International Journal of Technology Assessment in Health Care, 34:2 (2018), 163–171.
© Cambridge University Press 2018
doi:10.1017/S0266462318000090

DECISION-COMPONENTS OF NICE'S TECHNOLOGY APPRAISALS ASSESSMENT FRAMEWORK

Joost de Folter*
National Institute for Health and Care Excellence (NICE)
joost.defolter@nice.org.uk; folterj@gmail.com

Mark Trusheim*
Massachusetts Institute of Technology (MIT)

Pall Jonsson
National Institute for Health and Care Excellence (NICE)
Pall.Jonsson@nice.org.uk

Sarah Garner
National Institute for Health and Care Excellence (NICE)

Objectives: Value assessment frameworks have gained prominence recently in the context of U.S. healthcare. Such frameworks set out a series of factors that are considered in funding decisions. The UK's National Institute of Health and Care Excellence (NICE) is an established health technology assessment (HTA) agency. We present a novel application of text analysis that characterizes NICE's Technology Appraisals in the context of the newer assessment frameworks and present the results in a visual way.
Methods: A total of 243 documents of NICE's medicines guidance from 2007 to 2016 were analyzed. Text analysis was used to identify a hierarchical set of decision factors considered in the assessments. The frequency of decision factors stated in the documents was determined and their association with terms related to uncertainty. The results were incorporated into visual representations of hierarchical factors.
Results: We identified 125 decision factors, and hierarchically grouped these into eight domains: Clinical Effectiveness, Cost Effectiveness, Condition, Current Practice, Clinical Need, New Treatment, Studies, and Other Factors. Textual analysis showed all domains appeared consistently in the guidance documents. Many factors were commonly associated with terms relating to uncertainty. A series of visual representations was created.
Conclusions: This study reveals the complexity and consistency of NICE's decision-making processes and demonstrates that cost effectiveness is not the only decision-criteria. The study highlights the importance of processes and methodology that can take both quantitative and qualitative information into account. Visualizations can help effectively communicate this complex information during the decision-making process and subsequently to stakeholders.

Keywords: Decision making, Evidence-based medicine, Social values, Technology assessment

Value assessment frameworks are used to provide a comprehensive understanding of the benefits that a new medicine brings to patients and health systems, and to ensure those benefits are reflected in the price. This involves accurately capturing all the benefits and harms, and ensuring that, therefore, the "value" of new products is not under or overestimated. Several such frameworks have been put forward in the United States recently and reviewed in detail (1–5). To support this development, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) created the Initiative on U.S. Value Assessment Frameworks to inform the shift toward a value-driven healthcare system by promoting the development and dissemination of high-quality unbiased value assessment frameworks (1).

The frameworks are controversial when linked to funding decisions about the "value" of high-cost pharmaceuticals and have, therefore, been subject to analysis and comparison (2;3). The analyses note potential areas of concerns and outline guiding principles in relation to the implementation of these frameworks.

Oortwijn and colleagues (4) reviewed developments in value frameworks and their potential. They concluded that these value frameworks primarily intend to inform patient–clinician conversations, or specifically to inform value-based pricing. They defined the characteristics of good-quality value frameworks, including the need to define core components and to adhere to the principles of transparency, predictability, and broad stakeholder involvement.

In Europe, frameworks have emerged in the form of best-practice guidance. The European network for Health Technology Assessment (EUnetHTA), for instance, has developed practical tools and methods for health technology assessments (6;7). The motivation for EUnetHTA's work was primarily to harmonize a diverse approach to technology assessments in Europe, which could lead to variation in patient access to new healthcare interventions.

The United Kingdom became one of the earliest adopters of health technology assessment with the establishment of its HTA agency. The National Institute for Health and Care

*Joost de Folter and Mark Trusheim were joint primary authors.
This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

https://doi.org/10.1017/S0266462318000090 Published online by Cambridge University Press

Excellence (NICE) provides the National Health Service with a range of advice on effective, good value healthcare. As an HTA body, its approach to decision making and guidance development is firmly underpinned by the principles of evidence-based medicine. However, its deliberative approach to assessments inevitably requires a consideration of a wide range of aspects relating to value, both to patients and the health service.

This research puts NICE's Technology Appraisals in the context of the recently described value assessment frameworks and uses textual analysis to systematically explore what factors influence decisions made by NICE's Technology Appraisal committees about the value of a pharmaceutical product. Here we attempt to: reveal the complexity of the decision making, identify how many potential factors are considered, explore how NICE's methods inform decisions, and assess how much a formal framework is able to capture the assessment process while allowing for case-to-case considerations. To characterize the analysis, this work includes the development of novel visual representations to illustrate the results.

NICE TECHNOLOGY APPRAISALS

The scope of this research is focused on the area of NICE's technology appraisals. NICE's independent advisory committees are the key decision makers in the technology appraisal process. The process is implemented using an iterative, inclusive process that requires participation from the public, patients, academic and clinical experts, as well as the manufacturer and NICE officials through multiple stages. Briefly, the typical steps in NICE's single technology assessment are (i) scoping, which sets out the objectives and parameters of the assessment; (ii) a submission of evidence by the manufacturer, relating to cost and clinical effectiveness; (iii) a review of the evidence, performed by an independent evidence review group; (iv) consideration of the evidence and subsequent preliminary recommendation by the independent advisory committee; (v) a public consultation; (vi) a review of comments received through consultation and a formulation of a final appraisal determination by the independent advisory committee; (vii) the publication of NICE's guidance document. This process is described in detail in the Guide to the Processes of Technology Appraisal (8).

Analysis of NICE's Value Assessment Framework

Since 2008, the NICE Guide to the Methods of Technology Appraisal (9;10) and its document on Social Value Judgement (11) form a cornerstone of NICE's assessment of new and established medical technologies. Collectively, these documents describe the principles of appraisal methods, and outline the broad principles that NICE should follow in developing NICE guidance, including those beyond clinical and cost effectiveness. While these guides describe the normative NICE processes, they do not set out the whole range of potential factors that can inform the decision making by the independent advisory committee assessment.

In their 2009 paper, Rawlins and colleagues described a concise list of factors commonly used in NICE technology appraisals decisions (12). The list comprised scientific value judgements regarding clinical effectiveness and cost effectiveness as well as social value judgements regarding factors such as the severity of underlying illness, end of life treatments, stakeholder persuasion, significant innovation, disadvantaged populations, and children. While much work exists regarding HTAs and how they should be conducted, few analyses have been published on characterizing HTAs in the context of value assessment frameworks and how they have been applied to multiple products and therapeutic areas. Studies that have been performed focus on a product subset such as companion diagnostics (13), on coverage statistics in a therapeutic area such as cancer (14), or on use of a particular method such as comparative effectiveness research (15). A recent study was performed on social value judgement factors across a number of European agencies including NICE, although this was limited to a sample of ten orphan drugs (16).

METHODS

To understand what comprises the value assessment framework defined by NICE's appraisal methods for medicines, we used multiple techniques to identify the key decision factors involved in the decision making by NICE's advisory committees. The context of analysis was the independent advisory "committee discussion" section of the final appraisal determination (FAD) documents. For convenience this is what the term "document" refers to in this section.

A set of generalized decision factors was determined based on the "committee discussion" section of a selection of documents and expert revision. Five documents were selected with an emphasis placed on relevance and representativeness (see Supplementary Table 1 for details). The former was met by selecting relatively recently published appraisals. The latter was met by selecting appraisals with a combination of the following characteristics: an original appraisal, a review of a previous appraisal, a particular complex case (multiple consultations, and resulting a change in NICE's methodology), and a negative recommendation appraisal.

One reviewer (J.d.F.) reviewed the content of the committee discussion section of each of the documents. Thematic analysis was performed and the key decision factors were identified. Members of the NICE technical team performed additional review, and identified and arbitrated discrepancies. The final decision factors can be organized in a grouped structure or a relational one as some factors overlap. For simplicity and giving a clearer overview, a hierarchical set of relationships was chosen to represent the findings, grouping the decision factors into domains. The groups were initially based on the structure of

https://doi.org/10.1017/S0266462318000090 Published online by Cambridge University Press

key headings in technology appraisal documents (Clinical Effectiveness and Cost Effectiveness), and logical relationships and dependencies between factors, determined and refined in collaboration with technical team members. Forms of visual representations were explored to help communicate the set of decision factors.

In this analysis, we applied automated text analysis to the 243 NICE technology appraisals of medicines publicly released from January 2007 through December 2016 (see Supplementary Table 2 for the full list). The set of decision factors was matched against the documents to determine how often each factor was part of an appraisal. The matched documents for each decision factor also included the matched documents from their sub-factors. The text matching was implemented using Python and the Natural Language Toolkit (NLTK), based on common methods for text analysis (17). Using this tool, document texts were first extracted, common stop-words were removed, and the remaining text underwent text stemming. Each decision factor term was also stemmed, and additionally word roots and synonyms were derived taking word context into account. The decision factor terms were matched against the documents on sentence level. This resulted in information on which decision factors were indicated for which documents. Next, for each decision factor, the number of relating documents was counted. These absolute document counts were converted to more meaningful normalized values. The normalized document frequency is defined by the number of documents a decision factor appears in, divided by the total number of documents. A corresponding visualization was created to show the results of the analysis visually.

Next, the same matching method was used to determine indication of uncertainty in relation to the committees' decision making. This uncertainty relates to the committee's confidence in reaching a specific decision based on the evidence provided in a particular technology appraisal. This analysis was performed by searching the document for words indicating uncertainty, where these appeared in relation to the decision factors. Next, the normalized value of this uncertainty matching was calculated, relative to the previously obtained frequency for the decision terms. A similar visualization for this was generated.

The general context for this analysis was the committee discussion section, which comprises the core rationale of how the decision was reached. Focusing on this section enabled us to analyze the most impactful factors considered by the committees. To illustrate the decision factors in greater detail, the decision factors of two specific technology appraisals were also visualized, by mapping the relevant decision factors and excluding all remaining factors. The appraisals chosen for this purpose were a positive recommendation (TA235 *mifamurtide for the treatment of osteosarcoma*), and a negative one (TA338 *pomalidomide for relapsed and refractory multiple myeloma previously treated with lenalidomide and bortezomib*).

RESULTS

Decision Factors

The analysis of decision factors resulted in a set of 125 detailed decision factors that were grouped into eight domains (Condition, Current Practice, Clinical Need, New Treatment, Studies, Clinical Effectiveness, Cost Effectiveness, and Other Factors) with up to two intermediate grouping levels (see Figure 1 and Supplementary Table 3). Some of the decision factors are repeated among the branches highlighting a level of interdependency.

The factors regarding Clinical Effectiveness and Cost Effectiveness domains together make up over 50 percent of the decision factors, emphasizing their importance. We also compare with Rawlins' set of decision factors (12), which separates judgements into two categories: Scientific value judgements and Social value judgements. The results of the current study remain consistent with Rawlins' Scientific value judgements. The Other Factors domain incorporates Rawlins' Social value judgements, and additionally includes factors from NICE's current social value judgements.

Decision Factors Matching in Technology Appraisals Documents

The complete set of decision factors was matched against the 243 NICE appraisal documents from 2007–16, to determine how often each factor was part of an appraisal. A visualization was developed to illustrate the results of this analysis by encoding this additional information as color intensity (Figure 2), where darker colors indicate higher frequency. This result shows that the individual decision factors were consistently used in the different appraisals, measured by the relative document frequency.

The Clinical Effectiveness and Cost Effectiveness domains appear in all NICE technology appraisals. In all, seven of the eight decision factor domains, not just Clinical and Cost effectiveness, were consistently used in NICE recommendations (above 99 percent of the technology appraisals). The eighth, Clinical Need, was cited in fewer, but still a majority of cases (approximately 91 percent of the technology appraisals). This lower frequency is due to the textual analysis method used that relies on synonym matching rather than being able to recognizing meaning independently of the specific words that were used. A manual inspection of the appraisals where Clinical Need had not been detected by the automated text mining, revealed that it had indeed also been considered for each of those cases.

As opposed to the domains, in the lower level sub-factors, there is greater variability of specific decision factors being cited. As one might expect, treatment effect, incremental cost effectiveness (ICER) and comparative treatment occur in nearly all technology appraisals. Also as one might expect the second level special conditions captured in Other Factors did not consistently appear across assessments—an

https://doi.org/10.1017/S0266462318000090 Published online by Cambridge University Press

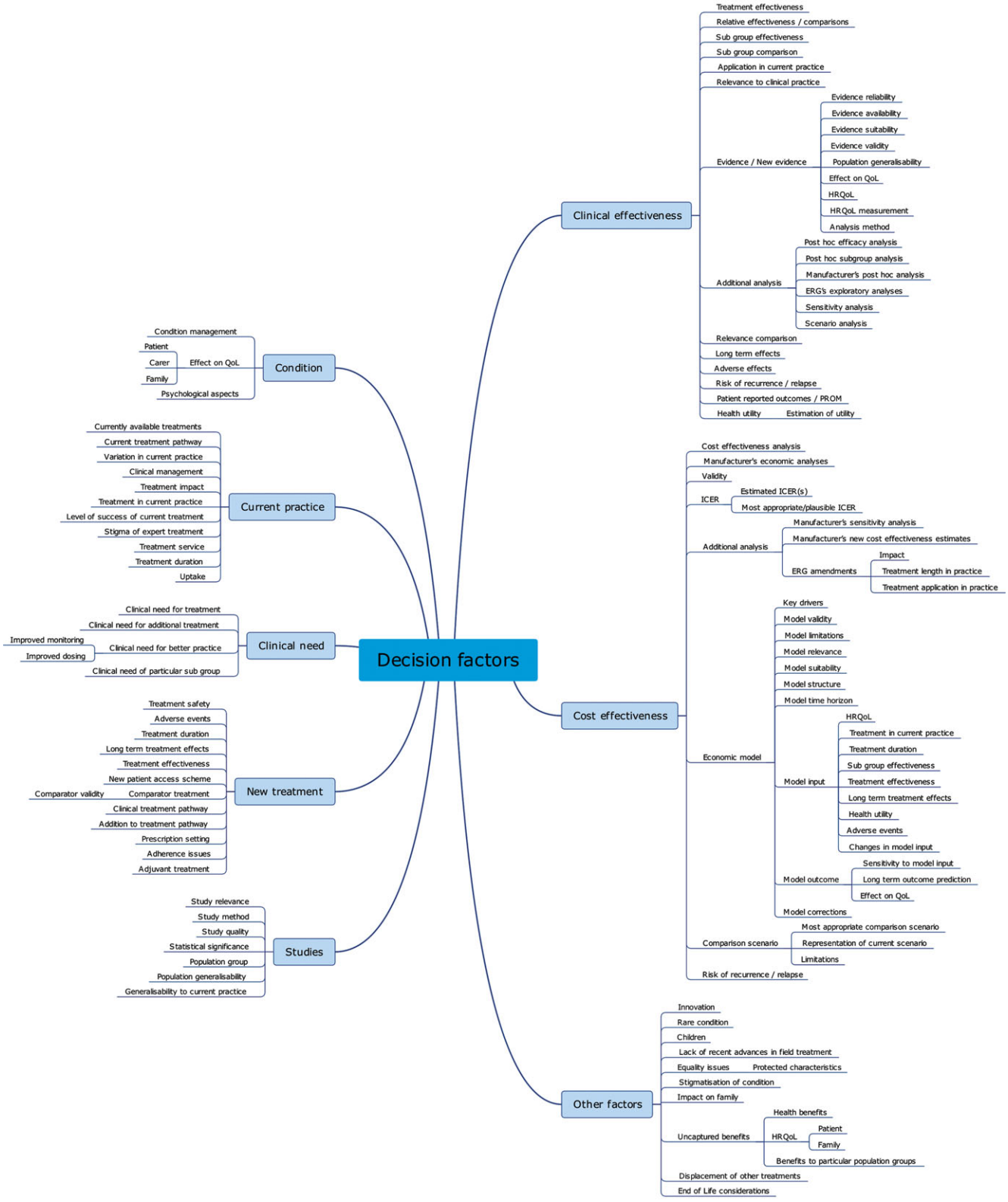


Figure 1. Hierarchical visualization of the 125 NICE technology appraisal decision factors identified by means of textual analysis. Top level domains are shown in boxes, with sub factors branching toward the outer edges of the tree.

exception being Equality Issues. There are factors with very low frequencies. As with Clinical Need in the top grouping, this may be due to the limitations of the textual analysis algorithm.

Decision Factors Uncertainty Matching

The visualization shown in Figure 3 illustrates how many times a decision factor appeared in relation to uncertainty, relative to the total number of times a decision factor appeared. The

https://doi.org/10.1017/S0266462318000090 Published online by Cambridge University Press

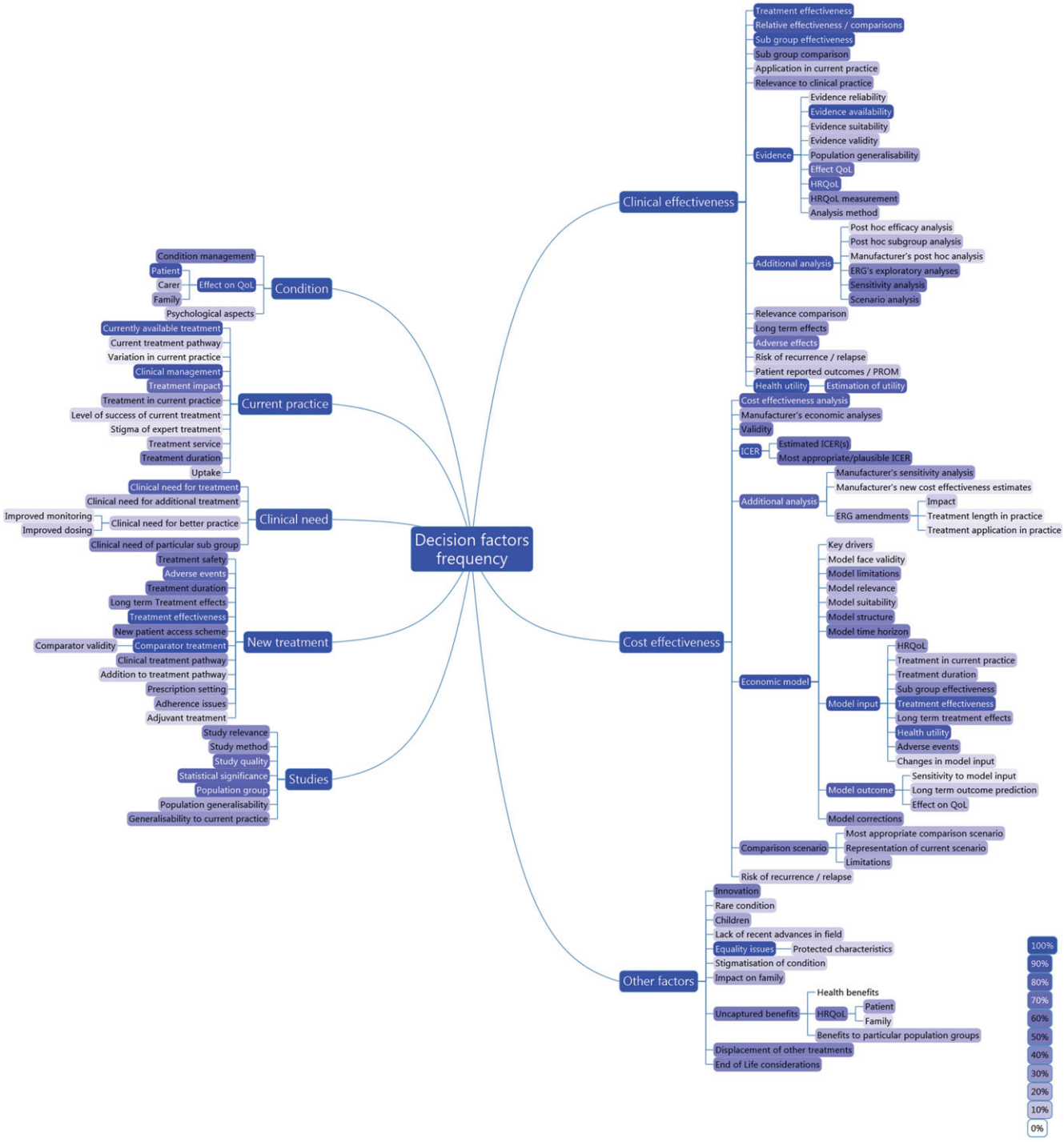


Figure 2. Decision factor frequency of use: Visualization of the occurrence of the decision factors in technology appraisal documents. The color intensity represents the normalized decision factor frequency.

results of this automated analysis highlight consistent areas of committees' concerns regarding uncertainty about the decision factors. Uncertainty was most often associated with the top level decision factor groups of Clinical Effectiveness, Cost Effectiveness, and New Treatment. These uncertainty areas corresponded to some of the most frequently cited decision factor groups. At the decision factor sub-tiers, uncertainty was most frequent regarding clinical effectiveness evidence, health utility, ICER, economic modeling, and comparators. This analysis highlights uncertainty is commonly associated with many of the decision factors.

Analysis of Individual Technology Appraisals

To illustrate the nuances of individual assessments, the decision factors of two Technology appraisals were considered.

https://doi.org/10.1017/S0266462318000090 Published online by Cambridge University Press

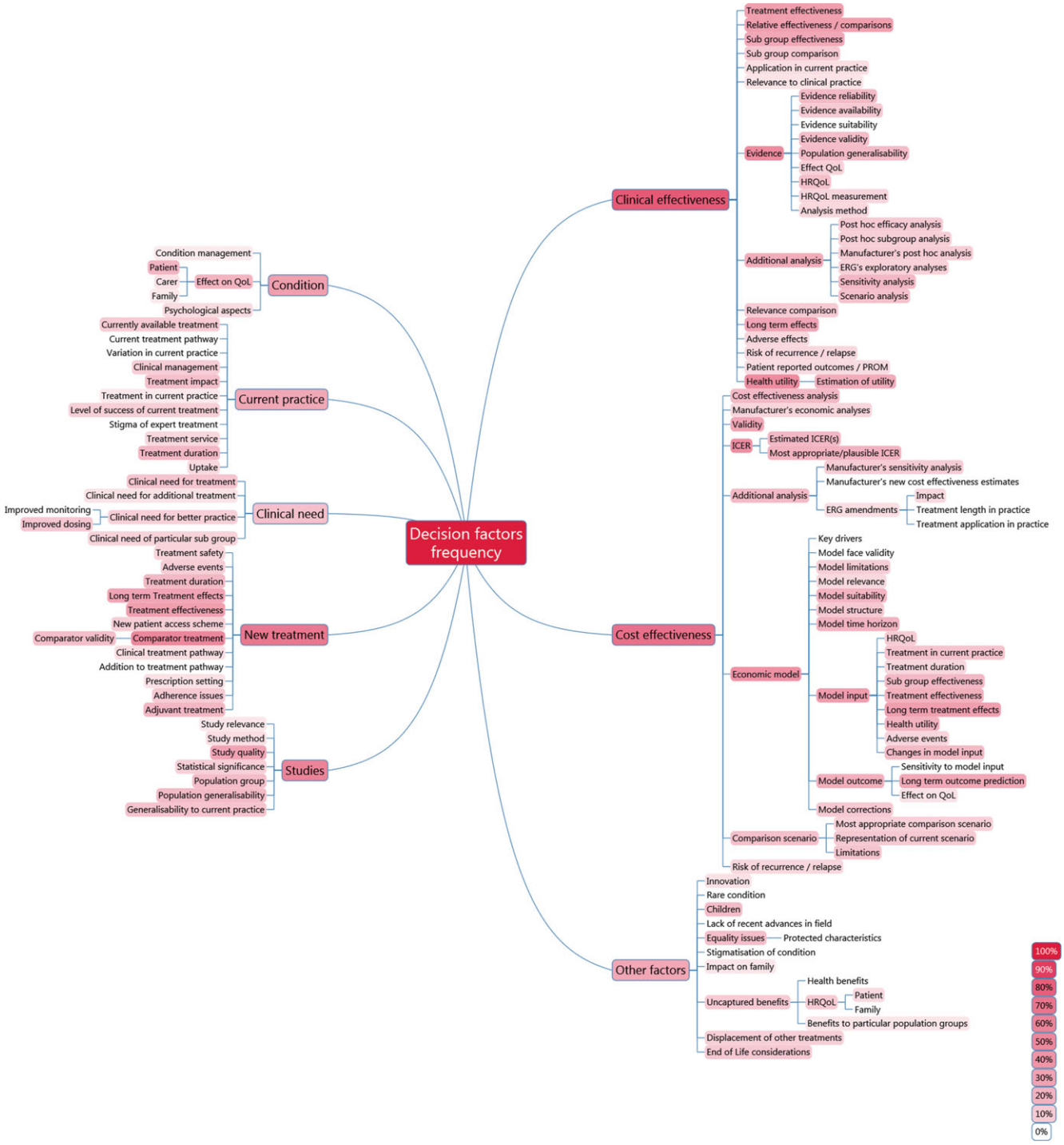


Figure 3. Uncertainty and decision factors: Visualization of occurrence of decision factors in the context of uncertainty. The color intensity represents the normalized frequency of uncertainty/term matches.

These are the positive recommendation TA235, *mifamurtide for the treatment of osteosarcoma;* and the negative recommendation TA338, *pomalidomide for relapsed and refractory multiple myeloma previously treated with lenalidomide and bortezomib.* In TA235, the decision factors reported in the technology appraisal committee discussion section contain only a sub-set of the full set of decision factors. The reported factors are visualized in Figure 4a. Importantly, TA235 features many Other Factors, which were described as social factors by Rawlins and colleagues (12), including innovation, rare condition, children, and lack of recent treatment advances.

The decision factors reported for the negative recommendation in technology appraisal TA338 spanned seven of the eight decision factor groups, as illustrated in Figure 4b. As with TA235, the decision factors were a sub-set of the second level factors. However, unlike TA235 this negative TA338

https://doi.org/10.1017/S0266462318000090 Published online by Cambridge University Press

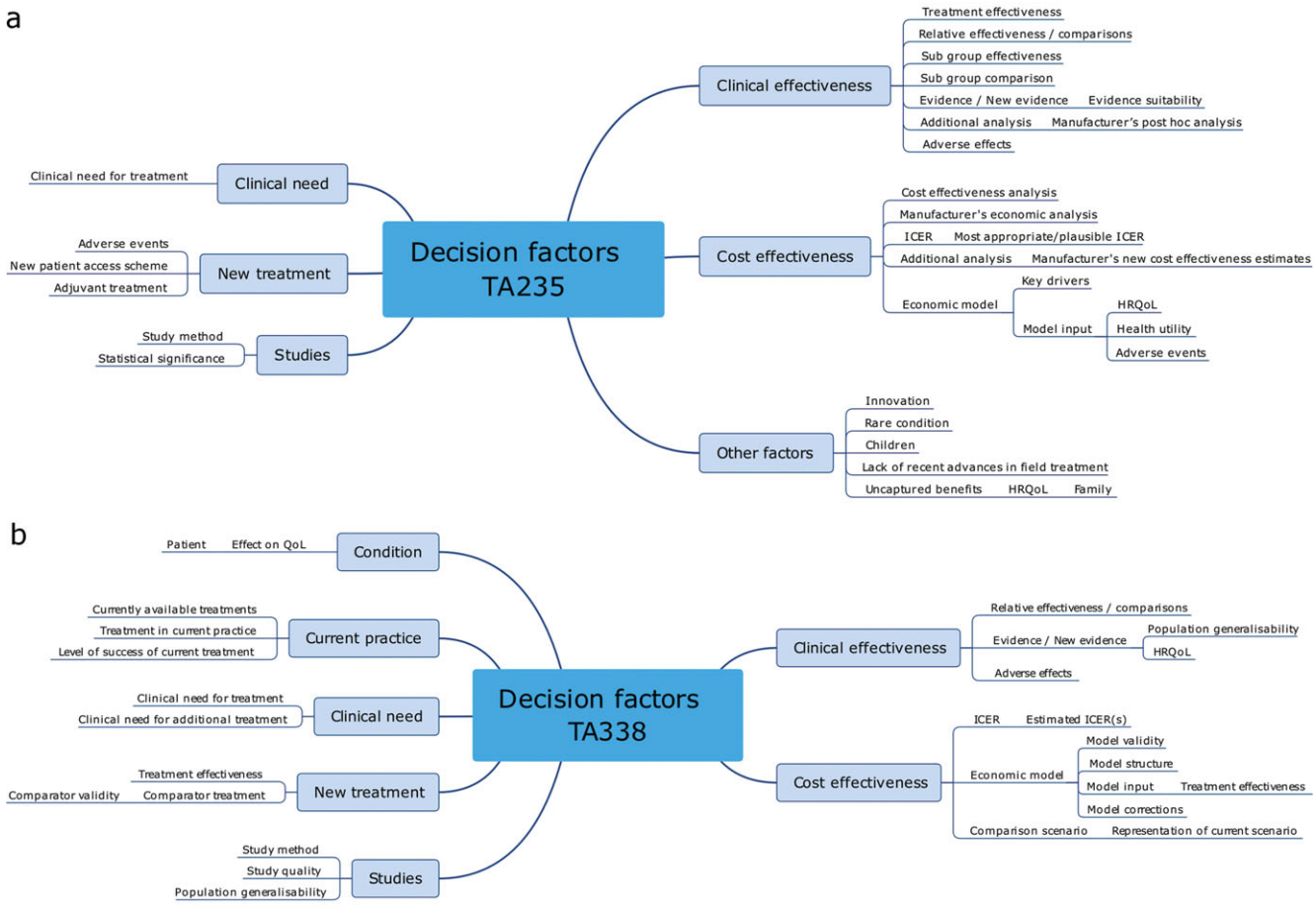


Figure 4. Decision factors in example technology appraisals. Visualization of decision factors most cited in two specific technology appraisals a: TA235: Mifamurtide for the treatment of osteosarcoma, b: TA338: Pomalidomide for relapsed and refractory multiple myeloma previously treated with lenalidomide and bortezomib. Decision factor hierarchy trimmed to show only decision factors cited.

assessment has more of the core second level factors of treatment effectiveness, ICER and economic models, evidence quality, new patient access schemes, adverse events, and comparators.

DISCUSSION

NICE's Value Assessment Framework

The value assessment framework described here includes decision-making domains that comprise factors that are either considered qualitatively or quantitatively. Results show seven of the eight decision factor domains, not just clinical and cost effectiveness, were consistently used in NICE recommendations. The eighth (Clinical Need) was found using the textual analysis algorithm in a majority of cases. The fact of this not being universally found is due to the limitations of the automated literal text matching. Textual analysis revealed that the large, deeply tiered super set of 125 decision factors were reported in practice across the NICE technology assessment collection. It is important to note that this analysis does not provide insight on the relative weight applied to the decision factors.

This work includes development of visual representations to provide suitable overviews of various results. As the decision factors follow a hierarchical structure, a corresponding visual that shows this hierarchy appears very suitable. This then allowed additional information to be incorporated, such as frequency of decision factors and uncertainty relating to those factors in a simple yet insightful way. Finally, the same structure was used to summarize the most relevant decision factors of a single appraisal. The decision factors provide a source for development of various types of visuals, which NICE is currently exploring further for implementation in practice.

Although many NICE decision factors are quantitative, the majority of the factors are more qualitative in nature requiring deliberation. Considering this and the many areas of uncertainty, this study also reveals the complexity of the decision making. Over this complete collection of agency's medicines recommendations in a rapidly evolving HTA field, the analysis indicates significant consistency of decision factors, and by implication decision criteria and agency priorities.

The appearance of the Social Value Judgements factors cited among the NICE technology appraisals (such as

https://doi.org/10.1017/S0266462318000090 Published online by Cambridge University Press

decision factors regarding the young, elderly, and orphan drug populations) suggests that the appraisal committees and relevant stakeholders also seek a flexible, compassionate approach that considers the context of the disease and the social values surrounding the medical technology, for example, to encourage innovation for conditions that have few alternatives.

This work observes NICE's value assessment framework demonstrates a multi-stakeholder, case by case process for appraisals of UK technology. This research also demonstrates the complexity of NICEs decision-making process, which is not just based on cost effectiveness.

Future Possibilities

It should be noted that the success of this analysis was enabled by the structured document formats that NICE makes publicly available. This, and continuation of the current research, will enable moving to more sophisticated metrics of decision factor influence.

This analysis looks at a 10- year period in its entirety. Thus, while the technology appraisal field has evolved over the time, this analysis was not structured to detect any trends that may exist within that period. Analyses by time period might elucidate such trends. Similarly, sub-set analyses by therapeutic area, regulatory path, or product type might illustrate adaptations to special circumstances.

For the purpose of comparison, this study includes the decision factors related to two individual appraisals as illustration. Further development of these methods could be performed to include in-depth application, comparison, and validation of appraisals. Extending this type of analysis to additional HTA collections might provide an evidence base to understand international HTA practice variations, inform HTA framework construction and perhaps encourage their harmonization while acknowledging local contextual differences. Additional research could gain insight into relative weight or impact for each decision factor, and we hope this work will inform future discussions on potential value frameworks.

This analysis further suggests that the value-based decision making cannot be reduced to a mathematical combination of clinically and financially measured parameters. Data uncertainty infuses many, even most, of the decision factors most of the time. Flexibly applying the set of decision factors to fit the context of the disease, therapeutic approach, and social values also confounds a prospective, mathematically prescriptive appraisal process. Technology appraisal processes, like the medical approaches they examine and the societies they serve, must be informed by evidence but will continue to require processes that combine such data with judgments and values in the context of each technology.

Health technology appraisal is a rigorous yet evolving and judgement filled field in which transparency of process and its framework is desired but often difficult to communicate. Modern analysis and visualization techniques could help HTA agencies improve the consistency, efficiency, utility and acceptance of their findings. Achievements which, in turn, can accelerate the adoption of the best treatments that provide the highest patient outcomes a compassionate society can afford, while simultaneously providing incentives for improved and breakthrough medicines.

SUPPLEMENTARY MATERIAL

Supplementary Table 1:
https://doi.org/10.1017/S0266462318000090
Supplementary Table 2:
https://doi.org/10.1017/S0266462318000090
Supplementary Table 3:
https://doi.org/10.1017/S0266462318000090

CONFLICTS OF INTEREST

Three of the authors were employed by the National Institute of Health and Care Excellence at the time this study was performed. The research and its conclusions are the authors' own; they were not influenced or approved by the National Institute of Health and Care Excellence.

REFERENCES

1. Doshi JA, Willke RJ. Advancing high-quality value assessments of health care interventions. *Value Health*. 2017;20:181-184.
2. Westrich K. *Current landscape: Value assessment frameworks*. Washington DC: National Pharmaceutical Council; 2016.
3. Sorenson C, Lavezzari G, Daniel G, et al. Advancing value assessment in the United States: A multistakeholder perspective. *Value Health*. 2017;20:299-307.
4. Oortwijn W, Sampietro-Colom L, Habens F. Developments in value frameworks to inform the allocation of healthcare resources. *Int J Technol Assess Health Care*. 2017;33:1-7.
5. Angelis A, Kanavos P. Multiple Criteria Decision Analysis (MCDA) for evaluating new medicines in health technology assessment and beyond: The advance value framework. *Soc Sci Med*. 2017;188:137-156.
6. Kristensen FB, Lampe K, Chase DL, et al. Practical tools and methods for health technology assessment in Europe: Structures, methodologies, and tools developed by the European Network for Health Technology Assessment, EUnetHTA. *Int J Technol Assess Health Care*. 2009;25(Suppl 2):1-8.
7. EUnetHTA. HTA Core Model Version 3.0. 2016 [updated January 25, 2016]. http://www.htacoremodel.info (accessed February 26, 2018).
8. NICE. *Guide to the processes of technology appraisal*. London: National Institute for Health and Care Excellence; 2014.
9. NICE. *Guide to the methods of technology appraisal*. London: National Institute for Health and Care Excellence; 2008.
10. NICE. *Guide to the methods of technology appraisal*. London: National Institute for Health and Care Excellence; 2013.
11. NICE. *Social value judgements: Principles for the development of NICE guidance*, Second edition. London: National Institute for Health and Care Excellence; 2008.

https://doi.org/10.1017/S0266462318000090 Published online by Cambridge University Press

12. Rawlins M, Barnett D, Stevens A. NICE's approach to decision-making. *Pharmacoeconomics*. 2009;70:346-349.
13. Byron SK, Crabb N, George E, Marlow M, Newland A. The health technology assessment of companion diagnostics: Experience of NICE. *Clin Cancer Res*. 2014;20:1469-1476.
14. Mason A, Drummond M, Ramsey S, Campbell J, Raisch D. Comparison of anticancer drug coverage decisions in the United States and United Kingdom: Does the evidence support the rhetoric? *J Clin Oncol*. 2010;28:3234-3238.
15. Sorenson C, Drummond M, Chalkidou K. Comparative effectiveness research: The experience of the National Institute for Health and Clinical Excellence. *J Clin Oncol*. 2012;30:4267-4274.
16. Nicod E, Kanavos P. Scientific and Social value judgments for orphan drugs in health technology assessment. *Int J Technol Assess*. 2016;32:218-232.
17. Manning C, Raghavan P, Schütze H. *An introduction to information retrieval*. Cambridge, England: Cambridge University Press; 2009.

https://doi.org/10.1017/S0266462318000090 Published online by Cambridge University Press