


RESEARCH ARTICLE

Cancer Therapy and Prevention

Health utility values of breast cancer treatments and the impact of varying quality of life assumptions on cost-effectiveness

Lindy M. Kregting¹  | Noëlle J. M. C. Vrancken Peeters²  | Marloes E. Clarijs²  |
 Linetta B. Koppert²  | Ida J. Korfage¹  | Nicolien T. van Ravesteyn¹ 

¹Department of Public Health, Erasmus MC, University Medical Center, Rotterdam, The Netherlands

²Academic Breast Cancer Center, Department of Oncologic and Gastro-intestinal Surgery, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Correspondence

Nicolien T. van Ravesteyn, Erasmus MC, Department of Public health, Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands.
 Email: n.vanravesteyn@erasmusmc.nl

Funding information

Rijksinstituut voor Volksgezondheid en Milieu

Abstract

In breast cancer research, utility assumptions are outdated and inconsistent which may affect the results of quality adjusted life year (QALY) calculations and thereby cost-effectiveness analyses (CEAs). Four hundred sixty four female patients with breast cancer treated at Erasmus MC, the Netherlands, completed EQ-5D-5L questionnaires from diagnosis throughout their treatment. Average utilities were calculated stratified by age and treatment. These utilities were applied in CEAs analysing 920 breast cancer screening policies differing in eligible ages and screening interval simulated by the MISCAN-Breast microsimulation model, using a willingness-to-pay threshold of €20,000. The CEAs included varying sets on normative, breast cancer treatment and screening and follow-up utilities. Efficiency frontiers were compared to assess the impact of the utility sets. The calculated average patient utilities were reduced at breast cancer diagnosis and 6 months after surgery and increased toward normative utilities 12 months after surgery. When using normative utility values of 1 in CEAs, QALYs were overestimated compared to using average gender and age-specific values. Only small differences in QALYs gained were seen when varying treatment utilities in CEAs. The CEAs varying screening and follow-up utilities showed only small changes in QALYs gained and the efficiency frontier. Throughout all variations in utility sets, the optimal strategy remained robust; biennial for ages 40–76 years and occasionally biennial 40–74 years. In sum, we recommend to use gender and age stratified normative utilities in CEAs, and patient-based breast cancer utilities stratified by age and treatment or disease stage. Furthermore, despite varying utilities, the optimal screening scenario seems very robust.

KEYWORDS

breast neoplasms, cost-effectiveness analysis, health utility values, mass screening

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *International Journal of Cancer* published by John Wiley & Sons Ltd on behalf of UICC.

What's new?

Utility scores facilitate calculations of quality-adjusted life years (QALYs), which represent a measure of the value of health outcomes. With regard to breast cancer, however, utility scores are inconsistent, with consequences for QALY evaluation and cost-effectiveness analysis (CEA). In our study, new utility values for breast cancer treatment were determined at different time points and stratified by different measures. Evaluation shows that normative utilities stratified by age and gender most reflected real-world situations and were therefore most meaningful for CEA. Likewise, patient-based breast cancer quality-of-life parameters stratified by age and treatment or by disease stage were most appropriate for CEA.

1 | INTRODUCTION

With almost 2.3 million new cases diagnosed annually, breast cancer is the most prevalent cancer worldwide (11.7% of all cancers diagnosed).¹ Fortunately, both developments in breast cancer treatment (eg, adjuvant and neo-adjuvant treatment) and in breast cancer screening programmes have decreased breast cancer mortality.²⁻⁵ In the Netherlands, the 10 year survival rate of invasive breast cancer increased from 40% in 1961-1970 to 80% in 2011-2020.⁶

Researchers and physicians not only aim to reduce breast cancer mortality, they also aim to improve the health related quality of life (HrQoL) of patients with breast cancer. Especially with an increasing population of patients with a good prognosis, also the effects of treatment choices on HrQoL have become more important. Diagnosis and treatment of breast cancer have both physical and psychological effects on patients and impact their HrQoL. To measure the HrQoL in patients, multiple disease specific and generic patient reported outcome measures (PROMs) have been developed. Disease specific PROMs like the European Organisation for Research and Treatment of Cancer (EORTC) QLQ-BR23, and BREAST-Q are used to evaluate the impact of breast cancer specifically from a patients' perspective.⁷⁻⁹ Generic PROMs such as the 36-Item Short Form Survey (SF-36) and the EuroQol EQ-5D are non-disease specific and can be used to compare HrQoL between groups of people, irrespective of their health condition.^{10,11} Some generic PROMs can be used to calculate a utility score which represents the desirability of an individual's health state at a particular point in time and generally ranges between 1 (perfect health) and 0 (equal to death), but can also be negative (worse than death).¹² Subsequently, utility scores per specific health state can be applied in cost-effectiveness analyses to calculate quality adjusted life years (QALYs) of a specific group of people over time.

Utility scores have been reported for a range of health states, diseases and populations. However, utility scores on comparable health states have been found to vary between studies.¹³⁻¹⁷ The differences between these estimates can largely be explained by three issues. First, utility scores as used in studies are based on different HrQoL measures, such as the visual analogue scale (VAS) and the EQ-5D-5L.^{14-16,18} The United Kingdom National Institute for Health and Care Excellence (NICE) recommends the use of the EQ-5D, which has led

to an increase of the use of the EQ-5D over time.¹⁹ Second, studies differed in the population who completed the HrQoL measures on which utility scores were based. A systematic review on oncological cost-effectiveness studies found that only 25% of the studies used utility scores which were based on responses of people who actually experienced the specific health state.¹⁵ In breast cancer specifically, a systematic review found that 11% of the metastatic breast cancer utilities and 43% of the early breast cancer utilities were based on responses of people who actually experienced the specific health states.¹⁶ The other utilities were based on responses of clinicians, researchers, the community, or patients who were asked to value a hypothetical health state. Such utilities have been shown to be lower than patients' own values.¹⁴⁻¹⁶ Third, utility scores were found to differ between countries even when populations, socioeconomic status, health systems, and attitudes toward health were expected to be quite similar.²⁰ It would be most accurate to use utility scores derived from the same population as the target population of a study. Therefore, the EQ-5D has country-specific value sets to translate the answers to the questionnaire into a country-specific utility score.²¹

In cost-effectiveness analyses with QALYs as effectiveness measure, utility values are applied to health states relevant to the disease or intervention studied. Cost-effectiveness analyses concerning cervical cancer, colorectal cancer, and oesophageal adenocarcinoma screening showed that outcomes depended on which utility scores were used.^{18,22} For instance, in a study by de Kok et al, varying HrQoL parameters in health states related to cervical cancer screening and treatment led to variation in the optimal screening strategy with differences in preferred primary test, number of life-time tests and screening interval.¹⁸ This shows that the chosen utility values in cost-effectiveness analyses may influence the results of the analyses and thereby possibly policy recommendations. For breast cancer health states a variety of utility values are available. However, there are uncertainties which values are best to use.^{13,14,17} These uncertainties regard utilities of the normative population, disutilities of breast cancer states (ie, the decrease in utility value due to a specific health state), and disutilities concerning screening and follow-up.

Therefore, the aim of our study is to determine new utility values for various breast cancer treatment options at different time points during treatment stratified by age. In addition, our study aims to quantify the impact of different sets of normative, breast cancer treatment

and screening and follow-up utility values on the cost-effectiveness of a large set of breast cancer screening strategies.

2 | METHODS

Our study consisted of two parts: (1) a longitudinal prospective cohort study on health utility values during breast cancer treatment and (2) a methodological microsimulation study on the effects of varying quality of life assumptions on cost-effectiveness.

2.1 | Breast cancer treatment health utility values

2.1.1 | Data collection

Breast cancer utility scores were calculated using an institutional database including questionnaires responses of 734 women with breast cancer who received axillary treatment for lymph node staging and/or metastasis in the Erasmus MC Cancer Institute in the Netherlands between November 2015 and December 2021.²³ Inclusion criteria were women with a breast cancer diagnosis, who received axillary treatment and completed a set of quality of life related PROMs in the 'patient data platform', the Institute's online PROM collection tool. This tool included the generic HrQoL questionnaire EuroQol EQ-5D-5L.²⁴ The EQ-5D-5L is the updated version of the originally developed EQ-5D-3L and offers five answer options per HRQoL dimension which makes the results more sensitive to small variations in HrQoL than those of the EQ-5D-3L.^{25,26} Exclusion criteria were patients undergoing proton therapy or palliative treatment, or having a history of previous breast cancer. Data of patients were also excluded if clinical data were unavailable or if patients had not completed the EQ-5D-5L. Informed was obtained during the first questionnaire, as part of routine care protocol.^{27,28} Specific information about the cohort was published elsewhere.²³

The database contained self-reported sociodemographic data, clinical treatment data from medical records and EQ-5D-5L data from questionnaires completed before surgery (baseline), 6 months post-surgery and 1-year post-surgery.

2.1.2 | Statistical analyses

The answers to the EQ-5D-5L questionnaires were transformed into a utility score using the Dutch value set.²⁹ Average utility scores were calculated for patients stratified by treatment allocation and age. Treatment was categorised using four characteristics of treatment (1) type of surgery (breast conserving surgery [BCS], mastectomy without reconstruction, mastectomy with reconstruction and no surgery), (2) neoadjuvant and/or adjuvant chemo therapy (Y/N), (3) neoadjuvant and/or adjuvant endocrine therapy (Y/N) and (4) neoadjuvant and/or adjuvant chemo and/or endocrine therapy

(only chemo therapy, only endocrine therapy, both chemo and endocrine therapy, neither chemo nor endocrine therapy). Age was categorised as <35, 35–45, 45–55, 55–65, 65–75 and ≥75 years of age.

2.2 | Effects on cost-effectiveness

2.2.1 | Microsimulation modelling

The microsimulation screening analysis breast (MISCAN-Breast) model was used to simulate individual life histories of a population of women from birth to death and, in a subset of women, the natural history of breast cancer (Figure S1).^{4,30} Breast cancer treatment parameters were updated with data up to 2013, the natural history and breast cancer survival rates were calibrated with data up to 2015.³¹ Additional information about the MISCAN-Breast model can be found in the Supplementary Methods Section in Data S1 and elsewhere.^{4,31,32}

In our study, we simulated a cohort of 10 million women at average risk of developing breast cancer. All women were assumed to be born on 1 January 1980 and life tables, breast cancer parameters and screening parameters were based on data from the Netherlands. Outcomes were calculated for the women from age 40 until death. To calculate the full potential of the screening strategies, attendance rates were set at 100%. In total, 920 breast cancer screening policies were simulated varying in age groups eligible for screening and screening interval. The simulations with the MISCAN-Breast model were performed as in the cost-effectiveness analyses performed previously; additional details on parameters and assumptions can be found in the original article.³²

The results from the simulations were used to calculate the number of QALYs gained and additional costs compared to no screening per screening policy. Furthermore, incremental cost effectiveness ratio's (ICERs) were calculated to form efficiency frontiers.

2.2.2 | Quality of life parameters

To evaluate the effect of different utility sets on cost-effectiveness, different sets of (1) normative, (2) breast cancer treatment and (3) screening and follow-up utility values were used to calculate QALYs gained and ICERs.

Normative utility values

First, cost-effectiveness analyses were performed using different utility sets for the normative health state (ie, the average health state of a comparative population without the disease of interest). Normative utility values which were applied were perfect health (utility score of 1), gender specific (0.858)²⁹ and gender and age specific³³ (Table 1). Breast cancer treatment and screening and follow-up utility values were used as in the study by Kregting et al.³²

TABLE 1 Quality of life parameters applied to health states in the different cost-effectiveness analyses.

Utility category	Utility set name	Utility and disutility values		
Normative	Perfect health	1		
	Gender specific (Versteegh et al ²⁹)	0.858		
	Age and gender specific (Clarijs et al ³³)	Age	Utility	
		<35	0.93	
		35–44	0.93	
		45–54	0.91	
		55–64	0.91	
65–74	0.89			
≥75	0.88			
Breast cancer treatment	Stratified by disease stage ^{32,34}	Stage at diagnosis	Disutility factor on normative utility	
		DCIS or localised cancer	0.1 for 2 years	
		Regional or distant cancer	0.25 for 2 years	
		Breast cancer leading to death	0.4 for 2 years	
	Stratified by age and type of surgery	Treatment utilities as reported in Table 3 for a duration of 2 years. Breast cancer leading to death: factor on normative utility of 0.4 for years 3 to terminal year, and factor of 0.7 for terminal year.		
Stratified by age and use of chemo therapy	Treatment utilities as reported in Table 3 for a duration of 2 years. Breast cancer leading to death: factor on normative utility of 0.4 for years 3 to terminal year, and factor of 0.7 for terminal year.			
Stratified by age and use of endocrine therapy	Treatment utilities as reported in Table 3 for a duration of 2 years. Breast cancer leading to death: factor on normative utility of 0.4 for years 3 to terminal year, and factor of 0.7 for terminal year.			
Stratified by age and use of chemo and/or endocrine therapy	Treatment utilities as reported in Table 3 for a duration of 2 years. Breast cancer leading to death: factor on normative utility of 0.4 for years 3 to terminal year, and factor of 0.7 for terminal year.			
Screening and follow-up	De Haes et al ³⁷	Screening participation	0.006 for 1 week	
		Follow-up testing	0.105 for 5 weeks	
	Disutilities double	Screening participation	0.012 for 1 week	
		Follow-up testing	0.210 for 5 weeks	
	Disutilities half	Screening participation	0.003 for 1 week	
		Follow-up testing	0.0525 for 5 weeks	
	Disutility screening participation doubles	Screening participation	0.012 for 1 week	
		Follow-up testing	0.105 for 5 weeks	
	Disutility screening participation halves	Screening participation	0.003 for 1 week	
		Follow-up testing	0.105 for 5 weeks	
	Disutility follow-up testing doubles	Screening participation	0.006 for 1 week	
		Follow-up testing	0.210 for 5 weeks	
	Disutility follow-up testing halves	Screening participation	0.006 for 1 week	
Follow-up testing		0.0525 for 5 weeks		
Time for follow-up testing disutility shortens	Screening participation	0.006 for 1 week		
	Follow-up testing	0.105 for 3 weeks		
No disutility for screening participation ⁴⁰	Screening participation	0.000		
	Follow-up testing	0.105 for 5 weeks		

Breast cancer treatment utility values

Second, cost-effectiveness analyses were performed using different utility sets for breast cancer treatment health states. These utility values were (1) equal to the values from Kregting et al³² which were based on Stout et al,³⁴ and (2–5) based on the values from part 1 of the current study (Table 1). Utility scores at diagnosis were used in the analyses for a duration of 1 month, treatment scores for the following 11 months and recovery scores for the subsequent 1 year.

Furthermore, screening and follow-up utility values were similar to Kregting et al and normative utility values were age and gender specific.^{32,33}

Screening and follow-up utility values

Lastly, cost-effectiveness analyses were performed using different utility sets for breast cancer screening and follow-up health states. These utility values were varied based upon the values by de Haes et al

(Table 1). Therefore, the analyses of these utility sets can be seen as a sensitivity analysis. Furthermore, breast cancer treatment utility values were similar to Kregting et al and normative utility values were age and gender specific.^{32,33}

For some treatment categories, no patient data were available to calculate utility values for. This was the case for some treatment options in the age group 75 years and older. Therefore, the assumption was made that the effect of the treatment on utility values for women over 75 years of age would be the same as for women aged between 64 and 75 years receiving the same treatment with a correction based on the difference in normative utility score between the two age groups (factor 0.99). Furthermore, there were no utility data for patients with breast cancer who had no surgery. Therefore, the assumption was made that women without surgery under the age of 75 years had a poor prognosis, possibly due to metastasis or comorbidity, and therefore a poor quality of life with an assumed disutility factor of 30% compared to normative utilities. Women aged 75 years and older who did not get breast cancer surgery were assumed to have latent tumours which would probably not lead to breast cancer death. Therefore, the disutility factor for this group was assumed to be 15% compared to normative utility values.

Data from the Netherlands Cancer Registry (NKR) and Netherlands Comprehensive Cancer Organisation (IKNL) on the use of breast cancer treatment options in the total population of patients with breast cancer diagnosed in 2017 in the Netherlands were used to determine treatment usage per age group and stratified by mode of detection (Tables S2–S5). The modes of detection were detected in screening, interval cancer (maximally 30 months after a screening examination) or clinically detected cancers (in women who did not attend screening in at least 30 months).

2.2.3 | Cost-effectiveness analyses

A healthcare payer perspective was adopted and direct medical costs were calculated, including costs of screening, diagnostics and treatment. Cost parameters were similar to the analyses of Kregting et al in which they were largely based on a study by Geuzinge et al (Table S1 and Supplementary Methods in Data S1).^{32,35} Per utility category, multiple cost-effectiveness analyses were performed differing in utility scores only. All other parameters remained the same and are equal to the cost-effectiveness analyses previously performed by Kregting et al. This includes discounting at 3.5% per year for QALYs and costs from 2020. Subsequently, ICERs were calculated by dividing the difference in costs by the difference in QALYs between screening strategies. Therefore, the ICER reflects the costs required to gain one QALY compared to the previous strategy. ICERs were not calculated for strategies that were dominated by another strategy (ie, another strategy gained more QALYs and required less costs). Per utility score set, an efficiency frontier was drawn with all strategies which were not dominated and therefore had an ICER. The ICERs were compared to a conservative willingness to pay (WTP) threshold of €20,000 per QALY gained.³⁶ Strategies that did not exceed this threshold were considered to be cost-effective. Per utility

category, the efficiency frontier were compared between the different utility score sets. Moreover, the optimal strategies, according to the WTP threshold, for each utility set were compared.

3 | RESULTS

In total, 734 patients were identified in the institutes' online patient database until 31 December 2021 (Figure 1). After removal of duplicates, 731 patients remained. Of those patients, 27 non-responders (<5%) were excluded and 152 patients were excluded based on the exclusion criteria, see Figure 1. For each of the three time points, patients were excluded if no EQ-5D-5L scores were available leaving 464 patients at baseline, 442 patients 6 months after surgery and 456 patients 12 months after surgery.

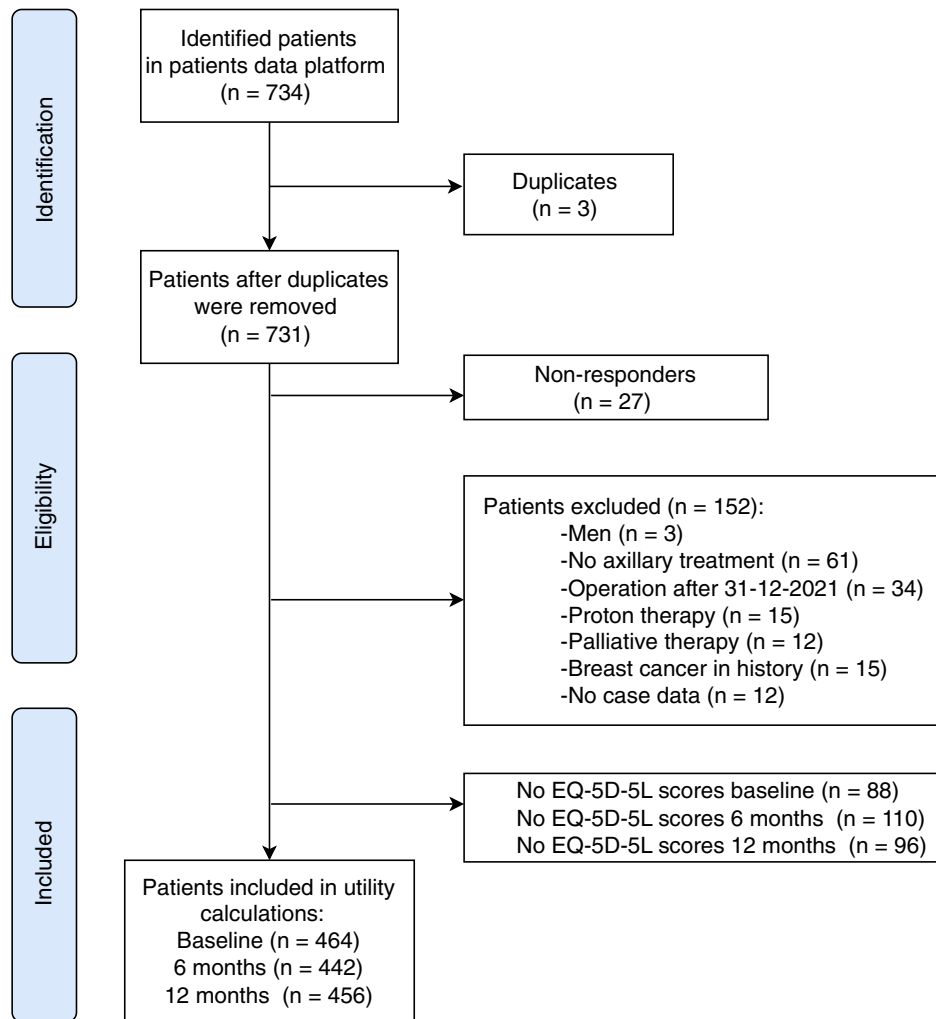
Among the 464 patients included at baseline, the mean age was 54 years (standard deviation [SD] 14.2) (Table 2). More than half of the patients had T stage 1 and negative lymph nodes. When looking at the treatment characteristics, 53.2% underwent BCS, 40.7% received chemotherapy and 52.6% received endocrine therapy.

The Erasmus MC dataset showed differences in utility scores between age groups (Table 3). Shortly after diagnosis, utility values were in general a little lower than normative values. During treatment, utility scores decreased in most categories. During recovery, utility scores in most categories increased generally approaching the level at diagnosis. However, differences in this pattern of decreasing and increasing utility scores were seen between treatment options and between age groups.

Results of the cost-effectiveness analyses showed that variation in normative utilities had an effect on the number of QALYs gained per modelled screening strategy (Figure 2A). The 'perfect health' utility set resulted in a higher amount of QALYs gained compared to the other two sets. Only small differences in QALYs gained were observed between the 'gender specific' and 'gender and age specific' utility sets where the first one was slightly lower. Compared between all three utility sets, the efficiency frontiers were very similar in regard to which strategies were on the frontier. Taking all considerations into account, the optimal strategy was biennial screening for ages 40–76 for all three utility sets (Table S6).

The analyses varying breast cancer treatment utilities showed small differences in QALYs gained per strategy (Figure 2B). For the triennial and quadrennial strategies, the efficiency frontier overlapped largely (Table S7). However, for the biennial strategies the differences in QALYs gained were a little larger, and these differences were even larger for annual strategies. Also here, the strategies on the efficiency frontier were similar between utility sets used. Furthermore, the utility sets 'stratified by disease stage', 'stratified by age and chemotherapy' and 'stratified by age and chemo and/or endocrine therapy' led to an optimal strategy of biennial screening for ages 40–76, whereas the sets 'stratified by age and type of surgery' and 'stratified by age and endocrine therapy' led to biennial screening for ages 40–74 to be the optimal strategy.

The sensitivity analyses on the screening and follow-up utility values showed that the variation in these values did not substantially



change the number of QALYs gained (Figure 2C). For the biennial, triennial and quadrennial strategies, the efficiency frontiers mostly overlapped (Table S8). For the annual strategies, the differences were somewhat larger. Again, the strategies on the efficiency frontier were rather similar between utility sets and the optimal strategy was biennial screening for the ages 40–76 for seven out of the nine utility sets (Table S1). Only utility sets ‘disutilities double’ and ‘disutility follow-up testing doubles’ resulted in a different optimal strategy; biennial screening for the ages 40–74.

4 | DISCUSSION

In our study, average utility values were calculated for Dutch female patients with breast cancer over time stratified by age and treatment. Furthermore, it showed the impact of using varying health utility values for normative, breast cancer treatment and screening and follow-up health states. This showed that the use of different normative utility sets led to differences in the amount of QALYs gained in cost-effectiveness analyses. On the other hand, variances in breast cancer treatment, and screening and follow-up utility sets only had a

limited effect on the number of QALYs gained. Moreover, the optimal screening strategies resulting from the cost-effectiveness analyses were rather robust at biennial for ages 40 to 74 despite the variation in utility sets.

In the cost-effectiveness analyses with varying normative utility sets, the ‘gender and age specific’ set can be considered as closest to the real world situation. Therefore, this set can be seen as the best option to use in cost-effectiveness analyses. When comparing the other utility sets to this, the ‘perfect health’ set leads to an overestimation of QALYs gained, while the ‘gender specific’ set leads to a small underestimation. In the cost-effectiveness analyses with varying breast cancer treatment utility sets, the set ‘stratified by disease stage’ was based on assumptions by Stout et al, whereas the sets ‘stratified by age and type of surgery’, ‘stratified by age and use of chemo therapy’, ‘stratified by age and endocrine therapy’ and ‘stratified by age and chemo and/or endocrine therapy’ were all based on the same dataset from the Erasmus MC Cancer Institute. Because the latter four were based on real-world data, these may be favoured over the first one. However, since results were rather similar, the choice of utility sets did not seem to have a large impact. Finally, the varying screening and follow-up utility sets were hypothetical values in order

TABLE 2 Population characteristics at baseline.

Population at baseline (n)	464
Mean age (SD)	54.0 y (14.2 y)
Age categorical (n, %)	
<35	56 (12.1%)
35–44	77 (16.6%)
45–54	103 (22.2%)
55–64	111 (23.9%)
65–74	91 (19.6%)
≥75	26 (5.6%)
T stage (n, %)	
pT0	42 (9.1%)
pT1	260 (56.0%)
pT2	90 (19.4%)
pT3	14 (3.0%)
pT4	1 (0.2%)
pTis	50 (10.8%)
pTmi	7 (1.5%)
N stage (n, %)	
pN0	354 (76.3%)
pN1	84 (18.1%)
pN2	19 (4.1%)
pN3	7 (1.5%)
Surgery (n, %)	
BCS ^a	247 (53.2%)
Mastectomy	129 (27.8%)
Reconstruction	88 (19.0%)
Chemo therapy (n, %)	
No	275 (59.3%)
Yes	189 (40.7%)
Endocrine therapy (n, %)	
No	220 (47.4%)
Yes	244 (52.6%)

^aBCS; breast conserving surgery.

to perform a sensitivity analyses. Therefore, no preferred set can be appointed. However, since the amount of QALYs gained, the strategies on the efficiency frontier, and the optimal strategy were very similar for all utility sets, it can be concluded that the screening and follow-up utility values as found by de Haes et al give robust results.³⁷

The effect of differing quality of life parameters was most substantial for normative utility values because of the large number of simulated person-years that were experienced in normative health. In contrast, only a proportion of individuals was simulated to develop breast cancer with a duration of a limited period of time, resulting in fewer person-years with breast cancer than in normative health. Therefore, the variations in normative utility values had the largest effect on total number of QALYs gained. Given that the normative

utility values from Clarijs et al were retrieved from a large population of women in the Netherlands and were stratified by both gender and age, these parameters can be considered best estimates.³³ Because the QALY estimates in the analyses using only gender specific utility values were quite similar to QALY estimates using gender and age specific utility values, we conclude that? gender specific utility values can be used in the absence of gender and age stratified normative utility values. Also of importance is the use of country specific utility parameters, especially for normative utilities which should preferably be calculated using the same measure and the same value set as the data on the disease or intervention of interest.^{38,39} A smaller effect on the number of QALYs gained was seen when varying breast cancer treatment quality of life parameters. This was expected for the utility sets which were all based on the same dataset and stratified differently on treatment. However, a bigger difference was expected for the 'Kregting et al' utility set which used rather different utility parameters based on stage at diagnosis. The analyses showed that disutilities attributed to the health state 'breast cancer leading to death' had the biggest effect on the number of QALYs gained. Furthermore, sensitivity analyses of the screening and follow-up utility parameters showed that the results of the cost-effectiveness analyses are robust. Although the estimates of de Haes et al have been made more than 30 years ago and are debated sometimes, they were shown to lead to robust conclusions when varied in sensitivity analyses from half up to double the value estimates resulting in robust QALY results on population level.³⁷ Moreover, all cost-effectiveness analyses with variations in utility sets resulted in a very similar list of screening strategies on the efficiency frontier and two quite similar optimal screening strategies of biennial screening for ages 40–74 or ages 40–76 (when using a WTP threshold of €20,000 per QALY gained). This showed that the benefit-harm balance of these strategies was robustly advantageous over other investigated screening strategies.

Earle et al reported utilities for patients with breast cancer undergoing BCS, chemotherapy or endocrine therapy to be between 0.97 and 1.0 compared to 1.0 for a normative population.¹⁵ In the current study, the utility values for these categories are much lower ranging between 0.71 (age 75+, chemotherapy yes, during recovery) and 0.91 (age 45–54, chemotherapy yes endocrine therapy no, at diagnosis). When taking into consideration the gender and age specific normative utilities, the differences become smaller, but remain present. An explanation for these differences can be found in the timing of the valuation compared to the treatment, the quality of life instrument used and the population asked to value the health states.¹⁶

A few studies investigated the effects of varying quality of life parameters in cost effectiveness analyses. A study on colorectal and oesophageal cancer screening strategies found that varying utility parameters could substantially impact the number of QALYs gained.²² However, this did not seem to impact which strategies were on the efficiency frontier, which is comparable to the finding of the current study in breast cancer screening strategies. On the contrary, a study by de Kok et al found different optimal strategies for cervical cancer screening when using different utility sets.¹⁸ Our study even found different screening modalities to be optimal depending on the set of

TABLE 3 Breast cancer treatment utility values.

Utility set name	Subcategories	Treatment stage	Age					
			<35	35–44	45–54	55–64	65–74	≥75
Type of surgery	BCS	Diagnosis	0.82	0.83	0.87	0.85	0.88	0.80
		Treatment	0.77	0.78	0.81	0.78	0.88	0.83
		Recovery	0.80	0.80	0.84	0.81	0.85	0.82
	Mastectomy	Diagnosis	0.79	0.82	0.82	0.77	0.80	0.81
		Treatment	0.69	0.78	0.71	0.77	0.77	0.85
		Recovery	0.79	0.80	0.80	0.78	0.73	0.83
	Mastectomy + reconstruction	Diagnosis	0.78	0.87	0.90	0.89	0.81	0.80 ^a
		Treatment	0.76	0.86	0.77	0.81	0.91	0.91 ^a
		Recovery	0.81	0.86	0.84	0.88	0.89	0.89 ^a
	No surgery ^b	Diagnosis	0.65 ^b	0.65	0.64 ^b	0.64 ^b	0.62 ^b	0.75 ^b
		Treatment	0.65 ^b	0.65 ^b	0.64 ^b	0.64 ^b	0.62 ^b	0.75 ^b
		Recovery	0.65 ^b	0.65 ^b	0.64 ^b	0.64 ^b	0.62 ^b	0.75 ^b
Chemotherapy	Chemotherapy yes	Diagnosis	0.80	0.83	0.87	0.83	0.76	0.75 ^a
		Treatment	0.74	0.79	0.77	0.78	0.74	0.73 ^a
		Recovery	0.80	0.80	0.82	0.80	0.71	0.71 ^a
	Chemotherapy no	Diagnosis	0.79	0.86	0.86	0.83	0.86	0.81
		Treatment	0.81	0.83	0.78	0.78	0.86	0.84
		Recovery	0.87	0.84	0.83	0.81	0.84	0.83
Endocrine therapy	Endocrine therapy yes	Diagnosis	0.81	0.85	0.84	0.85	0.86	0.78
		Treatment	0.75	0.81	0.78	0.77	0.91	0.84
		Recovery	0.77	0.83	0.83	0.80	0.89	0.78
	Endocrine therapy no	Diagnosis	0.79	0.83	0.90	0.81	0.84	0.83
		Treatment	0.74	0.79	0.77	0.79	0.78	0.84
		Recovery	0.83	0.80	0.83	0.83	0.76	0.90
Chemo and/or endocrine therapy	Chemotherapy yes, endocrine therapy no	Diagnosis	0.78	0.78	0.91	0.81	0.74	0.73 ^a
		Treatment	0.74	0.73	0.73	0.80	0.62	0.61 ^a
		Recovery	0.83	0.78	0.81	0.77	0.60	0.59 ^a
	Chemotherapy no, endocrine therapy yes	Diagnosis	0.67	0.85	0.84	0.86	0.87	0.78
		Treatment	0.78	0.79	0.77	0.77	0.91	0.84
		Recovery	0.89	0.87	0.82	0.78	0.89	0.78
	Chemotherapy yes, endocrine therapy yes	Diagnosis	0.83	0.85	0.85	0.84	0.79	0.78 ^a
		Treatment	0.75	0.82	0.79	0.77	0.89	0.88 ^a
		Recovery	0.76	0.82	0.83	0.81	0.86	0.85 ^a
	Chemotherapy no, endocrine therapy no	Diagnosis	0.91	0.87	0.88	0.81	0.86	0.83
		Treatment	0.87	0.85	0.79	0.79	0.80	0.84
		Recovery	0.85	0.83	0.85	0.84	0.78	0.90

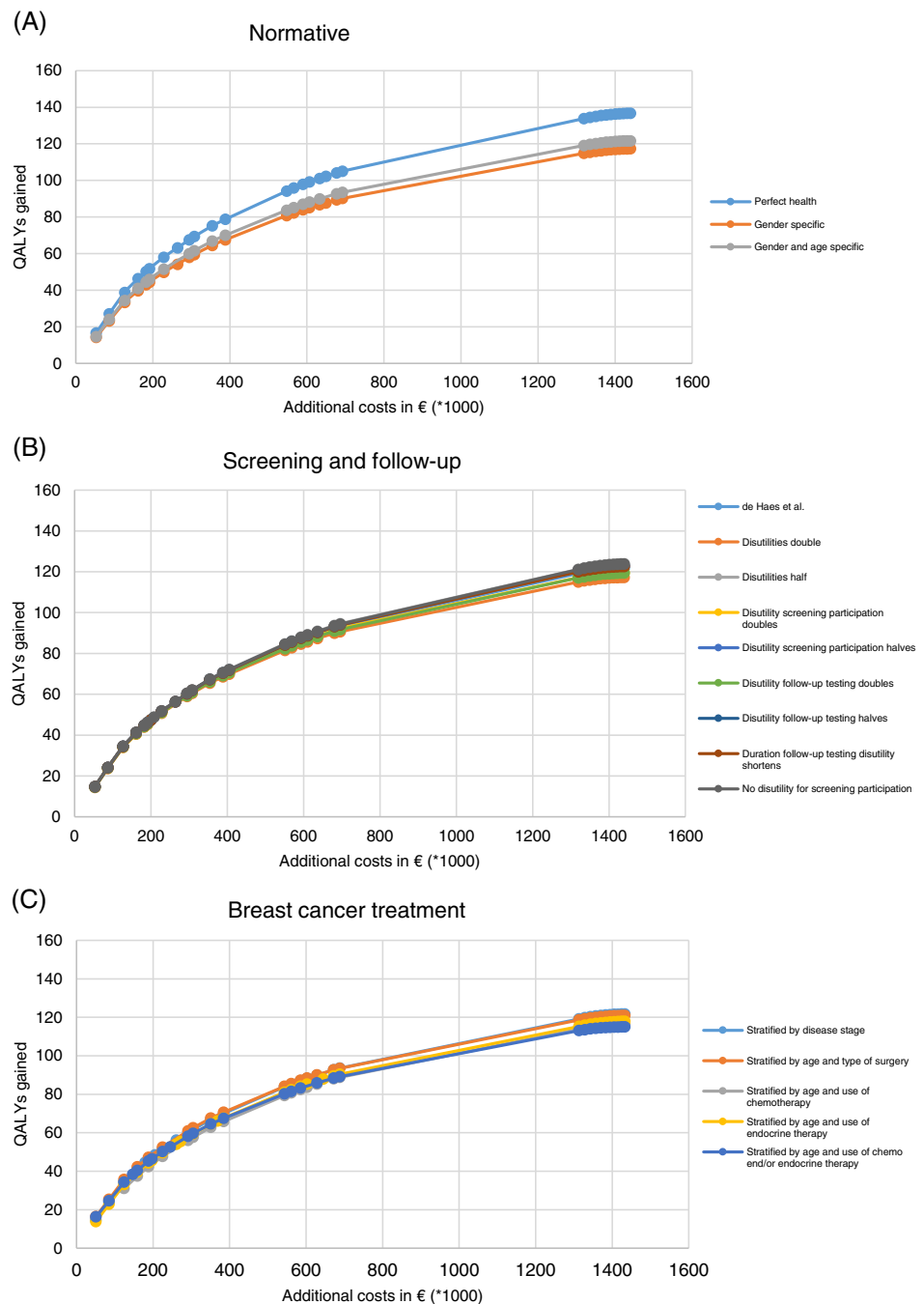
^aSome categories did not have any patient in the age group ≥75. Therefore utility values were calculated based on the utility in this treatment category and stage at age 64–75 and the normative value for age group ≥75.

^bThe Erasmus MC dataset consisted of patients who all received surgery as part of their treatment. Therefore, no data was available on patient who did not receive surgery. However, in the total population of Dutch patients with breast cancer, there were patients who did not receive surgery, therefore assumption on utility values based on normative utilities were necessary.

utility parameters used. A possible explanation for the difference between the studies could be that de Kok et al used a wide variety of utility sets used in published cost-effectiveness analyses, whereas the current study used a more restricted number of utility sets.

Furthermore, de Kok et al also varied the duration of disutilities which was only the case for one utility set of screening and follow-up utility parameters in the current study. It is expected that effects on QALYs will be larger in case durations of disutilities will be varied as well

FIGURE 2 Efficiency frontiers of screening strategies for each utility set for (A) normative, (B) breast cancer treatment and (C) screening and follow-up health states.



as the size of the disutilities. In addition, the patient treatment utilities in the current study were largely based on the same set of data which caused them to be rather similar. Moreover, compared to the studies on colorectal, oesophageal and cervical cancer screening, the current study modelled many more screening strategies in the cost-effectiveness analyses. Despite the higher number of strategies, the strategies on the efficiency frontier and the optimal strategy based on the WTP threshold were more robust.

A major strength of the current study was that both normative and breast cancer utility parameters in the cost-effectiveness analyses were based on prospective, longitudinal real-world data collected

using PROMs. Moreover, the use of the validated EQ-5D-5L in combination with the Dutch value set caused the utility values to be of high quality and representative for the Dutch population.²⁹ The utility values for patients with breast cancer were calculated using prospective cohort data from the population who actually experienced the specific health states studied. Therefore, the utility values were more representative of the actual health state patients experience than in studies where doctors, nurses or the general population valued health states that they never experienced themselves. A limitation of the dataset used was that all patients were treated in Erasmus MC, which is an academic hospital. Possibly, this could have influenced the representativeness of this patient population compared to the all patients

with breast cancer in the Netherlands. In addition, due to the inclusion criteria of the dataset, only patients who received surgery were included. Therefore, no quality of life data were available on patients with breast cancer who did not receive surgery and assumptions had to be made on utility parameters for this group. Also in literature, utility parameters for this group were hard to find, therefore the assumptions made are less reliable. Nevertheless, it is a strength of our study that this population was taken into consideration because data from IKNL showed that there is a substantial group of patients with breast cancer who did not receive surgery, especially the patients who did not participate in screening. Another limitation of the dataset was that, despite the high number of patients, some categories included only small numbers of patients or none at all after stratification by age and treatment type. These could all be explained by the fact that certain treatments are very uncommon in certain age groups (eg, chemotherapy in patients over the age of 75). To provide for utility parameters in the cost-effectiveness analyses, the utility values from the closest age group and a factor based on normative differences between the age groups were used.

In conclusion, our study provided new data-based utility values for patients with breast cancer stratified by age and treatment options which can be used in cost-effectiveness analyses. Furthermore, it showed that the use of gender and age stratified normative utilities and patient-based breast cancer quality of life parameters stratified by age and treatment, or disease stage are recommended. In addition, number of QALYs gained was not sensitive to variations in screening and follow-up utilities and efficiency frontiers and optimal screening strategies were found to be very robust.

AUTHOR CONTRIBUTIONS

Lindy M. Kregting: Conceptualisation; data curation; formal analysis; investigation; methodology; validation; visualisation; writing – original draft. **Noëlle J. M. C. Vrancken Peeters:** Conceptualisation; data curation; formal analysis; investigation; methodology; validation; visualization; writing – review and editing. **Marloes E. Clarijs:** Conceptualisation; data curation; methodology; writing – review and editing. **Linetta B. Koppert:** Conceptualisation; data curation; methodology; funding acquisition; supervision; writing – review and editing. **Ida J. Korfage:** Conceptualisation; methodology; supervision; validation; writing – review and editing. **Nicolien T. van Ravesteyn:** Conceptualisation; data curation; methodology; funding acquisition; supervision; validation; writing – review and editing. The work reported in the article has been performed by the authors, unless clearly specified in the text.

ACKNOWLEDGEMENTS

The authors thank the registration team of the Netherlands Comprehensive Cancer Organisation (IKNL) for the collection of data for the Netherlands Cancer Registry (NKR).

FUNDING INFORMATION

This research was funded by the Centre for Population Screening (CvB) of the Dutch National Institute for Public Health and the Environment (RIVM). The funding source had no involvement in the study

design, modelling decisions, interpretation of the results, writing of the report or in the decision to submit the article for publication.

CONFLICT OF INTEREST STATEMENT

Dr. van Ravesteyn reports receiving fees for consulting from Wickenstones (paid to institution). The other authors have no conflicts to declare.

DATA AVAILABILITY STATEMENT

Utility data and detailed modelling results are available from the corresponding author upon reasonable request. Data used as input for the MISCAN-Breast model can be requested from the primary source.

ETHICS STATEMENT

The Erasmus MC Research Ethics Committee has confirmed that no ethical approval is required. Informed consent from participants for storing and using information for research purposes was obtained during the first questionnaire, as part of routine care protocol.

ORCID

Lindy M. Kregting  <https://orcid.org/0000-0002-8262-0080>

Noëlle J. M. C. Vrancken Peeters  <https://orcid.org/0009-0007-3558-3946>

Marloes E. Clarijs  <https://orcid.org/0000-0001-7118-3324>

Ida J. Korfage  <https://orcid.org/0000-0001-6538-9115>

Nicolien T. van Ravesteyn  <https://orcid.org/0000-0003-3531-466X>

REFERENCES

- International Agency for Research on Cancer. *Breast Fact Sheet*. World Health Organisation; 2020.
- Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med*. 2005; 353:1784-1792.
- Vilapriño E, Puig T, Rue M. Contribution of early detection and adjuvant treatments to breast cancer mortality reduction in Catalonia, Spain. *PLoS One*. 2012;7:e30157.
- de Gelder R, Heijnsdijk EA, Fracheboud J, Draisma G, de Koning HJ. The effects of population-based mammography screening starting between age 40 and 50 in the presence of adjuvant systemic therapy. *Int J Cancer*. 2015;137:165-172.
- Sankatsing VDV, van Ravesteyn NT, Heijnsdijk EAM, et al. The effect of population-based mammography screening in Dutch municipalities on breast cancer mortality: 20 years of follow-up. *Int J Cancer*. 2017; 141:671-677.
- Nederlandse Kankerregistratie (NKR), Netherlands Comprehensive Cancer Organisation (IKNL). Overleving, Invasief mammacarcinoom, periode van diagnose (10 jaar). 2022.
- Pusic AL, Klassen AF, Scott AM, Klok JA, Cordeiro PG, Cano SJ. Development of a new patient-reported outcome measure for breast surgery: the BREAST-Q. *Plast Reconstr Surg*. 2009;124:345-353.
- Sprangers MA, Groenvold M, Arraras JI, et al. The European Organization for Research and Treatment of cancer breast cancer-specific quality-of-life questionnaire module: first results from a three-country field study. *J Clin Oncol*. 1996;14:2756-2768.
- Salas M, Mordin M, Castro C, Islam Z, Tu N, Hackshaw MD. Health-related quality of life in women with breast cancer: a review of measures. *BMC Cancer*. 2022;22:66.

10. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30:473-483.
11. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol group. *Ann Med*. 2001;33:337-343.
12. Torrance GW. Utility approach to measuring health-related quality of life. *J Chronic Dis*. 1987;40:593-603.
13. Wang Y, Gavan SP, Steinke D, Cheung K-L, Chen L-C. The impact of age on health utility values for older women with early-stage breast cancer: a systematic review and meta-regression. *Health Qual Life Outcomes*. 2022;20:169.
14. Kaur MN, Yan J, Klassen AF, et al. A systematic literature review of health utility values in breast cancer. *Med Decis Making*. 2022;42:704-719.
15. Earle CC, Chapman RH, Baker CS, et al. Systematic overview of cost-utility assessments in oncology. *J Clin Oncol*. 2000;18:3302-3317.
16. Peasgood T, Ward SE, Brazier J. Health-state utility values in breast cancer. *Expert Rev Pharmacoecon Outcomes Res*. 2010;10:553-566.
17. Paracha N, Thuresson PO, Moreno SG, MacGilchrist KS. Health state utility values in locally advanced and metastatic breast cancer by treatment line: a systematic review. *Expert Rev Pharmacoecon Outcomes Res*. 2016;16:549-559.
18. de Kok I, Korfage IJ, van den Hout WB, et al. Quality of life assumptions determine which cervical cancer screening strategies are cost-effective. *Int J Cancer*. 2018;142:2383-2393.
19. National Institute for Health and Care Excellence UK. *Guide to the Methods of Technology Appraisal 2013*. National Institute for Health and Care Excellence UK; 2013.
20. Janssen B, Szende A. *Population Norms for the EQ-5D Self-Reported Population Health; An International Perspective Based on EQ-5Ded*. Springer; 2014.
21. EQ-5D Valuation: Standard Value Sets, vol. 2022: EuroQo.
22. Omidvari AH, Lansdorp-Vogelaar I, de Koning HJ, Meester RGS. Impact of assumptions on future costs, disutility and mortality in cost-effectiveness analysis; a model exploration. *PLoS One*. 2021;16:e0253893.
23. Vrancken Peeters NJMC, Kaplan ZLR, Clarijs ME, et al. Health-related quality of life (HRQoL) after different axillary treatments in women with breast cancer: a 1-year longitudinal cohort study. *Qual Life Res*. 2024;33:467-479.
24. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20:1727-1736.
25. Buchholz I, Janssen MF, Kohlmann T, Feng YS. A systematic review of studies comparing the measurement properties of the three-level and five-level versions of the EQ-5D. *Pharmacoeconomics*. 2018;36:645-661.
26. Janssen MF, Bonsel GJ, Luo N. Is EQ-5D-5L better than EQ-5D-3L? A head-to-head comparison of descriptive systems and value sets from seven countries. *Pharmacoeconomics*. 2018;36:675-697.
27. van Egdom LSE, Lagendijk M, van der Kemp MH, et al. Implementation of value based breast cancer care. *Eur J Surg Oncol*. 2019;45:1163-1170.
28. Huberts AS, Clarijs ME, Pastoor H, van Rosmalen M, Koppert LB. Sexual well-being in patients with early-stage breast cancer at 1- and 2-year follow-up. *J Sex Med*. 2023;20:507-514.
29. Versteegh MM, Vermeulen KM, Evers SMAA, de Wit GA, Prenger R, Stolk EA. Dutch tariff for the five-level version of EQ-5D. *Value Health*. 2016;19:343-352.
30. de Gelder R, Bulliard JL, de Wolf C, et al. Cost-effectiveness of opportunistic versus organised mammography screening in Switzerland. *Eur J Cancer*. 2009;45:127-138.
31. Sankatsing VDV, van Ravesteyn NT, Heijnsdijk EAM, Broeders MJM, de Koning HJ. Risk stratification in breast cancer screening: cost-effectiveness and harm-benefit ratios for low-risk and high-risk women. *Int J Cancer*. 2020;147:3059-3067.
32. Kregting LM, Sankatsing VDV, Heijnsdijk EAM, de Koning HJ, van Ravesteyn NT. Finding the optimal mammography screening strategy: a cost-effectiveness analysis of 920 modelled strategies. *Int J Cancer*. 2022;151:287-296.
33. Clarijs ME, Kregting LM, van Ravesteyn NT, Koppert LB, Korfage IJ. Health-related quality of life using the EQ-5D-5L: normative utility scores in a Dutch female population. *Qual Life Res*. 2023;32:373-381.
34. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. *J Natl Cancer Inst*. 2006;98:774-782.
35. Geuzinge HA, Bakker MF, Heijnsdijk EAM, et al. Cost-effectiveness of magnetic resonance imaging screening for women with extremely dense breast tissue. *J Natl Cancer Inst*. 2021;113:1476-1483.
36. Zwaap J, Knies S, van der Meijden C, Staal P, van der Heiden L. Cost-effectiveness in practice. *Zorginstituut*. 2015.
37. de Haes JC, de Koning HJ, van Oortmarssen GJ, van Agt HM, de Bruyn AE, van Der Maas PJ. The impact of a breast cancer screening programme on quality-adjusted life-years. *Int J Cancer*. 1991;49:538-544.
38. Lamers LM, McDonnell J, Stalmeier PF, Krabbe PF, Busschbach JJ. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. *Health Econ*. 2006;15:1121-1132.
39. Norman R, Cronin P, Viney R, King M, Street D, Ratcliffe J. International comparisons in valuing EQ-5D health states: a review and analysis. *Value Health*. 2009;12:1194-1200.
40. Grann VR, Patel P, Bharthuar A, et al. Breast cancer-related preferences among women with and without BRCA mutations. *Breast Cancer Res Treat*. 2010;119:177-184.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kregting LM, Vrancken Peeters NJMC, Clarijs ME, Koppert LB, Korfage IJ, van Ravesteyn NT. Health utility values of breast cancer treatments and the impact of varying quality of life assumptions on cost-effectiveness. *Int J Cancer*. 2024;1-11. doi:10.1002/ijc.34899