Delta Hat

Powered by Petauri™

NICE decision-making ICER threshold analysis

NICE decision-making ICER threshold analysis

Contents

Table of figures

Table of tables

NICE decision-making ICER threshold analysis

Authors

Kurt Taylor
Petauri Evidence (formerly Delta Hat)

Rachael Batteson
Petauri Evidence (formerly Delta Hat)

Simone Critchlow
Petauri Evidence (formerly Delta Hat)

Throughout the research, the authors were employed by Delta Hat, which has since been rebranded as Petauri Evidence.

Funding

This report was commissioned and funded by the Association of the British Pharmaceutical Industry (ABPI).

May 2025

Background

The National Institute for Health and Care Excellence (NICE) is responsible for evaluating the clinical and cost effectiveness of healthcare interventions within the National Health Service (NHS) in England. To guide resource allocation, NICE health technology assessments (HTAs) typically apply cost-effectiveness thresholds set at £20,000–£30,000 for each additional year of health (measured in quality-adjusted life years [QALYs]) gained by a patient.[1] Central to this evaluation is the incremental cost-effectiveness ratio (ICER), which represents the additional cost required to gain one additional QALY when comparing a new technology with the current standard of care. The acceptability of technologies with ICERs exceeding these thresholds depends on various factors, including the degree of uncertainty surrounding the estimate and any additional, non-health benefits that may not be fully captured in QALY-based analysis.

While the £20,000–£30,000 ICER threshold provides a general benchmark, decisions within or above this range remain subject to NICE committee discretion. However, the rationale behind decisions within the threshold range is often unclear, particularly regarding the relative influence given to different decision modifiers and how they impact the acceptable ICER for a technology to be deemed cost effective.

The aim of this work was to analyse published NICE final guidance documentation to determine whether there has been a trend in changes to NICE's decision-making ICER thresholds over time. By examining historical trends, this report seeks to provide insights into the consistency, transparency, and evolution of NICE's approach to cost-effectiveness evaluations.

Methods

1.1 Searches

To conduct this analysis, relevant technology appraisals (TAs) guidance published between January 2010 and December 2024 was identified using the NICE website (https://www.nice.org.uk/guidance/published). The NICE database allows for direct extraction of key information based on the 'last updated date', including the TA title, reference number, published date, and last updated date. Searches covered January 2010 to December 2024.

Only TAs that had 'published guidance' within the specified date range were included in the analysis. The analysis excluded multiple technology appraisals (MTAs), appraisals marked as 'terminated' or 'withdrawn', and those identified as duplicates or superseded guidance. This approach ensured that only finalised and relevant NICE TAs were considered for further evaluation.

1.2 Data extraction

The decision-making ICER threshold was extracted for TAs where stated and grouped into 13 distinct decision-making ICER threshold categories (presented in Table 1). Several assumptions were made for areas of subjectivity. Multiple scenario analyses were conducted to test the impact of these assumptions and assess the robustness of the findings.

Decision-making ICER thresholds were categorised into three distinct time periods based on key changes to NICE methods and processes:

1. **2010 - 2015**
   start of the data search up to the change in the Cancer Drugs Fund (CDF) framework
2. **2016 - 2021**
   In 2016, the pre-existing CDF framework was adapted to provide a vehicle for interim funding of promising new cancer medicines/indications, via managed access arrangements, while further evidence was collected to address clinical uncertainty
3. **2022 - 2024**
   This is representative of the introduction of the updated methods manual (following NICE's extensive Methods Review) up until the end of 2024

This classification allowed for an examination of how decision-making ICER thresholds have evolved over time in response to these changes.

1.3 Analysis

For each decision-making ICER threshold category, a specific value was assigned so that each category had a defined number that could be used to calculate an average. For example, the category 'lower end of £20,000–£30,000' was assigned a threshold value of £22,500. In some instances, to leverage the data available, assumptions related to subjective interpretations had to be made (e.g. use of a £50,000 threshold associated with meeting the end-of-life [EOL] criteria). To fully understand the impact of these assumptions, values were explored in three distinct scenario analyses (presented in Table 1). To assess changes in the distribution of decision-making ICER threshold categories across the key time periods, a chi-squared test was conducted. Charts were generated to visualise the distributions of decision-making ICER categories over time. In addition to this, an average decision-making ICER threshold was calculated for each of the three distinct time periods.

Results

A total of 735 TAs were identified from the targeted review within the specified date range. Of these, 17 were published before 2010 and excluded. Among the remaining 718, 52 were MTAs, and 118 were classified as 'terminated' appraisals, both of which were excluded from full data extraction. In total 548 individual TAs were considered relevant for review.

To avoid potential duplication, appraisals that had been reconsidered (e.g. were initially recommended for use within the CDF) and have since received superseded NICE guidance, the updated and most recent TA was used to inform the analysis. Of the 548 TAs considered, 33 included multiple decision criteria (e.g. different decision-making thresholds within a single TA for different subpopulations) and were analysed separately, resulting in 581 individual ICER threshold decisions. Among the 581 extractions, 52 were cost comparisons, leaving 529 decisions associated with cost effectiveness to inform the final analysis.

Most TA outcomes were either recommended (50%, n=272) or optimised (37%, n=204). Optimised TAs are those that were accepted but with a restricted indication, meaning they were approved for use under specific conditions or for a limited population. A total of 50 TAs (9%) were not recommended, while the remaining TAs were recommended through other mechanisms: 18 (3%) via the CDF, 3 (1%) for use only in research, and 1 (less than 1%) through another managed access route.

Table 1 presents a summary of the findings from the review with a disaggregated breakdown of how the TAs final guidance referenced a decision-making ICER threshold and what corresponding values were assumed to inform analyses. Out of the 529 individual ICER decisions, 38% (n=203) referred to the standard cost-effectiveness range (i.e. £20,000–£30,000), 19% (n=99) did not state a threshold, and 2% (n=8) were stated as having dominant results (with the treatment being more effective and less costly than the comparator). As these categories did not provide any insight into the explicit decision-making threshold used to inform decision-making, these results were removed from the base case analysis leaving a data set of 219 (41%) individual thresholds. Different scenarios were conducted to explore the impact of different assumptions related to the ICER threshold:

- **Scenario 1:**
  In 104 instances (20% of the total sample), the EOL criteria were considered to be met. For these TAs, the assumed £50,000 ICER threshold was down-weighted to the standard £30,000 threshold to align with the application of severity modifiers under the updated methods framework – effectively removing the additional weighting previously granted for EOL. This adjustment was made to ensure consistency across methodologies. Of these 104 TAs, 98 (19% of the total sample) met the EOL criteria but did not explicitly state that a £50,000 threshold was applied. These were classified as having an 'implied £50,000' threshold. In the base case analysis, a threshold of £27,000 was applied to these TAs, with sensitivity analyses exploring alternative assumptions of £30,000 or full exclusion to assess the impact
- **Scenario 2:**
  In 203 instances (38% of the total sample), reference was made to the standard £20,000–£30,000 threshold. Scenario 2 builds on Scenario 1 (applying the same EOL assumptions) but also incorporates the 203 instances of the £20,000–£30,000 range, with the assumption that the midpoint (£25,000) was the corresponding ICER decision-making threshold
- **Scenario 3:**
  In 98 instances (19% of total sample), the EOL was met and therefore the decision-making ICER threshold were assumed to be £50,000. In this scenario, these are removed from the analysis. In addition, as per scenario 1 and 2, the 10 confirmed EOL £50,000 ICER instances were cateorgised as £27,000

Results are presented for the base case analysis, and all three scenarios.Table 1 provides a summary of assumptions.

Table 1: Decision-making ICER categories and assumptions

#	Stated DMI category	N	%	Assumed DMI Base case	Assumed DMI Scenario 1	Assumed DMI Scenario 2	Assumed DMI Scenario 3
1	<£20,000	7	1%	£17,500	Same as base case	Same as base case	Same as base case
2	£20,000	29	5%	£20,000	Same as base case	Same as base case	Same as base case
3	Lower end of the £20,000 - £30,000	24	5%	£22,500	Same as base case	Same as base case	Same as base case
4	£25,000	8	2%	£25,000	Same as base case	Same as base case	Same as base case
5	<£30,000	3	1%	£27,500	Same as base case	Same as base case	Same as base case
6	Upper end of the £20,000 - £30,000	14	3%	£27,500	Same as base case	Same as base case	Same as base case
7	£30,000	20	4%	£30,000	Same as base case	Same as base case	Same as base case
8	£20,000 - £30,000	203	38%	Excluded	Same as base case	£25,000	Same as base case
9	<£50,000	10	2%	£24,000	£27,000	£27,000	£27,000
10	Implied £50,000†‡	98	19%	£27,000	£30,000	£30,000	Excluded
11	£50,000‡	6	1%	£30,000	Same as base case	Same as base case	Same as base case
12	Not stated/unclear	99	19%	Excluded	Same as base case	Same as base case	Same as base case
13	Dominant	8	2%	Excluded	Same as base case	Same as base case	Same as base case
Total data available		**529**	**100%**	**219 (41%)**	**219 (41%)**	**422 (80%)**	**121 (23%)**

Abbreviations: DMI, decision-making incremental cost-effectiveness ratio; ICER, incremental cost-effectiveness ratio; NICE, National Institute for Health and Care Excellence; QALY, quality-adjusted life year; TA, technology appraisal. † Implied £50,000 was often cited where TAs indicated that end-of-life criteria was met, but no specific threshold was specified. ‡ Under the old NICE methodology, a QALY weighting was applied under certain circumstances to account for end-of-life medicines. Given this has now been replaced with the NICE severity modifier, the assumed DMIs have been down-weighted accordingly to align with a traditional £20,000-£30,000 threshold.

Figure 1 illustrates the proportion of decision-making ICER threshold decisions in the base case across each key time period, highlighting a significant concentration of decisions at £27,000 in the 2010–2015 and 2016–2021 periods. In contrast, the distribution of decisions in the 2022–2024 period is more evenly spread across different decision-making ICER threshold values. The chi-squared test revealed a significant association between time periods and threshold categories ($p<0.001$), indicating that the distribution of decision-making thresholds differed across time periods.

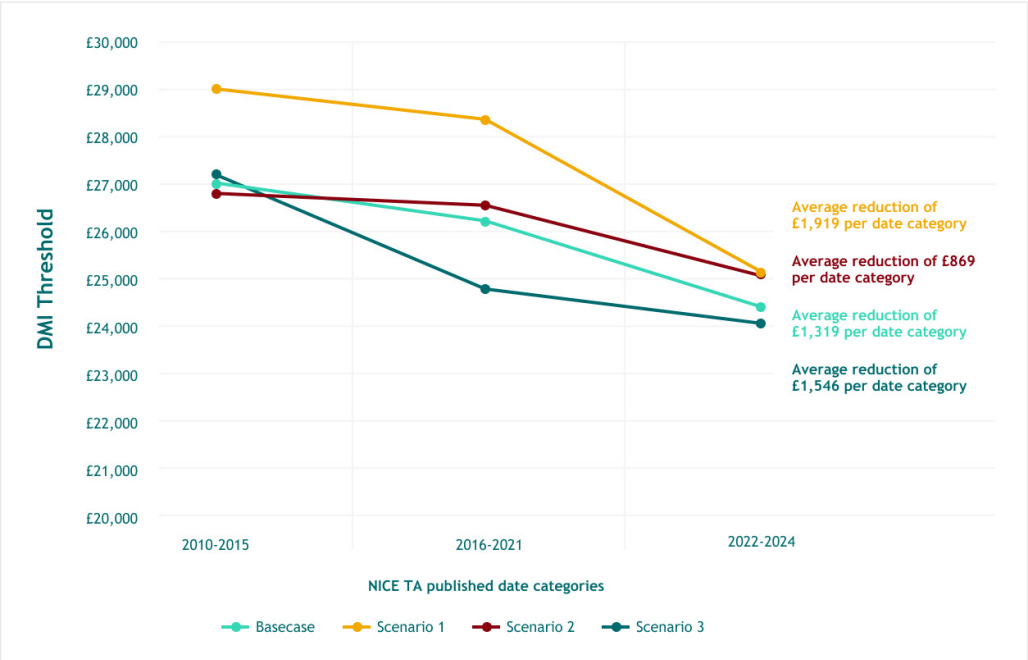
Results from the scenario analyses were generally consistent (see Appendix A), indicating a shift in the distribution of decision-making ICER values over time. Statistically significant associations were observed in Scenarios 1 and 2 based on chi-squared tests, whereas Scenario 3 did not show a significant association.

Figure 1:
Bar chart illustrating the proportion of decision-making ICERs between time periods

Abbreviations: DMI, decision-making incremental cost-effectiveness ratio; NICE, National Institute for Health and Care Excellence; TA, technology appraisal.

Figure 2 shows the average decision-making ICER threshold in each TA date category. The average threshold was estimated based on a weighted average of the percentage of appraisals that fell within each decision-making ICER threshold category. As shown in Figure 2 between the appraisals reported between 2010–2015 and the appraisals reported between 2016–2021, there was a decrease in the decision-making ICER threshold. This trend was maintained for TAs reported in 2022–2024. The trends were also consistent when examining oncology and non-oncology trends and across all three scenarios explored (see Appendix B).

Figure 2:
Line graph illustrating the weighted average decision-making ICERs across the three time periods



Abbreviations: DMI, decision-making incremental cost-effectiveness ratio; ICER, incremental cost-effectiveness ratio; NICE, National Institute for Health and Care Excellence; TA, technology appraisal.

Discussion and conclusions

This analysis reviewed published NICE final guidance documentation to assess whether there has been a shift in the overall ICER decision-making threshold over time. A total of 529 appraisal decisions from 2010 to 2024 were analysed.

The analyses indicated a clear and significant shift in the distribution of the NICE ICER decision-making threshold across the pre-defined time periods. In recent years, appraisal committees have increasingly applied decision-making thresholds towards the lower end of NICE's £20,000–£30,000 range. In the base case analysis, the weighted decision-making ICER values were £27,065 for 2010–2015, £26,210 for 2016–2021, and £24,427 for 2022–2024, reflecting a consistent downward trend. For the base case analysis, on average, the NICE decision-making ICER threshold has decreased by approximately £1,319 in each time period.

A key limitation of this analysis is the presence of missing data. A decision-making ICER threshold could be estimated for only 41% of appraisals in the base case, meaning that assumptions had to be made (i.e. for the chi-squared test, it was assumed that data were missing at random, which is untestable), which may impact the reliability of the findings. Despite this, the observed trends were consistent in sensitivity analyses. Notably, the proportion of appraisals with an identifiable decision-making ICER threshold have increased in recent years, likely due to improved clarity in NICE's reporting practices.

These findings suggest an evolving approach in NICE's decision-making, with greater caution in the application of higher decision-making thresholds. Future research could explore the implications of these shifts on patient access to treatments and healthcare resource allocation.

Appendix A: Scenario analysis

Figure 3: Bar chart illustrating the proportion of decision-making incremental cost-effectiveness ratios between time periods for each scenario

Scenario 1

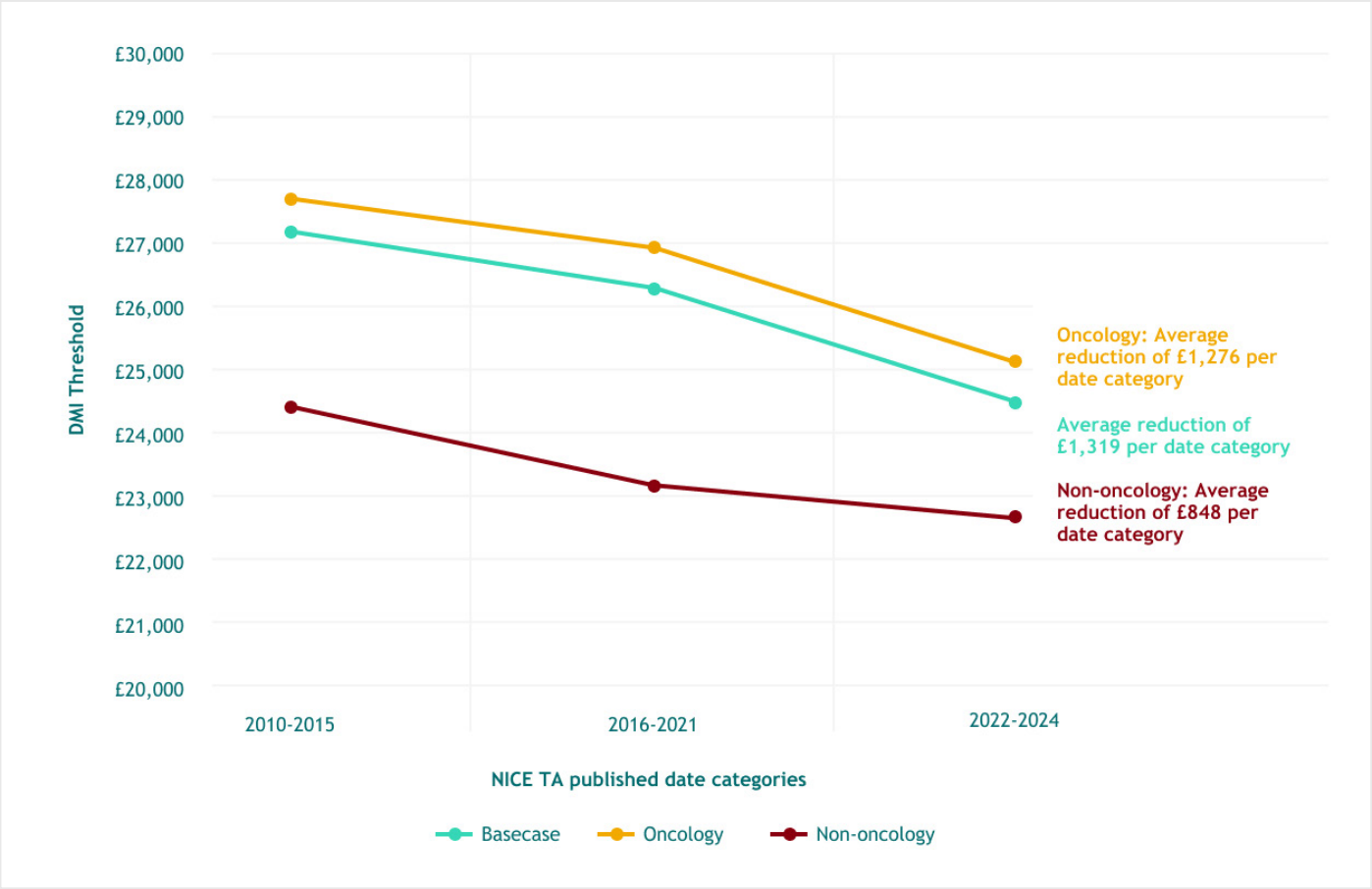
Scenario 2

Scenario 3

Abbreviations: DMI, decision-making incremental cost-effectiveness ratio; EOL, end-of-life; NICE, National Institute for Health and Care Excellence; TA, technology appraisal.

Appendix B: Oncology analysis

Figure 4: Line graph illustrating the weighted average decision-making ICER across the three time periods for the base case assumptions (oncology and non-oncology)



Abbreviations: DMI, decision-making incremental cost-effectiveness ratio; ICER, incremental cost-effectiveness ratio; NICE, National Institute for Health and Care Excellence; TA, technology appraisal.

References

1. National Institute for Health and Care Excellence (NICE). NICE health technology evaluations: the manual.