

# Quality of the End of Life — Utility Values in Advanced Solid Tumours in Technology Appraisals in the UK

*Noemi Muszbek, MSc, Senior Research Scientist, Modeling and Simulation;  
Agnes Benedict, MSc, MA, Senior Research Scientist, Modeling and Simulation;  
Linda Hortobagyi, MSc, Research Associate, Modeling and Simulation*

## INTRODUCTION

Set up in 1999 in the United Kingdom (UK), the National Institute for Health and Care Excellence (NICE) provides guidance to the National Health Service (NHS), local authorities, healthcare charities, and healthcare professionals on the most effective ways to prevent, diagnose, and treat disease and ill health, while offering the best value for money and reducing inequalities and variation.<sup>1</sup> These guidances incorporate the technology appraisals that assess the clinical and cost-effectiveness of health technologies according to a rigorous methodological guideline.<sup>2</sup> NICE guidance is mandatory for healthcare providers within the UK and often serves as an example for health technology assessment (HTA) agencies worldwide.

In the development of these guidances, manufacturer submissions are requested to demonstrate the clinical and cost-effectiveness of the healthcare technology, and, together with other stakeholder submissions, are reviewed by the Evidence Review Group (ERG) assigned to that appraisal; in the case of multiple technology appraisals (MTA), a new cost-effectiveness model is developed by the ERG. The decision in

recommending the technology is reached by the Review Committee based on the submitted evidence and its assessment according to the process guidelines.<sup>3,4</sup>

In recent years, the evaluation of cancer treatments has represented a large proportion (approximately 30%) of technology appraisals. This is due to the decision of the Department of Health at the end of 2007 to refer “all new cancer drugs and significant new licensed indications” to NICE, preferably in parallel with licensing, if “there is a sufficient patient population and evidence base on which to carry out an appraisal and that there is not a more appropriate alternative mechanism for appraisal.”<sup>5</sup> At the same time, the assessment of cancer treatments and the decision rules applicable for the assessment of end-of-life treatments have been in the centre of debates resulting in the establishment of the “end-of-life criteria” and in the setup of the Cancer Drugs Fund. The establishment of the end-of-life criteria in 2009 allows the differential treatment of technologies aiming to extend life in patients with short life expectancy, and those that are licensed for small numbers of patients with incurable illnesses,

by placing a higher value at the end of life, and, as a result, allowing for the use of a higher cost-effectiveness threshold.<sup>6</sup> The Cancer Drugs Fund was set up in 2011 for the reimbursement of cancer drugs that were not recommended by NICE.

These developments highlight the importance of discussions and research in the key aspects of technology appraisals of cancer treatments, such as the evaluation of quality of life through established, preference-based generic instruments resulting in utility values. In addition, issues regarding the methods and the face validity of the utility values have received more emphasis in the recent appraisals, for example, that some of the utility values measured in clinical trials in oncology patients are too similar to that of the general public.

Previous research highlighted the potential issues with the use of utility values in the cost-effectiveness studies in oncology and the application of the NICE reference case in the technology appraisals, as described in the NICE methodological guideline.<sup>7-10</sup> A review by Tosh *et al.*<sup>10</sup> examined the utility values in the NICE technology assessments up to 2008, comparing the methodology to the 2004 NICE

reference case. This review identified multiple issues and a large scope for improvement. A more recent review of NICE technology appraisals from the perspective of mapping found poor reporting of mapping methods and a falling proportion of appraisals using mapping.<sup>11</sup>

The aim of this study was to assess the use and elicitation of utility data in the current NICE technology appraisals of advanced oncology treatments in light of the current guidelines. Extrapolation methods will be assessed in an upcoming issue of *The Evidence Forum*.

## GUIDELINES FOR UTILITIES AND EXTRAPOLATION

Published in 2013, the current NICE methodological guideline is similar to the previous one in terms of the reference case for utility data (see *Table 1*). The preferred method for measurement of quality of life is still the EuroQoL-5D (EQ-5D) questionnaire, with the new version offering five levels in each dimension also mentioned in the new guidelines, although these guidelines do not detail the situations when EQ-5D is not available.

## METHODS

### Literature Review

A review was conducted to identify all completed NICE technology appraisals for the treatment of advanced solid tumours issued between 2011 and August 2013. Solid tumours were defined as tumours not containing cysts or liquid area.<sup>13</sup>

### Data Extraction

Final appraisal documentations, ERG reports, and, where available, manufacturer submissions were reviewed, and the following data were extracted and reviewed.

NICE REFERENCE CASE FOR UTILITIES <sup>2,12</sup>		
	2008 NICE reference case	2013 NICE reference case
<i>Preferred method of measurement</i>	EQ-5D When EQ-5D data are not available: <ul style="list-style-type: none"> <li>• Mapping to EQ-5D values</li> <li>• Direct valuation of descriptions of health states using time trade-off</li> </ul>	EQ-5D or EQ-5D-5L When EQ-5D data are not available: <ul style="list-style-type: none"> <li>• Mapping to EQ-5D values</li> </ul> If EQ-5D is proven to be inappropriate (e.g., lacks responsiveness & does poorly on tests), then alternative health-related quality of life (HRQoL) measures may be used if its validity is proven. Differences between EQ-5D and the alternative must be shown.
<i>Preferred method of measurement for children</i>	Standardised and validated preference-based measures of HRQoL for use in children, e.g., Health Utility Index 2 (HUI 2)	Standardised and validated preference-based measures of HRQoL for use in children
<i>Measured by</i>	Patients (and carers when impact of treatment on carers' health is important) When not possible, data need to be obtained from carer in preference to health care professionals.	Patients When not possible, data need to be obtained from carer in preference to health care professionals.
<i>Adjustments to values</i>	Not mentioned	Can be made if required (e.g., age, comorbidities)
<i>Preferred source of values</i>	No preference stated If from literature, it needs to be systematic and transparent.	Clinical trial If not available, systematic, transparent literature review.
<i>Valuation of preferences</i>	General public A representative sample of the UK population	General public A representative sample of the UK population
<i>Method of preference elicitation</i>	Choice-based method	Choice-based method

table 1

- Disease area
- Comparators
- ERG
- Issue time of the guidance
- Methods of utility elicitation
- Source of data
- Utility values used in the base case and in sensitivity analyses (both those submitted by manufacturers and the final versions accepted by the Review Committee)
- Criticisms and overall conclusions of the ERG and the Committee
- Final decision of the committee regarding the drugs appraised

Following the most common disease pathway for solid tumours, utility values were extracted for pre-progression (or stable disease) and post-progression health states and were organised into pre- and post-progression pairs. The methods of utility estimation have been compared against the NICE reference case from the 2008 and 2013 NICE guidance<sup>2,12</sup> (see Table 1). Data extracted by one reviewer was checked by an additional reviewer.

## RESULTS

Twenty-one technology appraisals were identified, 2 of which were terminated and 19 were extracted. There were 17 single technology appraisals and 2 MTAs. The indications where utility data were available were breast cancer, renal cell carcinoma, ovarian cancer, lung cancer, metastatic colorectal cancer, prostate cancer, melanoma, and urothelial tract carcinoma.

All technology appraisals included cost-utility models looking at patients passing through distinct health states (Markov models). These health states included, among others:

- Stable or pre-progression health state defined mostly by the progression-free survival (PFS) curve
- Post-progression or progressed health state
- Death defined by overall survival (OS)

## UTILITIES

In the 19 technology appraisals, there were 32 sets of complete, base case, pre- and post-progression

utilities reported. Two assessments (TA255, TA259) either did not report or reported only partial utility data. One MTA did not have the manufacturer submission for one of the comparisons.

The mean utility was 0.747 (standard deviation [SD]: 0.06, range: 0.65–0.87) and 0.55 (SD: 0.11, range: 0.25–0.79) for pre- and post-progression, respectively. Among the pre-progression utilities, the majority of utility values were used for patients ages 55–64 (83%); however, the utility values were equivalent to the values of the general UK population ages 75 and older (see Table 2). The results were similar for the utility values post-progression, with a slightly higher age.

Only 28% of utility values followed the preferred method of eliciting quality of life in the reference case presented in both the 2008 and 2013 guidelines, and were collected with the help of the EQ-5D questionnaire (see Table 3). More than half of the utilities were elicited using direct valuation. Among these, standard gamble was the most common

DISTRIBUTION OF UTILITY VALUES—According to age groups in the appraisal and age group equivalent in the general population				
	Distribution of utility values according to the median/mean* age in the trials used as primary source of efficacy data		Proportion of utility values matching the age specific utilities in the general population**	
Age	Pre-progression	Post-progression	Pre-progression	Post-progression
75+	0.00%	0.00%	82.54%	96.92%
65-74	0.00%	3.57%	1.59%	3.08%
55-64	85.19%	82.14%	11.11%	0.00%
45-54	14.81%	14.29%	4.76%	0.00%
Below 44	0.00%	0.00%	0.00%	0.00%

\*If reported, the median age was used in the analyses. In the absence of median age, the mean age was extracted from the trials, which was used as a proxy for the median age of the patient cohort in the model.  
 \*\*Source of values for the different age groups was Kind et al. 1999.<sup>14</sup>

table 2

METHOD OF UTILITY ELICITATION				
	Proportion of pre-progression utilities		Proportion of post-progression utilities	
	<i>All utilities</i>	<i>Utilities accepted by Committee</i>	<i>All utilities</i>	<i>Utilities accepted by Committee</i>
<i>Indirect valuation</i>	41.38%	43.33%	37.93%	40.00%
EQ-5D	29.31%	26.67%	27.59%	26.67%
HUI	10.34%	13.33%	10.34%	13.33%
Other	1.72%	3.33%	0.00%	0.00%
<i>Direct valuation</i>	58.62%	56.67%	62.07%	60.00%
SG	50.00%	53.33%	58.62%	56.67%
TTO	8.62%	3.33%	3.45%	3.33%
Other	0.00%	0.00%	0.00%	0.00%
<i>Mapping</i>	17.24%	10.00%	5.17%	6.67%
<i>Abbreviations: EQ-5D=EuroQol-5 Dimensions; HUI=Health Utilities Index; SG=standard gamble; TTO=Time Trade-off</i>				

table 3

method (54.31%). The majority of these direct preferences were elicited from the general population as per NICE reference case (see Figure 1). The methods used for the elicitation of values were similar for utility values submitted by the manufacturers and the ones used by the ERG and also accepted by the Committee. Mapping, however, was more common in the manufacturer submissions than among the values recommended by the ERG and the Committee (see Table 3).

The overwhelming majority of utility values came from the literature (60.3% and 80.7% pre- and post-progression, respectively), including previous technology appraisals. More pre-progression utilities were available from trials compared to post-progression values (37.9% and 19.3% pre- and post-progression, respectively).

Although pre- and post-progression utilities overlapped, there seems to be a trend for post-progression values being lower. Utility values from the

manufacturer submissions and recommended by the ERGs or Committees are similar, with the latter slightly lower (see Figure 2). Values were usually held highly uncertain, and in 75% of cases a univariate sensitivity analysis was conducted.

### MOST COMMON CRITICISM FROM THE ERG OR THE COMMITTEE

Utility values were criticized by both ERGs and the Review Committee in almost all cases. The most common criticisms were regarding the methodology and the face validity. Criticism regarding the methodology included the following:

- In many cases, the method of data collection was described in insufficient detail, leading to increased uncertainty.
- Utilities were not collected in the clinical trials.

- Utility data collected in the trials was not representative of the patients throughout the progression of the disease.
- Utilities in the model were not derived according to the NICE reference case.
- Disutilities for adverse events were not incorporated into the model.
- Literature review of utility values was not systematic.

Criticism regarding face validity centred on 1) doubts if the values were representative of the patient population evaluated (e.g., in terms of age, country, health status; and/or 2) the values not reflecting the impression and experience of the disease or the course of the disease in 42% of the technology appraisals. Values were compared to that of the general population and were expected to be significantly lower. Differences between pre- and post-progression utilities were also assessed and criticized if minor.

The expectations also varied according to the particular ERG. For example, in breast cancer, the commonly used source of data from Lloyd *et al.*<sup>15</sup> was accepted or required by some of the ERGs, yet was criticized by another ERG for not being in line with the NICE reference case.

## DISCUSSION

In light of the recent publications on the use of utility values in oncology, a review of the latest NICE technology appraisals of advanced cancer treatments was conducted to assess the use and elicitation of utility data. The results show that in the majority of cases the requirements of the NICE reference case were not met. EQ-5D was used in only 27% of cases, and, depending on progression status, clinical trials were the main source of data in only 19%–38% of cases. Although often criticized for lack of face validity, on average the difference between pre- and post-progression utilities was 0.197, and on average the values were lower than that of the general population in the same age group.

Despite the criticism of the utility values from the manufacturer submissions, they were very similar to the final values recommended by ERGs and accepted by the Committee, suggesting a lack of better alternatives. This was especially important for the post-progression utilities, where even the values elicited according to the NICE reference case raised concerns regarding face validity. Due to these concerns, the values elicited according to the NICE reference case were occasionally substituted with other types of utilities, such as with directly elicited values.

Thus, despite the new NICE guidance reinforcing the requirements for utility values, the methods used still vary, as in the previous finding by Tosh *et al.*<sup>10</sup> Meanwhile, the use of the EQ-5D is less than half (28%–29%) in these recent oncology appraisals

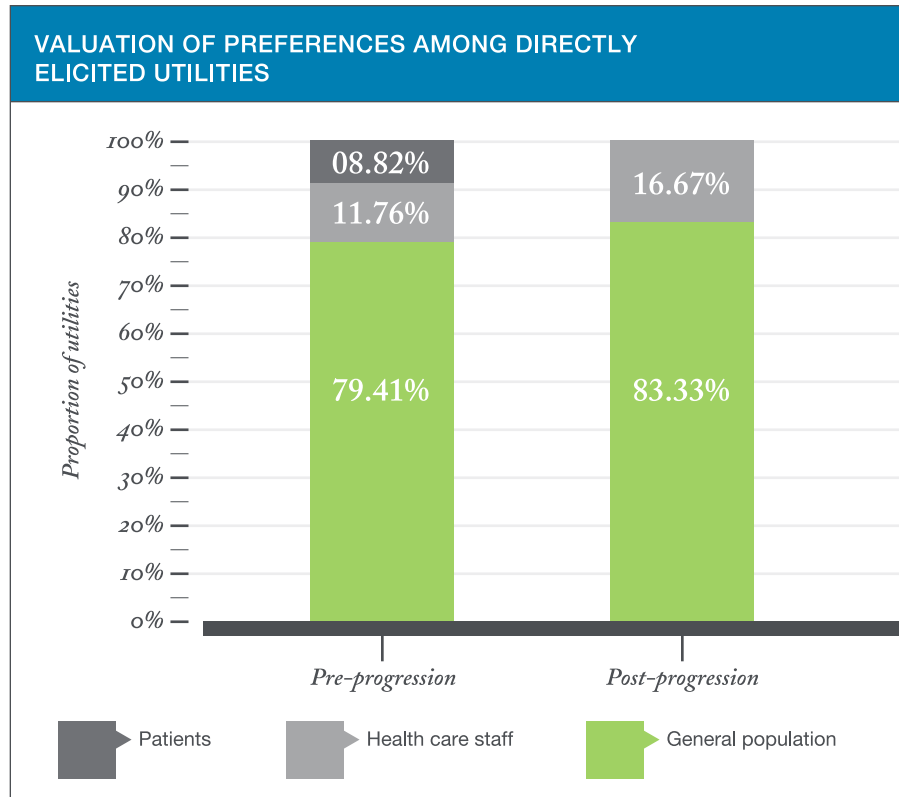


figure 1

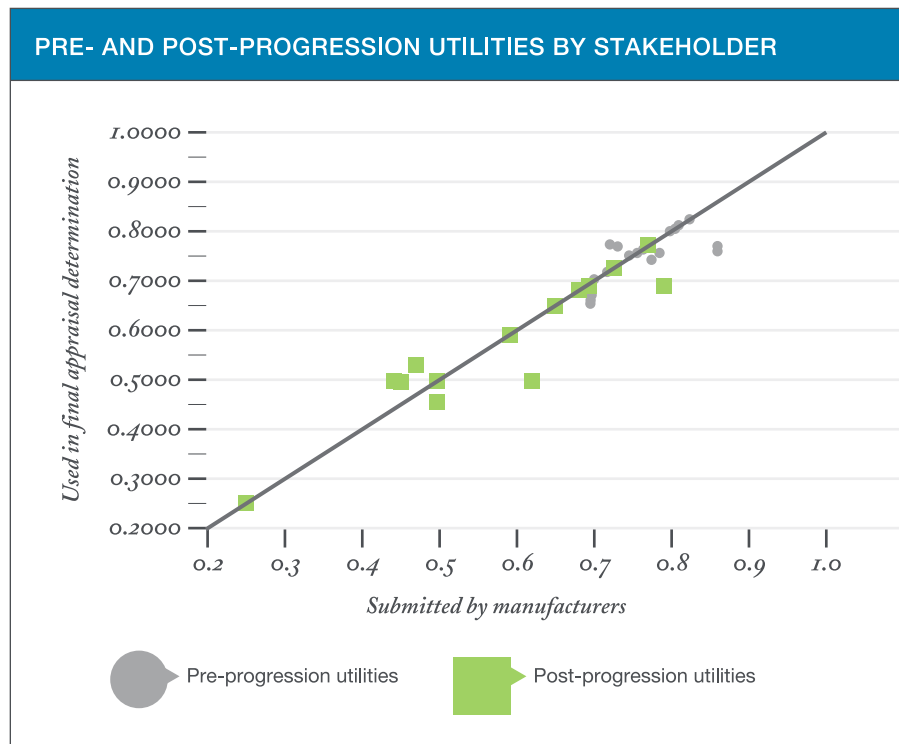


figure 2

---

**UNCERTAINTY AROUND THE UTILITY VALUES CONTRIBUTES TO THE UNCERTAINTY AROUND THE INCREMENTAL COST-EFFECTIVENESS RATIO IN ONCOLOGY; THIS NECESSARILY FOCUSES THE ATTENTION ON THE METHODOLOGIES AND FACE VALIDITY OF THESE UTILITY INPUTS.**

---

compared to all therapeutic areas assessed prior to 2009 (64%). This might be reflective of the concerns expressed in recent publications about the use of EQ-5D in cancer.

In a 2011 review, Garau *et al.*<sup>7</sup> assessed utility valuation in oncology, with special emphasis on EQ-5D in three methodological areas (the description of health states, the valuation of health states, and whose values are taken into account) and identified various potential flaws. Due to the five dimensions and three levels leading to limited number of unique health states (243), the EQ-5D lacks sensitivity. There is work ongoing exploring both the potential increase of the number of dimensions<sup>11</sup> and increasing the number of levels to five.<sup>16</sup> This, however, poses additional questions with regards to complexity and the need to evaluate more health states. Mapping from cancer-specific instruments, such as the Functional Assessment of Cancer Therapy (FACT) or the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire-Core 30 (QLQ-C30), have been explored to overcome this limitation.<sup>8,17</sup> However, mapping has its own issues.<sup>11</sup>

In the valuation of the health states determined by EQ-5D, one of the key issues is the assumption of constant proportional trade-off with the time trade-off method, i.e., the proportion of life expectancy traded off for an improvement of quality of life is constant for the individuals, independent of the length of life expectancy. This assumption may be violated when life expectancy is very short, as among patients with advanced cancer.<sup>14</sup> There is also the issue of the potential difference between the valuation of the general public and the patients themselves. The trade-off, between giving sufficient information to the general public to allow them to assess the health state without bias or misunderstanding and the provision of too much detail that can elicit misconceptions, is an important issue.


Of course, utility values for advanced oncology indications collected in clinical trials have their own set of issues. The timing of the assessment can be crucial with regards to toxicities.<sup>9</sup> Patients are not followed up until death regularly; data collection often stops soon after treatment discontinuation or progression. If they are followed-up, there is a very large attrition rate in quality of life measures even when other measures are available. Thus values that represent the quality of life of patients toward the end of life are usually scarce.

The use of values from trials in classic, three health state Markov models also assumes that patient quality of life changes after radiologic progression. This however may not be the case in indications where the symptomatic progression happens at a much later time point or where quality of life changes with the discontinuation or switching of treatments, rather than progression.

Although the results found in the review seem reflective of the methodological challenges debated in the literature, this study has various limitations. Utility values or method of elicitation were not always available publicly due to reporting or confidentiality (commercial or academic), and these missing data might bias the results. The average age of patients in the modeled cohorts was not available in most cases. Thus the median or mean age of patients in the clinical trials reported as the primary source of the efficacy data were used as a proxy. However, this might not accurately reflect the patient population used in the base case of the model. Thus the interpretation of these results is not straightforward. In addition, the precise effect of these uncertainties on the decisions (i.e., the link between the different aspects of the utilities and their acceptance by the Committee) has not been explored.

## CONCLUSIONS

Uncertainty around the utility values contributes to the uncertainty around the incremental cost-effectiveness ratio in oncology; this necessarily focuses the attention on the methodologies and face validity of these utility inputs. Although the 2013 NICE methods guidance reinforced the need for utilities measured in clinical trials with the help of the EQ-5D questionnaire from patients, methods of elicitation still often do not conform to the NICE reference case. In the post-progression

health state, even values elicited according to the NICE reference case have raised various concerns. These concerns regarding the source, measurement, and interpretation of utility values reflect the recent debates regarding the potential challenges of using EQ-5D values in oncology and stress the importance of methodological development. Also, although the assessment of utilities in advanced oncology indications is crucial in terms of cost-effectiveness, in many cases it is not incorporated or is not incorporated *appropriately* in the design of the Phase 3 clinical trials. 

### *Acknowledgements*

*The authors would like to thank Rebecca Baggaley, Chantelle Browne, Corina Chalkiadaki, David Gabriel, Panagiotis Orfanos, Zachary Smith, Dylan Trundell, and Meng Wang at Evidera for their help in the data extraction.*

For more information, please contact [Noemi.Muszbek@evidera.com](mailto:Noemi.Muszbek@evidera.com), [Agnes.Benedict@evidera.com](mailto:Agnes.Benedict@evidera.com) or [Linda.Hortobagyi@evidera.com](mailto:Linda.Hortobagyi@evidera.com).

### *References*

- 1 NICE. Who We Are. 2014. Available at: [http://www.nice.org.uk/aboutnice/whoweare/who\\_we\\_are.jsp](http://www.nice.org.uk/aboutnice/whoweare/who_we_are.jsp). Accessed Feb. 21, 2014.
- 2 NICE. Guide to the methods of technology appraisal 2013. Available at: <http://publications.nice.org.uk/pmg9>. Accessed 21 February 2014.
- 3 NICE. Guide to the Single Technology Appraisal (STA) Process. 2010. Available at: [http://www.nice.org.uk/aboutnice/howwework/devnicetech/developing\\_nice\\_single\\_technology\\_appraisals.jsp?domedia=1&mid=912F667C-19B9-E0B5-D43AD56E114A62D9](http://www.nice.org.uk/aboutnice/howwework/devnicetech/developing_nice_single_technology_appraisals.jsp?domedia=1&mid=912F667C-19B9-E0B5-D43AD56E114A62D9). Accessed Feb. 21, 2014.
- 4 NICE. Guide to the Multiple Technology Appraisal Process (reference N2022). 2010. Available at: [http://www.nice.org.uk/aboutnice/howwework/devnicetech/developing\\_nice\\_multiple\\_technology\\_appraisals.jsp?domedia=1&mid=9163D61A-19B9-E0B5-D49607127E65DECE](http://www.nice.org.uk/aboutnice/howwework/devnicetech/developing_nice_multiple_technology_appraisals.jsp?domedia=1&mid=9163D61A-19B9-E0B5-D49607127E65DECE). Accessed Feb. 21, 2014.
- 5 NICE. Cancer Reform Strategy. 2007. Available at: <http://www.nhs.uk/NHSEngland/NSF/Documents/Cancer%20Reform%20Strategy.pdf>. Accessed Feb. 21, 2014.
- 6 NICE. Appraising life-extending, end of life treatments. 2009. Available at: <http://www.nice.org.uk/media/E4A/79/SupplementaryAdviceTACEoL.pdf>. Accessed Feb. 21, 2014.
- 7 Garau M, Shah KK, Mason AR, Wang Q, Towse A, Drummond MF. Using QALYs in Cancer: a Review of the Methodological Limitations. *Pharmacoeconomics*. 2011; 29(8):673–685.
- 8 Rowen D, Brazier J, Young T, *et al*. Deriving a Preference-based Measure for Cancer Using the EORTC QLQ-C30. *Value Health*. 2011; 14(5):721–731.
- 9 Tappenden P, Chilcott J, Ward S, Egginton S, Hind D, Hummel S. Methodological Issues in the Economic Analysis of Cancer Treatments. *Eur J Cancer*. 2006; 42(17):2867–2875.
- 10 Tosh JC, Longworth LJ, George E. Utility Values in National Institute for Health and Clinical Excellence (NICE) Technology Appraisals. *Value Health*. 2011; 14(1):102–109.
- 11 Longworth L, Yang Y, Young T, *et al*. Use of Generic and Condition-specific Measures of Health-Related Quality of Life in NICE Decision-Making: a Systematic Review, Statistical Modelling and Survey. *Health Technol Assess*. 2014; 18(9):1–224.
- 12 NICE. Updated Guide to the Methods of Technology Appraisal – June 2008. Available at: <http://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalprocessguides/?domedia=1&mid=B52851A3-19B9-E0B5-D48284D172BD8459>. Accessed Feb. 21, 2014.
- 13 National Cancer Institute. Solid Tumor: Definition. Available at: <http://www.cancer.gov/dictionary?cdrid=45301>. Accessed Feb. 21, 2014.
- 14 Kind P, Hardman G, Macran S. UK Population Norms for EQ-5D. Discussion Paper 172: University of York, Centre for Health Economics; 1999
- 15 Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health State Utilities for Metastatic Breast Cancer. *Br J Cancer*. 2006; 95(6):683–690.
- 16 Pickard AS, De Leon MC, Kohlmann T, Cella D, Rosenbloom S. Psychometric Comparison of the Standard EQ-5D to a 5 Level Version in Cancer Patients. *Med Care*. 2007; 45(3):259–263.
- 17 Pickard AS, Ray S, Ganguli A, Cella D. Comparison of FACT- and EQ-5D-based Utility Scores in Cancer. *Value Health*. 2012; 15(2):305–311.