

Does Methodological Guidance Produce Consistency? A Review of Methodological Consistency in Breast Cancer Utility Value Measurement in NICE Single Technology Appraisals

Micah Rose¹  · Stephen Rice² · Dawn Craig²

Published online: 5 July 2017

© The Author(s) 2017. This article is an open access publication

Abstract Since 2004, National Institute for Health and Care Excellence (NICE) methodological guidance for technology appraisals has emphasised a strong preference for using the validated EuroQol 5-Dimensions (EQ-5D) quality-of-life instrument, measuring patient health status from patients or carers, and using the general public's preference-based valuation of different health states when assessing health benefits in economic evaluations. The aim of this study was to review all NICE single technology appraisals (STAs) for breast cancer treatments to explore consistency in the use of utility scores in light of NICE methodological guidance. A review of all published breast cancer STAs was undertaken using all publicly available STA documents for each included assessment. Utility scores were assessed for consistency with NICE-preferred methods and original data sources. Furthermore, academic assessment group work undertaken during the STA process was examined to evaluate the emphasis of NICE-preferred quality-of-life measurement methods. Twelve breast cancer STAs were identified, and many STAs used evidence that did not follow NICE's preferred utility score measurement methods. Recent STA submissions show companies using EQ-5D and mapping. Academic assessment groups rarely emphasized NICE-preferred methods, and queries about preferred methods were rare. While there appears to be a trend in recent STA submissions towards

following NICE methodological guidance, historically STA guidance in breast cancer has generally not used NICE's preferred methods. Future STAs in breast cancer and reviews of older guidance should ensure that utility measurement methods are consistent with the NICE reference case to help produce consistent, equitable decision making.

Key Points for Decision Makers

A comprehensive review of all completed National Institute for Health and Care Excellence (NICE) single technology appraisals (STAs) and their supporting documents identified 12 STAs with utility scores derived from 14 published utility sources and 3 redacted unpublished sources.

STA guidance published prior to 2016 lacked compliance or exhibited poor compliance with NICE preferred methods for measuring quality-of-life (utility scores), which reduces the consistency of decision making both within and between disease areas.

There appears to be a pattern of improved compliance, beginning in 2016. When NICE periodically updates guidance to reflect new evidence and treatments, it is important that quality-of-life data are concurrently updated to conform to NICE's preferred methods.

✉ Micah Rose
m.j.rose@southampton.ac.uk

¹ Southampton Health Technology Assessments Centre, University of Southampton, The University of Southampton Science Park, Alpha House, Enterprise Road, Southampton SO16 7NS, UK

² Institute of Health and Society, Newcastle University, The Baddiley-Clark Building, Richardson Road, Newcastle upon Tyne NE2 4AX, UK

1 Background

The role of the National Institute for Health and Care Excellence (NICE) is to improve health outcomes using the English National Health Service (NHS) and other public

health and social care services. This is achieved through several research avenues, one of which will be explored in this study: producing evidence-based guidance and advice for health, public health and social care practitioners through the single technology appraisal (STA) process. STAs assess the clinical and cost effectiveness of health technologies to ensure that all NHS patients have equitable access to the most clinically effective and cost-effective treatments available. STAs are designed to appraise a single health technology for a single indication [1].

The STA process consists of several stages designed to create rigorous, transparent decision making with ample stakeholder input into decisions. The primary evidence sources are company evidence submissions and work conducted by an independent Evidence Review Group (ERG) that assesses the validity and quality of the company submissions (CSs). The ERG requests and incorporates clarifications from the company, and conducts additional analyses to address uncertainty and areas where the CS is lacking.

Cost-utility analyses (CUAs) are the primary type of health economic analysis in NICE technology appraisals, a type of cost-effectiveness analysis where the primary measure of benefit is health-related quality-of-life (HRQoL), measured as quality-adjusted life-years (QALYs). In CUAs, HRQoL is generally measured through preference-based utility scores and decrements (utility values).

In 2004, NICE established a reference case defining the preferred methods for economic submissions; guidance on measuring utility scores were as follows: utility scores should be measured from patients or carers with valuation of health states conducted by the general public using a standardised and validated generic choice-based instrument (time trade-off or standard gamble), with EuroQol 5-Dimensions (EQ-5D) as the preferred method of utility measurement [2]. Alternative methods required justification. In 2008, NICE methods guidance recommended considering mapping to EQ-5D from other instruments when EQ-5D was unavailable; this transforms other methods of measuring quality-of-life into EQ-5D data and removed standard gamble from valuation methods [3]. The latest guidance (2013) provides a checklist explicitly asking whether EQ-5D was used [4].

The logic behind recommending one validated instrument is that different methods of measurement produce different results [2–4]. A systematic review in breast cancer utilities by Peasgood et al. found that measurement method substantially influenced utility score, as did whose preferences were measured and whose values were ascribed to those preferences [5]. This measurement inconsistency can

lead to inconsistent decision making, with potential adverse consequences for maximising population health.

2 Objective

This study assesses the consistency in utility measurement methods used in NICE STAs. Breast cancer was chosen as a case study for this review. Because the incidence of breast cancer is consistently among the highest of any cancer in England, with 15.4% of new cancer registrations in 2015 from breast cancer [6], we deemed that there would be a higher likelihood of attaining a good sample size of published STAs.

3 Methods

3.1 Evidence Selection

All published guidance produced by the NICE STA and multiple technology appraisal (MTA) programmes is available on the NICE website [7]. Title searches were conducted among these guidance to identify all published breast cancer STA guidance published between January 2006 and April 2017. The primary evidence documents for an STA are the CS and the ERG report. These documents, as well as the following additional documents, were obtained and checked for utility data: (1) company responses to requests for clarification from the ERG and NICE; (2) errata and addenda; and (3) documents produced by the NICE committee for public consultation and guidance production. Only STAs with published final funding decisions were included. Where any of the above documents referenced external sources, these publications were obtained and reviewed for utility data and additional sources.

3.2 Data Extraction

Data were extracted using a form developed in Microsoft Excel (Microsoft Corporation, Redmond, WA, USA). The extraction form was piloted on two STAs and further refined to ensure the capture of all relevant data. Data were extracted by one researcher (MJR) and checked by a second researcher (SJCR).

A streamlined process was used to extract utility data. First, data were extracted from the CS, and, second, the ERG report was examined. Utility scores and utility decrements were identified in the ERG report, but only extracted where they differed from the CS. NICE appraisal documents other than CSs and ERG reports were checked in the same way but none contained new utility data.

Lastly, studies identified by STAs as utility value sources were extracted.

Data on study and population characteristics, such as year of publication, patient age, menopausal status, breast cancer stage, receptor status, treatments compared, and previous treatments, were extracted. In addition, utility data were extracted for each health state and each adverse event. Health states were commonly represented by utility scores, while adverse events and complications were commonly represented by utility decrements. Collectively, we define these as utility values. Additional data were extracted on measurement methods, additional utility data references, and mapping techniques.

3.3 Synthesis and Analysis

The STAs were first divided into early breast cancer (EBC) and metastatic breast cancer (MBC). The sources and methods for the utility values were then compared with the relevant NICE reference case, which was defined by the year of STA submission. Next, utility scores identified in STAs were analysed for value consistency with their source. Additionally, an analysis of clarification requests by the ERG and NICE was undertaken. Feasibility of mapping was assessed by checking the availability of mapping algorithms using the Health Economics Research Centre (HERC) database of mapping studies [8]. Analyses included assessment of alternative utility values identified by ERGs and used in their exploratory analyses, where these differed from the CS.

4 Results

Twelve breast cancer STAs were identified, four evaluating EBC and eight evaluating MBC. The following sections analyse STA consistency with preferred methods in the NICE reference case, and then consistency with source data and methods.

4.1 Early Breast Cancer Technology Appraisals

There were four STAs relating to EBC: TA 107, TA 108, TA 109, and TA 424 [9–12]. EBC STAs and their utility score sources had broadly similar populations, with some differences in epidermal growth factor receptor, oestrogen receptor and axillary lymph node involvement status, which may affect prognosis but was considered unlikely to affect quality-of-life measurement [13]. No studies in the meta-analysis by Peasgood et al. assessed quality-of-life using the above factors [14].

Three EBC STAs were assessed against 2004 methodological guidance, and one was assessed against 2013

methodological guidance [2, 4]. Across the four EBC STAs, there were 39 utility values (utility scores and utility decrements). Three unpublished sources based on company trials and ten published sources were used for utility values [15–24]. Among utility values, 21% ($n = 8$) were redacted or not reported in the submission document, including all health-state utility scores in TA 107 (only utility decrements from Hillner and Smith were reported [16]) and health-state utility scores from TA 109 for remission (disease-free survival [DFS]) and first locoregional recurrence. Among the redacted or not reported scores, there were six utility scores that did not have the methods reported and two that were derived by mapping to EQ-5D from European Organisation for Research and Treatment of Cancer (EORTC) Quality-of-Life Questionnaire-Core 30 (QLQ-C30) data in the company primary effectiveness trial (the Breast Cancer International Research Group 001 Trial [BCIRG001]), both redacted [11, 25]. In summary, there were 33 utility scores that had methods reported in EBC. Table 1 shows the consistency of unredacted utility value sources, with the relevant NICE reference case for each utility source used in EBC STAs. The ERG on TA 424 listed Essers et al. as the source of alternative utility values in their analyses, however they did not identify the four original source papers these utility values were derived from in the study by Essers et al., i.e. Tengs and Wallace, Carter et al., Hayman et al. and Van Hanswijck de Jonge et al. [20, 21, 23, 24, 26]. Utility scores did not appear to be directly derived from the source papers as these papers were not cited by the ERG. We were unable to confirm utilities by Van Hanswijck de Jonge et al. as the poster presentation was not available to us; therefore, these values are not included in Table 1 or in the following totals. Values derived from the study by Hayman et al. were identical to those derived from the case report by Carter et al., and were therefore also not included [20, 21].

All percentages in this section are based on the 33 utility scores with clear methods reporting. Among the transparently reported utility scores, 36% ($n = 12$) of utility values were derived from patients and 39% ($n = 13$) were valued by the general public. The most consistent application of the reference case in EBC was the use of standard gamble or time trade-off methods, with 45% ($n = 15$) of utility scores using appropriate methods. Only 21% ($n = 7$) of utility scores used EQ-5D values. EBC utilities were often inconsistent with NICE reference case preferences.

Two ERG groups evaluated four CSs of EBC treatments. Of the four EBC STAs, only two included questions submitted to the company for clarification as the NICE STA process was still in its infancy in 2006. The clarification questions submitted to the company in TA 107 did not contain any questions about the derivation of utility scores, and TA 424 primarily used methodologically

Table 1 Consistency of EBC utility sources with the NICE reference case

Utility source, year	STAs that use the utility source	NICE-preferred assumptions (all should be yes)			
		Utilities measured by patients	Valuation by the general public	EQ-5D used	Choice-based method (not rating scale), TTO after 2008
Hillner and Smith 1991 [16]	107, 109 ^a	No	No	No	No
Sorensen et al. 2004 [33]	108 ^a	No	No	No	Yes
Ossa et al. 2004 [18]	109 ^a	No	Yes	No	Yes
Brown and Hutton 1998 [15]	109 ^a	No	No	No	Yes
Launois et al. 1996 [17]	109 ^a	No	No	No	No
Lloyd et al. 2006 [19]	424 ^b	No	Yes	No ^c	No
Lidgren et al. 2007 [22]	424 ^b	Yes	Yes	Yes	Yes
Tengs and Wallace 2000 [23]	424 ^{b,d}	No	Variable	No	No
Carter et al. 1998 [20]	424 ^{b,d}	No	No	No	No

EBC early breast cancer, *EQ-5D* EuroQoL-5 dimensions, *NICE* National Institute for Health and Care Excellence, *STAs* single technology appraisals, *TTO* time trade-off

^a Assessed against 2004 NICE reference case

^b Assessed against 2013 NICE reference case

^c No mapping methods used to derive EQ-5D

^d Referenced as Essers et al. [26], and used only in Evidence Review Group analyses, Essers et al. used multiple sources [33]

appropriate scores, resulting in no questions about NICE-preferred methods [27, 28]. While each of the ERG groups commented on the methods of deriving utility scores in their reports, no commentary indicated an NICE preference for EQ-5D data [29–32]. In fact, as Table 1 shows, one ERG group that examined company utility values derived from EQ-5D explicitly chose to test alternative utility values that consistently did not use NICE-preferred methods or comply with the 2013 NICE reference case [4, 32].

Utility scores used in EBC had utility values that were generally consistent with their cited sources, but sources rarely used NICE-preferred methods. Among studies assessed against the 2004 NICE reference case, Hillner and Smith used a rating scale [16], Sorensen et al. used a chained standard gamble method [33, 34], Brown and Hutton used the standard gamble method [15], Ossa et al. used a bespoke time trade-off method based on the EQ-5D [18], and redacted scores in submissions used undisclosed scales or EQ-5D. One STA was assessed against the 2013 NICE reference case, TA 424. Between the CS and the ERG report, there were five sources for utility scores reported, i.e. Lidgren et al., Lloyd et al., Tengs and Wallace, Carter et al., and Van Hanswijck de Jonge et al. [19, 20, 22–24]. The company used five utility scores from Lidgren et al. and one score from Lloyd et al. All five scores from Lidgren et al. were derived by mapping EORTC QLQ-C30 to EQ-5D. Lloyd et al. conducted a standard gamble with regression to derive scores.

The ERG used alternative utility scores from Essers et al. [26], which were derived from three studies not

cited by the ERG [20, 23, 24]. Tengs and Wallace gathered together a large sample of utility scores derived from a variety of studies using a variety of methods; however, none that were cited in the study by Essers et al. complied with the 2013 NICE guidance [4, 23, 26]. Carter et al. reported on a case study of a 74-year-old woman who used standard gamble elicitation derived from healthcare workers to assess her quality-of-life [20]. The utility value reported by Essers et al. does not perfectly match the value reported by Carter et al. [20, 26]. The third study was a poster that we could not acquire [24]. Mixing these different instruments could provide inconsistent results as it is well-known that alternative methods of utility measurement produce different results [2].

4.2 Metastatic Breast Cancer Technology Appraisals

There were eight technology appraisals of MBC: TA 116, TA 214, TA 239, TA 250, TA 263, TA 295, TA 371, and TA 423 [12, 35–41]. Some appraisals were for locally advanced breast cancer (LABC) or MBC. Similar to how human epidermal growth factor receptor 2 (HER2) and oestrogen receptor status in EBC were handled, the same utility scores were used for LABC as those used for MBC. The populations for the STAs were broadly similar, with most technology appraisals using utility scores derived from one study, i.e. Lloyd et al. [19]. No utility values were redacted in MBC STAs.

All STAs were assessed against the NICE reference case relevant to the year of their production. Only TA 116 was assessed against the 2004 reference case [40], and six STAs were assessed against the 2008 reference case: TA 214, TA 239, TA 250, TA 263, TA 295 and TA 371 [35–39, 46]. While TA 371 was published after the 2013 NICE reference case had been published, the STA began before the 2013 methods guidance had been published, therefore it has also been assessed against the 2008 reference case. One MBC STA was assessed against the 2013 NICE reference case, i.e. TA 423 [41]. In total, 135 utility values were reported.

Six sources were used among the eight MBC STAs. Evaluating the six sources for consistency with the NICE reference case is more parsimonious than evaluating each of the 135 utility values. Table 2 presents the six MBC utility sources and assesses them against the NICE reference case. Five of six sources used regression methods to derive utility scores [5, 19, 41–43]. Lloyd et al. built a regression based on chained standard gamble results from the general population [19], while Cooper et al. [42] used a meta-regression of utility scores for MBC, with most values derived from three sources, all of which used the standard gamble approach, with oncology nurses and oncologists providing responses and valuations [15, 17, 44]. Peasgood et al. conducted a systematic review with separate meta-regressions of utility scores in EBC and MBC, derived using a variety of measurement methods and a variety of different groups of individuals valuing health states; the meta-regressions included many covariates, including valuation methods [5]. TA 423 used the mapping algorithm developed by Crott and Briggs to map EORTC-QLQC30 trial data to EQ-5D utility scores for both health states and for utility decrements due to adverse events [41, 43, 45].

As Table 2 shows, most utility scores used in STAs of MBC did not use the preferred methods of utility measurement specified in the relevant NICE methods guidance. From 2006 onwards, seven of eight CSs for MBC STAs used Lloyd et al., a source that did not use NICE-preferred methods [35–41, 46]. All eight ERG reports for MBC STAs used utility values derived from Lloyd et al. [32, 47–53]. As a consequence, many of the utility values are out of line with NICE-preferred methods used in other STAs in other disease areas, making the utility values in breast cancer difficult to compare with those in other areas—the currency of comparison is not universal and there is no means of conversion [19].

Peasgood et al. found that measurement methodology (who measured, who valued, what instrument) was often highly influential on utility values [5]. While the work of Peasgood et al. was used for sensitivity analysis in TA 214, no explanation was given as to which regression variables

were utilised and how the utility scores were generated [5]. As such, no STA conducted on breast cancer has produced cost-effectiveness results that can be compared like-for-like with other disease areas.

The eight MBC STAs included three ERG groups and five companies [35–41, 46]. Only one ERG group explicitly expressed NICE's preference for EQ-5D data and queried whether mapping algorithms to EQ-5D were considered [50]. The company reply indicated that no search had been conducted but that a current trial was collecting both EQ-5D and FACT-B data with a goal of producing such an algorithm, with data expected in the next 18–24 months (that submission was in 2010 [54]). The company has since had three NICE STA guidance published using the study by Lloyd et al. for utilities, with the latest submission produced in 2013 [19, 36, 46, 55].

In 2008, NICE officially advocated mapping methods to convert other quality-of-life scales to EQ-5D [3]. Only the last of the seven breast cancer STAs produced after the 2008 NICE methods guidance used mapping to produce EQ-5D values [35–39, 46]. Each of the six STAs that did not use mapping algorithms used trial evidence with effectiveness measures that could be mapped to EQ-5D at the time of submission to NICE.

Four clinical effectiveness studies used in STAs used the FACT-B disease-specific quality-of-life instrument. The E2100 trial was used by TA 214 (submitted May 2010) and TA 263 (submitted December 2011) for effectiveness evidence [36, 54, 56], while the CONFIRM trial was used for effectiveness evidence in TA 239 (submitted November 2011) [38, 57]. The study referred to as Study 201 was used in the CS for TA 250 (submitted March 2011) [37, 58], and the EMILIA trial was used for effectiveness evidence in TA 371 (submitted December 2013) [46, 59, 60]. FACT-B contains the FACT-G cancer questionnaire plus additional breast cancer-specific questions [61]. Cheung et al. and Teckle et al. mapped FACT-G to EQ-5D in a breast cancer population [62, 63]. During TA 214, the company in TA 263 indicated that an in-process trial using FACT-B and EQ-5D data would produce a mapping algorithm [54], and use of individual patient data in this algorithm could allow sophisticated control for adverse events and disease states. Examining the latest version of the HERC database of mapping studies identified no published algorithm based on that trial [8].

The Bolero-2 trial was used for effectiveness evidence in TA 295 (submitted February 2013); Bolero-2 used the EORTC QLQ-C30 and the breast cancer-specific questionnaire (QLQ-BR23) [2, 35, 64]. EORTC QLQ-C30 was mapped to EQ-5D by Kind [25], Crott and Briggs (breast cancer population) [45], and Kim et al. (MBC population) [65]. Additionally, Kim et al. included the QLQ-BR23 in mapping [65]. In 2013, mapping was possible at the time of

Table 2 Comparison of MBC utility methods with NICE reference case

Utility source, year	STAs utilising source	NICE preferred assumptions (all should be yes)				Mapping used?
		Utilities measured by patients	Valuation by the general public	EQ-5D used	Choice-based method (not rating scale); if after 2008, TTO	
Lloyd et al. 2006 [19] ^f	116	No	Yes	No	Yes	Not relevant
Lloyd et al. 2006 [19] ^{g,h}	214 ^a , 239, 250, 263 ^b , 295, 371 ^d , 423 (ERG)	No	Yes	No	No	No
Cooper et al. 2003 [42] ^g	214 ^c	No	No	No	No	No
Hutton et al., 1996 [44] ^g	250	No	No	No	No	No
Peasgood et al. 2010 [5] ^h	214 ^a	Partially	Partially	Partially	Partially	No
TA 423 CS [41] ^h	423	Yes	Yes	Yes	Yes	Yes ^e
Hudgens et al. 2014 [43] ^h	423	Yes	Yes	Yes	Yes	Yes ^e

MBC metastatic breast cancer, EQ-5D EuroQoL-5 dimensions, TTO time trade-off, CS company submission, ERG Evidence Review Group report, NICE National Institute for Health and Care Excellence, STAs single technology appraisals

^a Used only in sensitivity analysis

^b Cites Fleeman et al. [51], used the Lloyd et al. [19], regression but fixed age at 47 years and weighted by treatment response

^c Cited as Winstanley et al. [87] in the company submission but all utility values are derived from Cooper et al. [42]

^d TA371 began in 2012, before the 2013 NICE Methods Guidance was released [4]

^e Crott and Briggs [45]

^f Assessed against 2004 NICE reference case

^g Assessed against 2008 NICE reference case

^h Assessed against 2013 NICE reference case

submission for TA 295. No justification was provided as to why mapping methods were not used by the company.

Table 3 provides a checklist for assessing consistency with the NICE reference case with regard to utility measurement. The checklist is designed to force technology appraisal participants (whether the company or appraisal group) to fully examine consistency with the NICE reference case, help formulate clarification requests and requests for additional analyses to be sent to the company, and encourage ERGs to ensure that they are knowledgeable about available mapping algorithms for the disease area being studied in a technology appraisal.

5 Discussion

The results of this review show that NICE methodological guidance has not led to consistent application of NICE-preferred methods of utility measurement in breast cancer STAs. While mapping algorithms were often available, they were rarely assessed, recommended, or used. In one instance, a company stated that their trial data would allow mapping [54], and subsequently submitted five submissions

that did not use a mapping algorithm generated from that trial [12, 36, 46, 55, 66]. In addition to nonreference methods being prevalent across breast cancer STAs, clinical guideline (CG) 81 and an HTA of lapatinib plus trastuzumab with an aromatase inhibitor also used utility values that were not valued using EQ-5D or mapped to EQ-5D [47, 67]. These guidance were produced with direct NICE oversight or by an NICE-contracted technology appraisal group.

The findings of this review relate to breast cancer STAs conducted by NICE for the English NHS. It is unclear whether the results are generalisable to other disease areas. We have not made, nor can we make, any conclusions about methodological consistency in utility score measurement in NICE technology appraisals in general.

The review did not formally assess other forms of NICE guidance, such as MTAs, diagnostic guidance (DG), or CG. For completeness, we have conducted rapid informal reviews of the utility scores used in other breast cancer guidance produced by and for NICE. We identified two NICE CGs related to breast cancer: CG 80 and CG 81 [67, 68]. CG 80 did not use QALYs or utility scores [68], whereas CG81 used utility scores that were derived from

Table 3 Reference case assessment checklist for technology appraisal submissions

Was EQ-5D used to value health?	
Yes	No
Were utility scores derived from the relevant patient population?	Were utility scores derived from the relevant patient population?
Were values for these utility scores derived from the general population?	Were values for these utility scores derived from the general population?
Was EQ-5D measured directly, or mapped from another instrument?	Was the preferred time trade-off utility score measurement method used?
	Were any mapping algorithms available and not used?
	Was sufficient justification provided for the choice of utility score measurement?

EQ-5D EuroQoL-5 dimensions

Cooper and colleagues and were thus inconsistent with NICE-preferred methods, as stated earlier [42, 67]. Four additional technology appraisals were identified through the NICE website (TA 34) and through CG 81, which replaced TA 30, TA 54, and TA 62, all of which were not available from the NICE website, and, although TA 34 was available, it did not use QALYs or utility scores [69]. CG 81 did not use methodologically appropriate HRQoL data [67]. Two DGs were identified: DG 8 and DG 10. DG 8 used health-state utility scores that were inconsistent with preferred methods, but did include utility decrements that were consistent with NICE methodology; the majority of utility values used in DG 8 were inconsistent with NICE methods [70]. The utility values in DG 10 used NICE-preferred methods [71]. A limited amount of data were available for TAs produced prior to 2002, but it is clear from other published guidance in breast cancer that inconsistency with methodological guidance has historically been a problem in these evidence streams, in a similar fashion to STAs.

The findings of this study are similar to previous work evaluating utility value measurement in STAs. Tosh et al. [72] found that only 56% of values used in technology appraisals were compliant with the 2004 NICE reference case [2], and only 49% used EQ-5D [72]. Technology appraisal participants have shown a lack of awareness of the full specification of NICE's methodological preferences for utility measurement, and have not emphasised mapping to EQ-5D. This can be remedied in future technology appraisals by working from the full reference case guidance and using the framework laid out in Table 3 for assessing utilities.

The study by Lloyd et al. was used most across the 12 published STAs, and was used in some form by all eight MBC STAs and one EBC STA [19]. Their study was conducted by United BioSource Corporation, a consultancy company, and Eli Lilly, a pharmaceutical company and one of the companies that made a submission to NICE [19, 40].

In the study by Lloyd et al., EQ-5D was collected at baseline in a nonclinical valuation population but not in a clinical study for the disease population [19]. One ERG stated that Lloyd et al. might have been the best available utility data; however, this ERG did not consider whether a mapping algorithm could have been used to produce EQ-5D utilities. Using utility scores derived by mapping as the base case, and the values used in the study by Lloyd et al. utilities as a sensitivity analysis would have been in line with NICE guidance, would allow transparent discussion of any limitations of mapping algorithms, and would allow decision makers to decide whether they prefer consistency within breast cancer STAs or consistency with methodological guidance [3, 19].

There are many mapping studies but support for mapping methods is not universal. Doble and Lorgelly assessed the external validity of mapping algorithms from EORTC QLQ-C30 to EQ-5D and found that existing algorithms generally perform poorly, with seven of ten algorithms having statistically significantly different observed and predicted QALYs [73]. The best performing algorithm was the most complicated and computationally intensive algorithm, which limits its application [73, 74]. Pennington and Davis found that choice of mapping algorithm could significantly affect cost-effectiveness results [75]. Introducing mapping algorithms may add complexity to the technology appraisal process without producing better estimates of EQ-5D valuations. The impending transition from EQ-5D-3L to EQ-5D-5L now that a UK scoring algorithm for EQ-5D-5L is available adds further complications to the debate over the appropriateness of mapping as the greater sensitivity of EQ-5D-5L may make it more amenable to mapping [76, 77].

A search for STAs currently in development revealed eight STAs [78]; four were suspended, withdrawn, or discontinued before any evidence submissions were received from the company. In addition, one STA was suspended after evidence submission by the company and analysis by

the ERG; this STA used utility scores from the study by Lloyd et al., with no objection from the ERG and no request for additional analyses using NICE-preferred methods [79]. NICE currently has four guidance actively in development for MBC [78], two of which do not have publicly available submissions from the manufacturers [80, 81]. One STA began in 2013, and, in this STA, the CS and ERG report used the study by Lloyd et al. for utility data [19, 46, 82]. The most recent STA (CS provided in September 2016) to have CSs, clarification questions, and an ERG report conforms well to NICE-preferred methods [83, 84].

Why EQ-5D values were so infrequently used in CSs is an unanswered question. An in-progress systematic review, currently published only as a conference abstract, identified 17 studies that provided EQ-5D utility scores for breast cancer [85]. Only three published breast cancer STAs used EQ-5D values. Two EBC STAs used EQ-5D—one used mapping (TA 109, scores redacted) and one used directly measured EQ-5D scores (TA 424). The clinical evidence sources used in the companies' submissions for MBC all had data that could be mapped to EQ-5D at the time of submission. Only one MBC STA used mapping from EORTC QLQ-C30 to derive appropriate EQ-5D utility scores and received criticism for using these appropriate methods [41, 48]. We believe there are three potential explanations for why alternative utility values have not been used: (1) alternative values have been identified and are unfavourable to companies; (2) there is a lack of belief in mapping methodology; (3) and/or the values from the study by Lloyd et al. are viewed as accepted and further work viewed as unnecessary or inappropriate in spite of the dataset's inconsistency with NICE-preferred methods. None of these potential explanations helps foster informed consistent decision making across disease areas.

The last two published MBC guidance (TA 423 and TA 424) both used EQ-5D utility scores, with one STA using mapping to derive these values [12, 41]. The most recent breast cancer STA guidance in development also used EQ-5D [83]. It appears probable that NICE, ERGs and companies are improving compliance to NICE's methods; however, STA guidance issued prior to this improvement in compliance are still binding and still have the potential to cause inequitable decisions. The NICE should ensure that when these guidance are updated, the utility methods in these guidance are also made conformant with NICE-preferred methods.

While some ERGs have criticised utility score usage in STAs, the primary focus in breast cancer STAs has been on survival rather than utilities, which is not inappropriate. Time is limited in STAs. Only 8 weeks are available to evaluate a multi-hundred page document, formulate clarification questions to the company, incorporate new data

and analyses from the company clarification response into ERG analysis, check an executable model, and produce a report usually spanning more than 100 pages [86]. It is important for ERGs to focus on the factors most likely to affect cost-effectiveness conclusions, and survival data frequently have the largest impact on cancer model outcomes.

However, ERGs have frequently failed to comment on the consistency of utility scores in breast cancer STAs with NICE-preferred methods. In fact, in TA 423, the ERG advised using noncompliant data from the study by Lloyd et al. [19], and, in TA 424, the ERG used utility scores from sources that did not comply with the NICE reference case for their alternative utility value analyses [32, 48]. We believe that failure to use NICE preferences may be prevented through the use of Table 3.

In the current methodological guidance (2013), the section on valuing health effects for the reference cases has 12 subsections, with the use of EQ-5D being reiterated in nine of these [4]. The emphasis in the guidance is clearly on consistency across different disease areas and between NICE guidance. Unfortunately, the written emphasis has not, until very recently, been followed by consistent actions.

In breast cancer STAs, NICE and ERG groups have often not requested analyses, and have frequently not conducted supplemental analyses, using EQ-5D utilities. The quantity of EQ-5D mapping algorithms for breast cancer populations has increased, as evidenced in the HERC database of mapping studies [8]. This database facilitates the use of mapping methods to transform clinical effectiveness data into EQ-5D utility scores in current and future breast cancer STAs. NICE needs to clarify whether the use of mapping methodologies is their preferred second option, and, if it is, they should cease to accept utility values derived from noncompliant measurement methods in breast cancer STAs (and more broadly), and ensure that their preference for EQ-5D utility scores is adhered to, otherwise NICE's repeated emphasis of their preference for EQ-5D has little meaning and they increase the risk of inconsistent decision making.

6 Conclusion

Historically, breast cancer STAs have shown a broad lack of compliance with preferred methods for measuring utility scores in the NICE reference case. However, the last two guidance published in December 2016 adhere to the NICE reference case, and the latest STA guidance in development also conforms to NICE-preferred methods [12, 41, 83]. It will be the continued responsibility of the ERG groups and NICE to ensure that future guidance and

reviews of guidance maintain this forming trend. It would be inappropriate to let noncompliant utilities used in past guidance remain the preferred utilities in future guidance. To fail to update to methodologically appropriate utility scores when updating guidance would be to grant these guidance using methods that are inconsistent with the NICE reference case a durable precedent, which is a decision that should not stand.

Author Contributions MR was responsible for revision of the research proposal, drafting the protocol, collaboratively designing the data extraction form, piloting the extraction form, performing data extraction and checking, synthesising findings, and writing all drafts of this article. SR contributed to the production of the protocol, collaboratively designing the data extraction form, piloting the extraction form, performing data extraction and checking, providing feedback, and editing all drafts of this article. DC identified the topic, wrote the initial research proposal, contributed to the production of the protocol, had oversight of the review, and provided feedback and editing of all drafts of this article.

Compliance with Ethical Standards

This research involved no human or animal participation.

Funding This study applied for and received no funding.

Conflict of interest Micah Rose, Stephen Rice and Dawn Craig declare no conflicts of interest.

Data Availability Statement The data identified and analysed during the current study are available from the corresponding author upon reasonable request.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. National Institute for Health and Care Excellence (NICE). Guide to the single technology appraisal process 2009. London: NICE; 2009.
2. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2004. London: NICE; 2004.
3. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2008. London: NICE; 2008.
4. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. London: NICE; 2013.
5. Peasgood T, Ward SE, Brazier J. Health-state utility values in breast cancer. *Expert Rev Pharmacoecon Outcomes Res.* 2010;10(5):553–66.
6. Office for National Statistics. Cancer registration statistics, England: first release, 2015 Cancer diagnoses and age-standardised incidence rates for all cancer sites by age and sex: Office for National Statistics; 2017.
7. National Institute for Health and Care Excellence (NICE). NICE published guidance and advice list. 2017. Available at: <https://www.nice.org.uk/guidance/published?type=ta>. Accessed 3 Apr 2017.
8. Dakin H. Review of studies mapping from quality of life or clinical measures to EQ-5D: an online database. *Health Qual Life Outcomes.* 2013;11:151. doi:10.1186/1477-7525-11-151.
9. Bristol-Myers Squibb Pharmaceuticals Ltd. Taxol® (paclitaxel) for the adjuvant treatment of early breast cancer: submission to the National Institute for Health and Clinical Excellence (CS TA108); 2006.
10. Roche. Herceptin (trastuzumab): NICE submission. Achieving clinical excellence in the adjuvant treatment of her-2 positive breast cancer (CS TA107). 2006.
11. Sanofi-Aventis. Docetaxel for the adjuvant treatment of early node-positive breast cancer. Single Technology Appraisal. Submission to the National Institute for Health and Clinical Excellence (CS TA109). 2006.
12. Roche. Neoadjuvant Parjeta (pertuzumab) for the treatment of HER2-positive early breast cancer (CS TA424). 2016.
13. Cheang MCU, Chia SK, Voduc D, Gao D, Leung S, Snider J, et al. Ki67 Index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst.* 2009;101(10):736–50.
14. Peasgood T, Ward S, Brazier J. A review and meta analysis of health state utility values in breast cancer. Discussion paper. Sheffield: The University of Sheffield; 2010.
15. Brown RE, Hutton J. Cost-utility model comparing docetaxel and paclitaxel in advanced breast cancer patients. *Anticancer Drugs.* 1998;9(10):899–907.
16. Hillner BE, Smith TJ. Efficacy and cost effectiveness of adjuvant chemotherapy in women with node-negative breast cancer. a decision-analysis model. *N Engl J Med.* 1991;324(3):160–8.
17. Launois R, Reboul-Marty J, Henry B, Bonnetere J. A cost-utility analysis of second-line chemotherapy in metastatic breast cancer. Docetaxel versus paclitaxel versus vinorelbine. *PharmacoEconomics.* 1996;10(5):504–21.
18. Ossa DF, Briggs A, Cowell W, Littlewood T, Sculpher M. Utility associated with severity of cancer-related anaemia (CRA): a societal valuation. *Value Health.* 2004;7(6):680.
19. Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. *Br J Cancer.* 2006;95(6):683–90.
20. Carter K, Ritchey N, Castro F, Caccamo L, Kessler E, Erickson B, et al. Treatment of early-stage breast cancer in the elderly: a health-outcome-based approach. *Med Decis Mak.* 1998;18:213–9.
21. Hayman J, Hillner B, Harris J, Weeks J. Cost-effectiveness of routine radiation therapy following conservative surgery for early-stage breast cancer. *J Clin Oncol.* 1998;16(3):1022–9.
22. Lidgren M, Wilking N, Jonsson B, Rehnberg C. Health related quality of life in different states of breast cancer. *Qual Life Res.* 2007;16(6):1073–81.
23. Tengs T, Wallace A. One thousand health-related quality-of-life estimates. *Med Care.* 2000;38(6):583–637.
24. Van Hanswijck de Jonge P, Doyle S, Farina C, Walker M. Poster presentation 190 P: elicitation of UK Health utilities in primary, recurrent and metastatic breast cancer. In: 31st European Society Medical Oncology (ESMO), 29 Sep–3 Oct 2006, Istanbul.
25. Kind P. Measuring the value of quality of life in cancer: an index based on EORTC QLQC-30. *J Clin Oncol.* 2005;23(16):531S–S.
26. Essers B, Seferina S, Tjan-Heijnen V, Severens J, Novák A, Pompen M, et al. Transferability of model-based economic evaluations: the case of trastuzumab for the adjuvant treatment of

- HER2-positive early breast cancer in the Netherlands. *Value Health*. 2010;13(4):375–80.
27. National Institute for Health and Care Excellence (NICE). Trastuzumab for early breast cancer clarification questions (PFC1, TA107). 2006.
 28. National Institute for Health and Care Excellence (NICE). Trastuzumab for early breast cancer clarification questions (PFC2, TA107). 2006.
 29. Chilcott J, Lloyd Jones M, Wilkinson A. Docetaxel for the adjuvant treatment of early node-positive breast cancer: a single technology appraisal (ERG TA109). Sheffield: The School of Health and Related Research (ScHARR); 2006.
 30. Ward S, Pilgrim H, Hind D. Trastuzumab for the treatment of primary breast cancer in HER2 positive women: a single technology appraisal (ERG TA107). Sheffield: University of Sheffield, School of Health and Related Research (ScHARR); 2006.
 31. Centre for Health Economics (CHE), Regional Drug and Therapeutics Centre. The use of paclitaxel in the management of early stage breast cancer (ERG TA108). 2006.
 32. Squires H, Pandor A, Thokala P, Stevens J, Kaltenthaler E, Clowes M. Pertuzumab for the neoadjuvant treatment of HER2-positive breast cancer (ERG TA424). 2016.
 33. Sorensen SV, Brown R, Benedict A, Flood E, Revicki D. Patient-rated utilities in postmenopausal early breast cancer (EBC): a cross-country comparison. *Value Health*. 2004;7(6):641–2.
 34. Locker GY, Mansel R, Cella D, Dobrez D, Sorensen S, Gandhi SK, et al. Cost-effectiveness analysis of anastrozole versus tamoxifen as primary adjuvant therapy for postmenopausal women with early breast cancer: a US healthcare system perspective. The 5-year completed treatment analysis of the ATAC ('Arimidex', Tamoxifen Alone or in Combination) trial. *Breast Cancer Res Treat*. 2007;106(2):229–38.
 35. Novartis. Single technology appraisal (STA): Everolimus (Afinitor) in combination with exemestane for the treatment of advanced or metastatic HER2 negative, hormone receptor positive breast cancer after prior endocrine therapy (CS TA295): Novartis; 2012.
 36. Roche. Single technology appraisal (STA): Bevacizumab in combination with capecitabine for the treatment of metastatic breast cancer (CS TA263): Roche; 2011.
 37. Eisai Ltd. Single technology appraisal (STA): Eribulin for the treatment of locally advanced breast cancer or metastatic breast cancer (CS TA250). 2011.
 38. AstraZeneca UK Ltd. Single Technology Appraisal (STA): Fulvestrant 500 mg for the treatment of postmenopausal women with oestrogen receptor positive, locally advanced or metastatic breast cancer (CS TA239). 2010.
 39. Roche. Bevacizumab in combination with taxanes for the treatment of HER2-negative 1st line metastatic breast cancer. Roche: submission to the National Institute for Health and Clinical Excellence (CS TA214). 2010.
 40. Eli Lilly and Company Ltd. Gemcitabine for the treatment of metastatic breast cancer. Single technology appraisal (STA). Submission to the National Institute for Health and Clinical Excellence (CS TA116). 2006.
 41. Eisai Ltd. Eribulin for treating locally advanced or metastatic breast cancer after chemotherapy (CS TA 423). 2016.
 42. Cooper NJ, Abrams KR, Sutton AJ, Turner D, Lambert PC. A Bayesian approach to Markov modelling in cost-effectiveness analyses: application to taxane use in advanced breast cancer. *J R Stat Soc Ser A Stat Soc*. 2003;166:389–405.
 43. Hudgens S, Briggs A, Velikova G, Forsythe A, McCutcheon S, Kind P. Impact of treatment with eribulin or capecitabine for metastatic breast cancer (MBC) on EQ-5D utility derived from EORTC QLQ-C30. In: European Society of Medical Oncology; 2014.
 44. Hutton J, Brown R, Borowitz M, Abrams K, Rothman M, Shakespeare A. A new decision model for cost-utility comparisons of chemotherapy in recurrent metastatic breast cancer. *Pharmacoeconomics*. 1996;9(Suppl 2):8–22.
 45. Crott R, Briggs A. Mapping the QLQ-C30 quality of life cancer questionnaire to EQ-5D patient preferences. *Eur J Health Econ*. 2010;11(4):427–34.
 46. Roche. Trastuzumab emtansine for treating HER2-positive, locally advanced or metastatic breast cancer after treatment with trastuzumab and a taxane (CS TA371): Roche; 2013.
 47. Fleeman N, Bagust A, Boland A, Dickson R, Dundar Y, Moonan M, et al. Lapatinib and trastuzumab in combination with an aromatase inhibitor for the first-line treatment of metastatic hormone receptor-positive breast cancer which over-expresses human epidermal growth factor 2 (HER2): a systematic review and economic analysis. *Health Technol Assess*. 2011;15(42):1–93, iii–iv.
 48. Fleeman N, Bagust A, Richardson M, Houten R, Krishan A, Beale S, et al. Eribulin for treating locally advanced or metastatic breast cancer after chemotherapy [ID964] (ERG Report TA 423): Liverpool Reviews and Implementation Group; 2016.
 49. Fleeman N, Beale S, Davis H, Dickson R, Loncaster J, O'Reilly S, et al. Bevacizumab in combination with capecitabine for the first-line treatment of metastatic breast cancer: a Single Technology Appraisal (ERG TA263). Liverpool: Liverpool Reviews & Implementation Group (LRiG), University of Liverpool; 2011.
 50. Rodgers M, Soares M, Epstein D, Yang H, Fox D, Eastwood A. Bevacizumab in combination with a taxane for the first-line treatment of HER2-negative metastatic breast cancer (ERG TA214). York: Center for Reviews and Dissemination (CRD), CHE Technology Assessment Group; 2010.
 51. Fleeman N, Bagust A, Beale S, Blundell M, Dwan K, Pilkington G, et al. Everolimus in combination with an aromatase inhibitor for the treatment of breast cancer after prior endocrine therapy: a Single Technology Appraisal. LRiG, The University of Liverpool (ERG TA295). Liverpool: Liverpool Reviews and Implementation Group (LRiG), The University of Liverpool; 2013.
 52. Bagust A, Blundell M, Boland A, Dickson R, Dundar Y, Fleeman N, et al. Fulvestrant for the treatment of locally advanced or metastatic breast cancer (ERG TA239). Liverpool: Liverpool Reviews and Implementation Group (LRiG); 2011.
 53. Jones J, Takeda A, Tan SC, Cooper K, Loveman ECA, Murray N. Evidence Review Group Report commissioned by the NHS R&D HTA Programme on behalf of NICE: gemcitabine for metastatic breast cancer (ERG TA116). Southampton: Southampton Health Technology Assessments Centre; 2006.
 54. Roche. Bevacizumab in combination with a taxane for the first-line treatment of metastatic breast cancer: manufacturer response to points for clarification (PFC TA214). London. 2010.
 55. Roche. Single technology appraisal—bevacizumab in combination with capecitabine for the first-line treatment of metastatic breast cancer: points for clarification (PFC TA263). Manchester: Roche; 2012.
 56. Eastern Cooperative Oncology Group. Paclitaxel with or without bevacizumab in treating patients with locally recurrent or metastatic breast cancer. NCT00028990. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US); 2002.
 57. AstraZeneca. Comparison of fulvestrant (FASLODEX™) 250 mg and 500 mg in postmenopausal women with oestrogen receptor positive advanced breast cancer progressing or relapsing after previous endocrine therapy (CONFIRM). NCT00099437. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US); 2004.
 58. Eisai, Inc. A study of E7389 in advanced/metastatic breast cancer patients. NCT00097721. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US); 2004.

59. Verma S, Miles D, Gianni L, Krop IE, Welslau M, Baselga J, et al. Trastuzumab emtansine for HER2-positive advanced breast cancer. *N Engl J Med*. 2012;367(19):1783–91.
60. Welslau M, Dieras V, Sohn JH, Hurvitz SA, Lalla D, Fang L, et al. Patient-reported outcomes from EMILIA, a randomized phase 3 study of trastuzumab emtansine (T-DM1) versus capecitabine and lapatinib in human epidermal growth factor receptor 2-positive locally advanced or metastatic breast cancer. *Cancer*. 2014;120(5):642–51.
61. FACIT.org. FACT questionnaires. Available at: <http://www.facit.org/FACITOrg/Questionnaires>. Accessed 8 May 2014.
62. Cheung Y-B, Thumboo J, Gao F, Ng G-Y, Pang G, Koo W-H, et al. Mapping the English and Chinese versions of the Functional Assessment of Cancer Therapy-General to the EQ-5D Utility Index. *Value Health*. 2009;12(2):371–6.
63. Teckle P, McTaggart-Cowan H, Van der Hoek K, Chia S, Melosky B, Gelmon K, et al. Mapping the FACT-G cancer-specific quality of life instrument to the EQ-5D and SF-6D. *Health Qual Life Outcomes*. 2013;11:203.
64. Novartis Pharmaceuticals. Everolimus in combination with exemestane in the treatment of postmenopausal women with estrogen receptor positive locally advanced or metastatic breast cancer who are refractory to letrozole or anastrozole (BOLERO2). NCT00863655. *ClinicalTrials.gov*. Bethesda (MD): National Library of Medicine (US); 2009.
65. Kim EJ, Ko SK, Kang HY. Mapping the cancer-specific EORTC QLQ-C30 and EORTC QLQ-BR23 to the generic EQ-5D in metastatic breast cancer patients. *Qual Life Res*. 2012;21(7):1193–203.
66. Roche. Pertuzumab in combination with trastuzumab and docetaxel for the treatment of HER2-positive metastatic breast cancer. STA submission: Roche; 2013.
67. National Collaborating Centre for Cancer. Advanced breast cancer: diagnosis and treatment (CG 81). 2009.
68. National Collaborating Centre for Cancer. Early and locally advanced breast cancer: diagnosis and treatment (CG 80): National Institute for Health and Care Excellence; 2009.
69. Lewis R, Bagnall A-M, Forbes C, Shirran E, Duffy S, Kleijnen J, et al. A rapid and systematic review of the clinical effectiveness and cost-effectiveness of trastuzumab for breast cancer: National Institute for Health and Care Excellence; 2001.
70. Huxley N, Jones-Hughes T, Coelho H, Snowsill T, Cooper C, Cooper K, et al. A systematic review and economic evaluation of intraoperative tests [RD-100i one-step nucleic acid amplification (OSNA) system and Metasin test] for detecting sentinel lymph node metastases in breast cancer. *Health Technol Assess*. 2012;19(2):1–215.
71. Ward S, Scope A, Rafia R, Pandor A, Harnan S, Evans P. Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in breast cancer management. *Health Technol Assess*. 2011;17(44):1–302.
72. Tosh JC, Longworth LJ, George E. Utility values in National Institute for Health and Clinical Excellence (NICE) Technology Appraisals. *Value Health*. 2011;14(1):102–9.
73. Doble B, Lorgelly P. Mapping the EORTC QLQ-C30 onto the EQ-5D-3L: assessing the external validity of existing mapping algorithms. *Qual Life Res*. 2016;25(4):891–911.
74. Longworth L, Yang Y, Young T, Mulhern B, Hernandez Alava M, Mukuria C, 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.
75. Pennington B, Davis S. Mapping from the Health Assessment Questionnaire to the EQ-5D: the impact of different algorithms on cost-effectiveness results. *Value Health*. 2014;17(8):762–71.
76. Devlin N, Shah K, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England (OHE research paper 16/01). Office of Health Economics; 2016.
77. Feng Y, Devlin N, Shah K, Mulhern B, van Hout B. New methods for modelling EQ-5D-5L value sets: an application to English data (OHE research paper 16/02). Office of Health Economics; 2016.
78. National Institute for Health and Care Excellence (NICE). NICE In Development Technology Appraisal Guidance List. 2017. Available at: <http://www.nice.org.uk/guidance/indevelopment?type=TA>. Accessed 25 Apr 2017.
79. Jones J, Takeda A, Picot J, von Keyserlingk C, Clegg A. Evidence Review Group Report commissioned by the NHS R&D HTA Programme on behalf of NICE: lapatinib for HER2 over-expressing breast cancer. Southampton: Southampton Health Technology Assessments Centre; 2007.
80. National Institute for Health and Care Excellence (NICE). Breast cancer (brain metastases)—etirinotecan pegol [ID881]. 2017. Available at: <https://www.nice.org.uk/guidance/indevelopment/gid-ta10066>. Accessed 25 Apr 2017.
81. National Institute for Health and Care Excellence (NICE). Fulvestrant for untreated hormone-receptor positive metastatic breast cancer [ID951]. 2017 Available at: <https://www.nice.org.uk/guidance/indevelopment/gid-ta10106>. Accessed 25 Apr 2017.
82. Fleeman N, Bagust A, Beale S, Dwan K, Dickson R, Proudlove C, et al. Pertuzumab in combination with trastuzumab and docetaxel for the treatment of HER2 positive metastatic or locally recurrent unresectable breast cancer: a Single Technology Appraisal. Liverpool: Liverpool Reviews and Implementation Group (LRiG), The University of Liverpool; 2013.
83. Pfizer. Palbociclib in combination with an aromatase inhibitor for previously untreated metastatic, hormone receptor-positive, HER2-negative breast cancer [ID915 CS]. 2017.
84. Fleeman N, Stainthorpe A, Bagust A, Richardson M, Nolan S, Houten R, et al. Palbociclib in combination with an aromatase inhibitor for previously untreated metastatic, hormone receptor-positive, HER2-negative breast cancer [ID915 ERG Report]. 2017.
85. Hughes R, Mitchell CR, Bishop RS, Fotheringham I. Health-state utility values in breast and prostate cancer measured using the EQ-5D: a systematic review of the literature. *Value Health*. 2014;17(7):A645–6.
86. National Institute for Health and Care Excellence (NICE). Guide to the processes of technology appraisal 2014. London: NICE; 2014.
87. Winstanley J, Murray N. National Institute for Health and Care Excellence (NICE). CG81: Advanced breast cancer: diagnosis and treatment. Cardiff: National Collaborating Centre for Cancer. 2009.