.nice.org.uk/about/who-we-are/our-principles (accessed 10 November 2021).

NICE (2020*c*) Technology appraisal data: appraisal recommendations (online). Available at https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance/data/appraisal-recommendations (accessed 10 November 2021).

NICE (2022) *NICE Signals Commitment to Greater Flexibility in Its Evaluation of Promising New Health Technologies and Making Patient Access Fairer*. https://www.nice.org.uk/news/article/nice-signals-commitment-to-greater-flexibility-in-its-evaluation-of-promising-new-health-technologies-andmaking- patient-access-fairer (accessed 9 March 2022).

Nicod E and Kanavos P (2016) Scientific and social value judgments for orphan drugs in health technology assessment. *International Journal of Technology Assessment in Health Care* **32**, 218–232.

O'Connor A and Sargeant J (2015) Research synthesis in veterinary science: narrative reviews, systematic reviews and meta-analysis. *The Veterinary Journal* **206**, 261–267.

O'Mahony J and Paulden M (2014) NICE's selective application of differential discounting: ambiguous, inconsistent, and unjustified. *Value in Health* **17**, 493–496.

Paré G, Trudel MC, Jaana M and Kitsiou S (2015) Synthesizing information systems knowledge: a typology of literature reviews. *Information and Management* **52**, 183–199.

Paulden M (2017) Recent amendments to NICE's value-based assessment of health technologies: implicitly inequitable?. *Expert Review of Pharmacoeconomics and Outcomes Research* **17**, 239–242.

Paulden M, O'Mahony J, Culyer A and McCabe C (2014) Some inconsistencies in NICE's consideration of social values. *PharmacoEconomics* **32**, 1043–1053.

Raftery J (2001) NICE: faster access to modern treatments? Analysis of guidance on health technologies. *British Medical Journal* **323**, 1300–1303.

Raftery J (2006) Review of NICE's recommendations, 1999–2005. *British Medical Journal* **332**, 1266–1268.

Rawlins M and Culyer A (2004) National Institute for Clinical Excellence and its value judgments. *British Medical Journal* **329**, 224–227.

Rawlins M, Barnett D and Stevens A (2010) Pharmacoeconomics: NICE's approach to decision-making. *British Journal of Clinical Pharmacology* **70**, 346–349.

Schaefer R and Schlander M (2018) Is the National Institute for Health and Care Excellence (NICE) in England more 'innovation-friendly' than the federal joint committee (G-BA) in Germany? *Expert Review of Pharmacoeconomics & Outcomes Research* **17**, 1559732.
Scheunemann L and White D (2011) The ethics and reality of rationing in medicine. *Chest* **140**, 1625–1632.
Shah K, Cookson R, Culyer A and Littlejohns P (2013) NICE's social value judgements about equity in health and health care. *Health Economics, Policy and Law* **8**, 145–165.
Smith R (2004) The triumph of NICE. *British Medical Journal* **329**. doi: 10.1136/bmj.329.7459.0-g
Tappenden P, Brazier J, Ratcliffe J and Chilcott J (2007) A stated preference binary choice experiment to explore NICE decision making. *PharmacoEconomics* **25**, 685–693.
Timmons N, Rawlins M and Appleby J (2016) *A Terrible Beauty: A Short History of NICE.* Nonthaburi, Thailand: Health Intervention and Technology Assessment Program.
Ubel P (2000) *Pricing Life: Why It's Time for Health Care Rationing. Basic Bioethics Series*. Cambridge and London: MIT Press.
Williams I, Bryan S and McIver S (2007) How should cost-effectiveness analysis be used in health technology coverage decisions? Evidence from the National Institute for Health and Clinical Excellence approach. *Journal of Health Services Research and Policy* **12**, 73–79.
Wood E and Hughes D (2020) The New and Non-transparent Cancer Drugs Fund. *PharmacoEconomics* **38**, 1–4.
Yuasa A, Yonemoto N, Demiya S, Foellscher C and Ikeda S (2021) Investigation of factors considered by health technology assessment agencies in eight countries. *Pharmacoeconomics - Open* **5**, 57-69. doi: 10.1007/s41669-020-00235-6

Appendix 1: Search strategy

Search of academic literature

Search A: conducted on 19 October 2019

Title/abstract/key ('National Institute of Health and Care Excellence' OR 'National Institute of Health and Clinical Excellence' OR 'National Institute of Clinical Excellence' OR 'National Institute for Health and Care Excellence' OR 'National Institute for Health and Clinical Excellence' OR 'National Institute for Clinical Excellence')
AND
Title/abstract/key ('social value*' OR 'social norm*' OR 'societal value*' OR 'societal norm*' OR 'moral value*' OR 'core value*' OR 'other value*' OR 'other factor*' OR 'value judgement*' OR 'value judgment*' OR criterion OR criteria OR modifier* OR equity OR fair* OR justice OR 'trade off' OR trade-off OR tradeoff OR ethic* OR substantive OR normative)

Search B: conducted on 5 November 2019

Title (NICE)
AND
Title/abstract/key ('social value*' OR 'social norm*' OR 'societal value*' OR 'societal norm*' OR 'moral value*' OR 'core value*' OR 'other value*' OR 'other factor*' OR 'value judgement*' OR 'value judgment*' OR criterion OR criteria OR modifier* OR equity OR fair* OR justice OR 'trade off' OR trade-off OR tradeoff OR ethic* OR substantive OR normative)

Both searches were repeated on 16 November 2020, using the date range: '1 October 2019 to present'.

Search of grey literature

Multiple simple searches were conducted via Google Scholar on 19 October 2019 and were repeated on 5 November 2020. In each case, the first 10 pages of results based on relevance (i.e. the 100 results returned) were screened. Patents and citations were excluded and a date range of 1999–present was applied.
Search i: 'National Institute for Clinical Excellence', 'social value'
Search ii: 'National Institute for Care Excellence', 'social value'
Search iii: 'National Institute for Health and Care Excellence', 'social value'
Search iv: NICE, health, 'social value'
Search v: NICE, health, justice
Search vi: NICE, health, 'decision factor'
Search vii: NICE, health, 'other factor'

Appendix 2: Index of observed factors

Ref.	Factor	Studies in which factor was observed
1	Cost-effectiveness	Observed by every included study
2	Uncertainty	Rawlins and Culyer (2004); Devlin and Parkin (2004); Milewa and Barry (2005); Dakin *et al.* (2006); Raftery (2006; Tappenden *et al.* (2007), Clement *et al.* (2009); Mason and Drummond (2009); Chalkidou (2012); Cerri *et al.* (2014); Dakin *et al.* (2015); Calnan *et al.* (2017); de Folter *et al.* (2018); Charlton and Rid (2019); Schaefer and Schlander (2018); Kieslich (2020); Yuasa *et al.* (2021)
3	Budget impact	Devlin and Parkin (2004); Dakin *et al.* (2006); Mauskopf *et al.* (2013); Cerri *et al.* (2014)
4	Clinical need	Raftery (2001); Rawlins and Culyer (2004); Devlin and Parkin (2004); Dakin *et al.* (2006); Raftery (2006); Bryan *et al.* (2007); Tappenden *et al.* (2007); Mason and Drummond (2009); Rawlins *et al.* (2010); Chalkidou (2012); Mauskopf *et al.* (2013); Shah *et al.* (2013); Cerri *et al.* (2014); Dakin *et al.* (2015); Griffiths *et al.* (2015); Nicod and Kanavos (2016); de Folter *et al.* (2018); Charlton and Rid (2019); Schaefer and Schlander (2018); Chang (2020); Kleinhout-Vliek *et al.* (2020); Wood and Hughes (2020); Yuasa *et al.* (2021)
5	Innovation	Milewa and Barry (2005); Rawlins *et al.* (2010); Shah *et al.* (2013); Dakin *et al.* (2015); Griffiths *et al.* (2015); Kieslich (2020); Charlton and Rid (2019); Nicod and Kanavos (2016); de Folter *et al.* (2018); Chang (2020); Yuasa *et al.* (2021)
6	Rarity	Devlin and Parkin (2004); Chalkidou (2012); Shah *et al.* (2013); Dakin *et al.* (2015); Griffiths *et al.* (2015); Chang (2020); Wood and Hughes (2020); Yuasa *et al.* (2021); Nicod and Kanavos (2016); de Folter *et al.* (2018); Cerri *et al.* (2014)
7	Age	Tappenden *et al.* (2007); Bryan *et al.* (2007); Rawlins *et al.* (2010); Dakin *et al.* (2015); Calnan *et al.* (2017); de Folter *et al.* (2018)
8	Cause of disease	Rawlins and Culyer (2004); Rawlins *et al.* (2010); Chalkidou (2012); Shah *et al.* (2013)
9	Wider societal impacts	Rawlins *et al.* (2010); Nicod and Kanavos (2016); de Folter *et al.* (2018); Charlton and Rid (2019); Yuasa *et al.* (2021)
10	Stakeholder influence	Milewa and Barry (2005); Milewa (2006); Dakin *et al.* (2006); Rawlins *et al.* (2010); Chalkidou (2012); Cerri *et al.* (2014); Dakin *et al.* (2015); Kieslich (2020); Chang (2020); Yuasa *et al.* (2021)
11	Process factors	Dakin *et al.* (2006); Cerri *et al.* (2014); Dakin *et al.* (2015); Mason and Drummond (2009)

Cite this article: Charlton V (2022). The normative grounds for NICE decision-making: a narrative cross-disciplinary review of empirical studies. *Health Economics, Policy and Law* **17**, 444–470. https://doi.org/10.1017/S1744133122000032