Breast Cancer Res Treat (2012) 131:627–636
DOI 10.1007/s10549-011-1769-7

EPIDEMIOLOGY

Head-to-head comparison of the 70-gene signature versus the 21-gene assay: cost-effectiveness and the effect of compliance

Valesca P. Retèl · Manuela A. Joore · Wim H. van Harten

Received: 30 August 2011 / Accepted: 2 September 2011 / Published online: 24 September 2011

Abstract Both the 70-gene signature and the 21-gene assay are novel prognostic tests used to guide adjuvant chemotherapy decisions in patients with early breast cancer. Although the results of ongoing prospective trials will only become available in some years, the tests have already been included in clinical guidelines such as St. Gallen's. In literature, the cost-effectiveness (CE) of both tests as compared to conventional prognostic tests has been described. We report on a direct comparison of CE; as different compliance rates were reported, we also taken these into account. A Markov decision model with a time horizon of 20 years was developed to assess the effects, costs and CE of three alternatives; 21-gene, 70-gene, and St. Gallen (SG) or Adjuvant Online (AO), dependent on the dataset used in patients with early, node-negative, breast cancer. Sensitivity and specificity were based on two datasets, incorporating compliances rates based on literature. For both datasets, whereas the 70-gene signature yielded more quality adjusted life years (QALYs) and was less costly; the 21-gene amounted more life years (LYs) but was more costly. The decision uncertainty surrounding the probability of CE of the Thomassen-series amounted 55% for both cost/LY and cost/QALY, for the Fan-series 80% for LY and 65% for QALYs. Taking reported compliance with discordant test results into account, in general, the effect of all strategies decreased, while the costs increased, without relatively influencing the CEA performance. This comparison indicates that the performances of the 70-gene and the 21-gene based on reported studies are close. The 21-gene has the highest probability of being cost-effective when focusing on cost/LY, while focusing on cost/QALY, the 70-gene signature was most cost-effective. The level of compliance can have serious impact on the CE. With additional data, preferably from head-to-head outcome studies and especially on compliance concerning discordant test results, calculations can be made with higher degrees of certainty.

Keywords Head-to-head comparison · 70-gene signature · MammaPrint · 21-gene recurrence score · Oncotype DX · Cost-effectiveness analysis

Electronic supplementary material The online version of this article (doi:10.1007/s10549-011-1769-7) contains supplementary material, which is available to authorized users.

V. P. Retèl · W. H. van Harten (✉)
Department of Psychosocial Research and Epidemiology, Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital (NKI-AVL), Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands
e-mail: w.v.harten@nki.nl

V. P. Retèl
e-mail: v.retel@nki.nl

M. A. Joore
Department of Health, Organization, Policy and Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands
e-mail: m.joore@mumc.nl

M. A. Joore
Department of Clinical Epidemiology and Medical Technology Assessment, Maastricht University Medical Center, PO Box 5800, 6202 AZ Maastricht, The Netherlands

W. H. van Harten
School of Governance and Management, MB-HTSR, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

Introduction

Both the 70-gene prognosis signature [1] and the 21-gene recurrence score assay [2] are relatively new prognostic

tests used to guide adjuvant treatment decisions in patients with early breast cancer. They outperform current guidelines, which offer most patients adjuvant chemotherapy, while 60–70% have a fairly good survival with locoregional treatment alone [1, 2]. While there are many studies performed regarding both diagnostic tests separately, no head-to-head comparison has yet been made.

In the current running randomized clinical trials, the "Microarray In Node-negative Disease may Avoid ChemoTherapy" (MINDACT-trial) [3] and "Trial Assigning IndividuaLized Options for Treatment (Rx)" (TAILOR-X trial) [4], the additional clinical value of both diagnostic instruments is separately being tested. In the MINDACT-trial, patients with discordant test results (70-gene signature result versus the web tool Adjuvant Online (AO) [5]; 70-gene low/AO high or 70-gene high/AO low) are randomized between decisions of adjuvant chemotherapy based on the 70-gene or AO risk assessment. In the TAILOR-X trial, patients with an intermediate 21-gene assay score are randomized to either adjuvant chemotherapy in combination with endocrine therapy or only endocrine therapy. Although the results of the prospective trials will only become available in some years, the tests have already been included in guidelines such as the National Comprehensive Cancer Network (NCCN), American Society of Clinical Oncology (ASCO), Dutch CBO 2008 and St. Gallen. However, the exact clinical use has to be established and the profiles have to be used selectively in cases where risk prediction is equivocal based on clinical variables [6]. The 21-gene assay may be more user friendly by using formalin-based tissue while the 70-gene signature needs fresh frozen tissue; the 70-gene, however, is more "decision friendly" using its dichotomous result "low" or "high" risk whereas the 21-gene assay provides an "intermediate" result, in part of the cases where the additional value of the decision using the prognostic test on whether or not to give adjuvant chemotherapy is unclear.

In the field of cost-effectiveness (CE), six CE analyses (CEAs) have been performed regarding gene expression profiles in breast cancer; four regarding the 21-gene assay [7–10] versus clinical guidelines such as NCCN, St. Gallen, and two CEAs are performed regarding the 70-gene signature versus clinical guidelines such as St. Gallen, AO and the National Institute of Health (NIH) guidelines [11, 12]. In the reported CEAs regarding the 21-gene assay, all patients with an intermediate or high risk were assumed to undergo hormonal (if endocrine responsive) and chemotherapy. In one CEA of the 21-gene assay it was calculated in the sensitivity analysis (SA) that 50% of the patients with an intermediate risk test result would receive hormonal and chemotherapy [9]. Both CEAs of the 70-gene signature assumed that patients with a high risk test result would undergo hormonal (if endocrine responsive) and chemotherapy. In all CEAs the genomic profile in question was found to be cost-effective compared to the clinical guideline used.

A CEA shows the CE of a technology versus the next best alternative. A CEA should compare all relevant alternatives [13]. Unfortunately, there is no CEA performed comparing both tests in one analysis, because a comparison of the "original" 70-gene signature and the "original" 21-gene assay in one independent dataset is not available. Answering the question which test performs best will require comparative effectiveness research. Government and industry seldom fund such studies because they may not offer much additional therapeutic promise as new discoveries do, and because industry is not eager to fund direct comparisons with competitive products [14]. The only articles in which both diagnostic tests are compared are Thomassen et al. [15] and Fan et al. [16], however, they do not use the "original" assays.

Why is it still important to perform a CEA directly comparing the tests in this case? Physicians have to choose between the two tests and the question which of the tests is most (cost-)effective, is relevant especially in view of the fact that the available data are not yet optimal. Data available should not guide the analysis; the decision problem should guide the analysis [17].

Therefore, we performed a direct CE comparison using the sensitivity and specificity of the 70-gene signature, the 21-gene assay and the St. Gallen 2003 [18] based on the Thomassen-series [15], or using the sensitivity and specificity of the 70-gene signature, the 21-gene assay and the AO based on the Fan-series [16]. In addition, the impact of changes in compliance is calculated since it is known that there is seldom full compliance with guidelines and that for both prognostic tests compliance with the test result may be an issue, as shown in the pilot study of the MINDACT-trial [19].

Methods

Sensitivity and specificity of the genomic tests

Sensitivity and specificity of the 70-gene, 21-gene and St. Gallen (SG) or AO were derived from two available datasets; Thomassen et al. [15] ($n = 60$), and Fan et al. [16] ($n = 101$).

The Thomassen-series [15] assessed both gene expression profiles and clinical characteristics using the same algorithms on one platform. In this study, the comparison of prediction of metastasis in a low-malignant breast cancer group is made. The study is designed with pairs of metastasizing and non-metastasizing tumours matched according to classic prognostic markers, developing

classification algorithms reducing the effect of different platforms [15]. All tumours in this database were included in the current study. In the model, each strategy was based on the sensitivity and specificity of the prognostic test, these were derived from figures 1B and 1H from the Thomassen paper [15]. Patients were classified as having a true low, true high, false low, or false high risk of developing metastasis. In the Thomassen-series, this classification was generated by a "probability of poor outcome" cut-off of 0.5, applied to all classifications of the used diagnostic tests [15].

In the Fan-series [16], the gene expression data set containing 295 tumours was derived by researchers from the Netherlands Cancer Institute and Rosetta Inpharmatics—Merck using oligonucleotide microarrays (Agilent). Tumours with node-negative and ER-positive characteristics were selected from this database. We calculated the sensitivity and specificity of the 70-gene, 21-gene and AO (see Table 1). The intermediate risk patients of the 21-gene assay were grouped together with the high risk (as the former analysis also did), assuming that both intermediate and high risk patients received hormonal and chemotherapy.

Furthermore, for both datasets, it was assumed that the low risk patients received only hormonal therapy, and the high risk patients hormonal and chemotherapy (see more details in Retèl et al. [12]).

Compliance rates

We used the compliance rates regarding discordant test results from the clinical trial data of the MINDACT pilot (first 800 patients) [19]. The compliance rates were modelled for the discordant cases clinical low/genomic high risk (13%) and clinical high/genomic low risk (4%) for both strategies and both datasets. The compliance rates were incorporated in the sensitivity and specificity of the diagnostic tests (see Table 2).

Decision model

Previously, a Markov decision model was developed to assess the effects (life years (LYs) and quality adjusted LYs (QALYs)), health care costs and CE of the 70-gene signature as compared to clinical guidelines (such as SG and AO) in patients with early, node-negative, oestrogen receptor positive breast cancer patients [12]. A QALY is defined as a LY multiplied by a quality of life weight between 0 and 1, for instance 2 years with quality of life 0.8 amounts to 1.6 QALYs. The model was constructed with four mutually exclusive health states: disease free survival, relapse (including local and regional recurrences, secondary primary and contralateral breast cancer), distant metastasis, and death (Fig. 1). The study adopts a health care perspective. For further model details, see Retèl et al. [12]. In the current analysis, the 21-gene assay was added as a comparator. The calculations are performed per year, with a total simulated time horizon of 20 years. Future costs and effects were discounted to their present value by a rate of 4 and 1.5% per year, respectively, according to Dutch guidelines [20].

Cost and utility input

The costs of the 70-gene signature were €2675, provided by Agendia BV; full costs including transport, additional specimen processing at the local hospital and value added tax (VAT). The costs of the 21-gene assay were based on the manufacturer's retail price of $4075 (€3179), derived from the website of Genomic Health Inc. Costs were expressed in 2010 Euros. For other cost and utility input, see more details in Retèl et al. [12].

Analysis

Incremental CE ratios (ICERS) were calculated by dividing the incremental costs by incremental LYs and by incremental QALYs. Uncertainty in the input parameters was handled probabilistically, by assigning distributions to parameters (Table 1) [21]. The results of a simulation of 1000 patients representing the dataset are illustrated in a CE plane. In the CE plane each quadrant indicates whether a strategy is more or less expensive and more or less effective. Whether a strategy is deemed efficient depends on how much society is willing to pay for a gain in effect, which is referred to as the ceiling ratio [22]. In the Netherlands an informal ceiling ratio of €80,000 per QALY exists (Dutch Council for Public Health and Health Care 2006) [23]. This is a maximum ceiling ratio which applies when there is a high burden of disease. This is certainly the case for breast cancer. In the US this threshold is $ 50,000–100,000/QALY. And in the UK, the NICE handles a threshold of £ 20,000–30,000/QALY [24]. In this study, we handle the Dutch ceiling ratio of €80,000/QALY. In theory, when the differences in costs divided by the differences in outcomes is above this ceiling ratio, the strategy is not considered cost-effective. To show this decision uncertainty, CE acceptability curves (CEACs) are presented.

Sensitivity analyses

We performed four SA concerning the used dataset, to show the robustness of the results. First, because we expect that the 21-gene assay could be in disadvantage [25], we calculated a SA regarding higher sensitivity and specificity for the 21-gene assay. For each database, we improved the

Table 1 Test performance base case for Thomassen and Fan

Parameter	Risk group		*N*	se	sp	*P*	SE	Distribution	Source
Base case Thomassen									
70-G	Low	True	25	0.70	0.83	0.417	0.07	Dirichlet	15
		False	9			0.150	0.12	Dirichlet	
	High	True	21			0.350	0.08	Dirichlet	
		False	5			0.083	0.07	Dirichlet	
21-G	Low	True	22	0.73	0.73	0.367	0.07	Dirichlet	
		False	8			0.133	0.12	Dirichlet	
	High	True	22			0.367	0.08	Dirichlet	
		False	8			0.133	0.08	Dirichlet	
SG	Low	True	13	0.57	0.43	0.217	0.09	Dirichlet	
		False	13			0.283	0.09	Dirichlet	
	High	True	17			0.217	0.09	Dirichlet	
		False	17			0.283	0.09	Dirichlet	
Base case Fan									
70-G	Low	True	46	0.74	0.70	0.455	0.06	Dirichlet	16
		False	9			0.089	0.08	Dirichlet	
	High	True	26			0.257	0.07	Dirichlet	
		False	20			0.198	0.06	Dirichlet	
21-G	Low	True	29	0.89	0.44	0.287	0.06	Dirichlet	
		False	4			0.040	0.05	Dirichlet	
	High	True	31			0.307	0.05	Dirichlet	
		False	37			0.366	0.06	Dirichlet	
AO	Low	True	41	0.66	0.62	0.406	0.06	Dirichlet	
		False	12			0.119	0.08	Dirichlet	
	High	True	23			0.228	0.08	Dirichlet	
		False	25			0.248	0.06	Dirichlet	

70-G 70-gene signature, *21-G* 21-gene assay recurrence score, *SG* St. Gallen guidelines (2003), *AO* Adjuvant Online, *se* sensitivity, *sp* specificity, *P* probability, *SE* standard deviation

true low and true high group with one patient. Second, we used the compliance rates of the feasibility studies of Bueno-de-Mesquita et al. [26] and Lo et al. [27] as SA, to show the "worst case" scenario when including non-compliance. We incorporated non-compliance rates with the genomic test results based on two articles in which compliance was measured. The non-compliance rates were modelled for the discordant cases clinical low/genomic high risk and clinical high/genomic low risk. In Bueno-de-Mesquita et al. [26], the non-compliance rate for the 70-gene signature in case of a clinical high and genomic low risk was 60%, in case of clinical low and genomic high it was 43%. In Lo et al. [27], the non-compliance rate for the 21-gene assay of, respectively, clinical high/genomic low and clinical low/genomic high was 25 and 88% (see Table 2). In this calculation, we have taken together the intermediate and the high risk group who are assumed to receive chemotherapy. Third, because using QALY as an outcome in CEAs in oncology is a debated issue, as this has proven to be difficult to estimate health state utilities among cancer patients [28], we used different QoL-scores (utilities) for disease free survival with and without adjuvant systemic therapy [29]. Finally, because the costs of chemotherapy are likely to become higher in the future with the increase of novel regimens (e.g. Taxanes), the costs of chemotherapy were varied to €20,000 [6]. CEACs frontiers are used to show the impact of these changes in model input on the probability that the 70-gene is cost-effective.

Results

Mean results

For both series, whereas the 70-gene signature yielded more QALYs and was less costly, the 21-gene amounted more LYs but was more costly (see Table 3).

For the Thomassen-series, the number of LYs amounted to 14.76 for the 21-gene, 14.61 for the 70-gene and 14.04 for the SG. The QALYs of the 70-gene yielded 11.41,

Table 2 Test performance taking into account compliance

Parameter	Risk group		Probability	SE	Distribution	Source
Incorporated non-compliance rates in the Thomassen-series						
70-G low	SG high	Discordance	0.640	0.09	Beta	15
		Non-compliance	0.040	0.02	Beta	19
70-G high	SG low	Discordance	0.360	0.09	Beta	15
		Non-compliance	0.130	0.05	Beta	19
21-G low	SG high	Discordance	0.560	0.10	Beta	15
		Non-compliance	0.040	0.02	Beta	19
21-G high	SG low	Discordance	0.440	0.10	Beta	15
		Non-compliance	0.130	0.05	Beta	19
Incorporated non-compliance rates in the Fan-series						
70-G low	AO high	Discordance	0.530	0.08	Beta	16
		Non-compliance	0.040	0.02	Beta	19
70-G high	AO low	Discordance	0.470	0.08	Beta	16
		Non-compliance	0.130	0.05	Beta	19
21-G low	AO high	Discordance	0.430	0.08	Beta	16
		Non-compliance	0.040	0.02	Beta	19
21-G high	AO low	Discordance	0.570	0.08	Beta	16
		Non-compliance	0.130	0.05	Beta	19

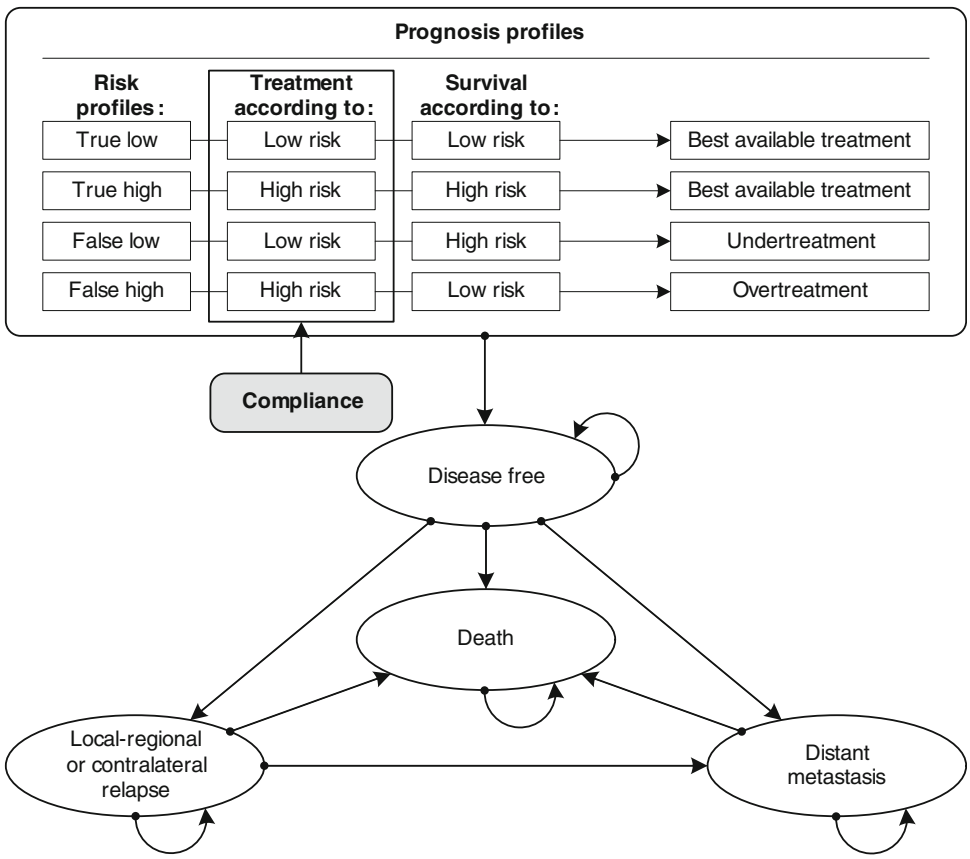

Fig. 1 The cost-effectiveness model structure

11.33 for the 21-gene and 10.41 for the SG. The total health care costs per patient were €40,393 for the 70-gene, €41,868 for the 21-gene and €44,232 for the SG. When focusing on survival, the 21-gene strategy has the highest probability of being cost-effective, with a willingness to pay of €1,475/LY and higher, taken into account a ceiling

Table 3 Mean results base case for Thomassen and Fan

Strategy	LYs	Costs	ΔLys (CI)	ΔCosts (CI)	ICER
Thomassen					
21-G	14.76	€41,868			
70-G	14.61	€40,393			
SG	14.04	€44,232			
21-G vs 70-G			0.14 (−0.99 to 1.27)	€1,475 (−7,988 to 10,920)	€1,475
21-G vs SG			0.72 (−0.51 to 1.90)	−€2,364 (−10,831 to 6,519)	DOM
Fan					
21-G	15.86	€38,799			
70-G	15.26	€34,858			
AO	15.00	€34,115			
21-G vs 70-G			0.40 (−0.73 to 0.77)	€3,941 (−3,969 to 8,945)	€9,272
70-G vs AO			0.26 (−0.52 to 1.05)	€743 (−5,967 to 6,727)	€2,913
Strategy	QALYs	Costs	ΔQALYs (CI)	ΔCosts (CI)	ICER
Thomassen					
70-G	11.41	€40,393			
21-G	11.33	€41,868			
SG	10.41	€44,232			
70-G vs 21-G			0.08 (−1.01 to 1.11)	−€1,475 (−10,920 to 7988)	DOM
70-G vs SG			1.00 (−0.06 to 1.91)	€3,839 (−13,307 to 5,256)	€3,839
Fan					
70-G	11.92	€34,858			
21-G	11.61	€38,799			
AO	11.61	€34,115			
70-G vs 21-G			0.31 (−0.49 to 0.90)	−€3,941 (−8,945 to 3,969)	DOM
21-G vs AO			0.00 (−0.61 to 0.86)	€4,684 (−4,088 to 9,457)	€1,6 mill

Thomassen and Fan Incremental cost-effectiveness results (ICER)

CI 95% confidence interval, *70-G* 70-gene signature, *21-G* 21-gene assay, *SG* St. Gallen guidelines (2003), *AO* Adjuvant Online, Δ incremental, *DOM* dominant, *mill* million, *LYs* life years, *QALYs* quality adjusted life years, *vs* versus

ratio of €80,000/QALY. In case of costs/QALY, the 70-gene signature has the highest probability of being cost-effective, with less cost and higher survival (see Fig. 2).

For the Fan-series, the number of LYs amounted to: 15.26 (70-gene), 15.86 (21-gene) and 15.00 (AO). The QALYs amounted to: 11.92 (70-gene), 11.61 (21-gene) and 11.61 (AO). The total health care costs per patient were: €38,779 (21-gene), €34,858 (70-gene) and €34,115 (AO). The difference in costs per LY gained of the 21-gene compared to the 70-gene resulted in equal LYs but more costs for the 21-gene. While focusing on costs/QALY, the 70-gene yields more QALYs and less costs than the other strategies.

The uncertainty surrounded by the Thomassen-series amounted 55% for the LYs and 55% for the QALYs, for the Fan-series 80% for the LY and 65% for the QALYs.

Taking reported compliance with discordant test results into account, in general, the effect of all strategies slightly decreased, the costs slightly increased and with more uncertainty around the decision (see Table 4; Fig. 3).

Sensitivity analyses

When improving the outcome for the 21-gene, the results of the costs/LY appeared stronger in both datasets, for the costs/QALY, the 70-gene signature remained most cost-effective in the Fan-series.

For the second sensitivity analyses regarding other compliance input, for the Thomassen-series, the 70-gene signature became the most cost-effective strategy when focusing on survival, when focusing on quality adjusted survival, the AO strategy became most cost-effective strategy. For the Fan-series, the AO became most cost-effective for both LY and QALYs. For both analyses the probability of CE was around 50%, which means that the decision of CE has substantial uncertainty in this case.

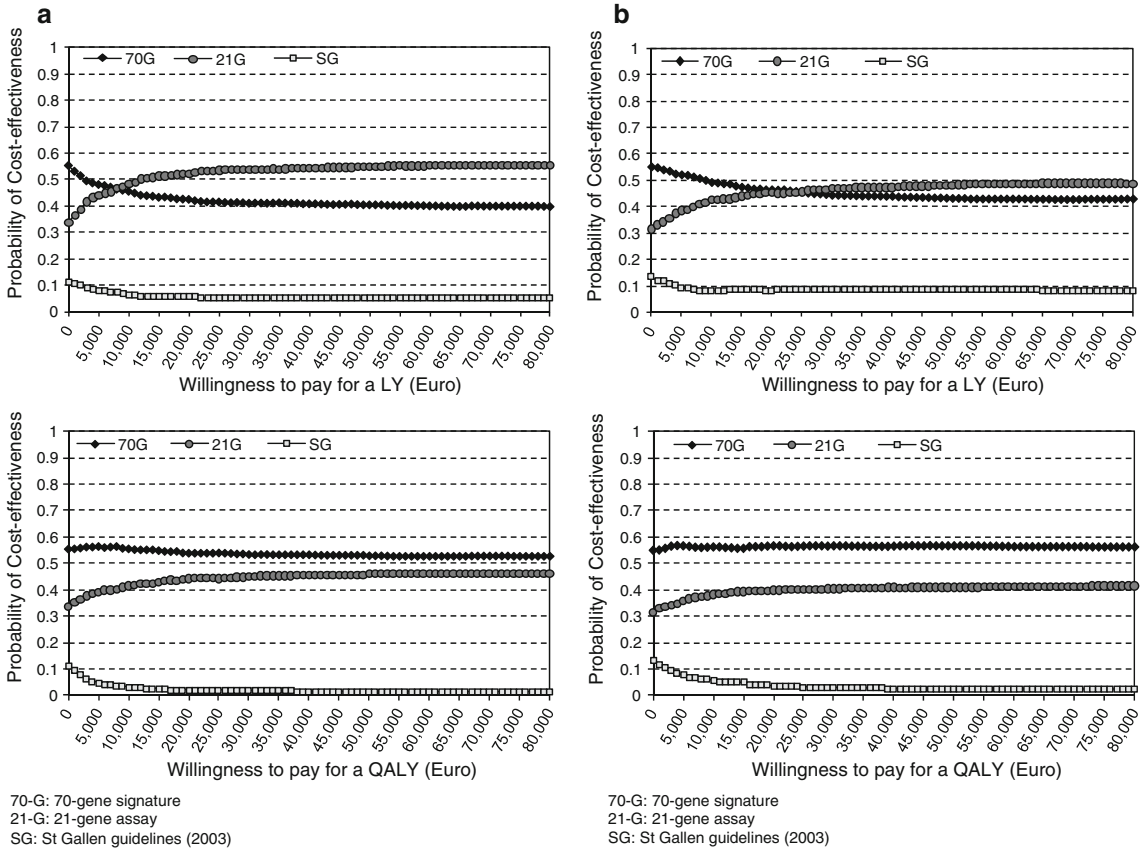

Fig. 2 Cost-effectiveness acceptability curves (LY and QALY) based on the Thomassen-series, for the base case (**a**) and including compliance (**b**); presenting the probability of cost-effectiveness for a range of values of thresholds (ceiling ratios, willingness to pay for one QALY)

Lower utilities and higher chemotherapy costs showed the same pattern as the base case, only slightly shifted (see Supplementary Figure 1).

Discussion

Based on the currently available data, and assuming that there was 100% compliance in case of discordant test results, the 21-gene has the highest probability of being cost-effective when focusing on cost/LY, however, while when focusing on cost/QALY, the 70-gene signature has the highest probability of being cost-effective, taking into account a threshold of €80,000/QALY. These results are concerning expected CE, as the outcome data of the currently ongoing trials are not yet available. The analyses yielded more uncertainty surrounding the Thomassen-series compared to the Fan-series, probably due to the small patient group. Using the reported non-compliance with discordant test results, the trend of the mean results remained, although a bit tempered and with higher uncertainty.

The data derived from both datasets have some remaining issues [25]. For the Thomassen-series [15], the profiles are performed on one algorithm, thus reducing the bias of different platforms, but on the other hand somewhat lower accuracy for both tests. In addition, hardly any patient has been treated with Tamoxifen, which is an eligibility criteria for the 21-gene assay. This could be in favour of the 70-gene signature [25]. The data derived from the Fan-series [16] has also remaining issues, such as the fact that the profiles are partly performed on the original dataset of the development of the 70-gene signature, whereas the 21-gene assay is performed on fresh frozen tissue instead of paraffin, which could also suggest that the results were in favour of the 70-gene signature [25]. This possible biases were the reason we performed the sensitivity analyses with improved outcome for the 21-gene, which showed that when focusing on survival, the 21-gene remained cost-effective, however, when focusing on quality adjusted survival, the 70-gene remained most cost-effective. We can conclude that this is a main driver for outcomes and that the most ideal design should be a head-to-head prospective trial where both diagnostic tests are

Table 4 Mean results base case taking into account compliance

Strategy	LYs	Costs	ΔLys (CI)	ΔCosts (CI)	ICER
Thomassen					
21-G	14.61	€42,227			
70-G	14.51	€40,813			
SG	14.04	€44,232			
21-G vs 70-G			0.10 (0.19 to 2.09)	€1,412 (−10,211 to 5,592)	€14,862
70-G vs SG			0.47 (−0.66 to 1.77)	−€3,419 (−10,862 to 7,196)	DOM
Fan					
70-G	15.14	€35,068			
21-G	15.11	€37,135			
AO	15.00	€34,116			
70-G vs 21-G			0.03 (−0.26 to 1.03)	−€2,067 (−4,585 to 7,558)	DOM
21-G vs AO			0.11 (−0.71 to 0.89)	€3,019 (−3,284 to 9,646)	€28,123
Strategy	QALYs	Costs	ΔQALYs (CI)	ΔCosts (CI)	ICER
Thomassen					
70-G	11.32	€40,813			
21-G	11.24	€42,227			
SG	10.41	€44,232			
70-G vs 21-G			0.08 (−0.31 to 1.86)	−€1,412 (−11,743 to 6,069)	DOM
21-G vs SG			0.82 (0.39 to 1.97)	−€2,007 (−12,437 to 1,829)	DOM
Fan					
70-G	11.86	€35,068			
21-G	11.64	€37,135			
AO	11.61	€34,116			
70-G vs 21-G			0.22 (−0.46 to 0.85)	−€2,067 (−8,714 to 4,435)	DOM
21-G vs AO			0.03 (−0.68 to 0.75)	€3,019 (−3,284 to 9,646)	€79,470

Thomassen and Fan Incremental cost-effectiveness results (ICER) of sensitivity analyses

CI 95% confidence interval, *70-G* 70-gene signature, *21-G* 21-gene assay, *SG* St. Gallen guidelines (2003), *AO* Adjuvant Online, Δ incremental, *DOM* dominant

being compared in one population. A next step would be to synthesize all available evidence, by using mixed treatment comparison (MTC). MTC allows for indirect comparisons and can therefore provide useful information for clinical and reimbursement decision making in the absence of head-to-head data [30].

We incorporated compliance rates from the MINDACT-trial pilot [19], however, one could dispute whether these compliance rates are reflecting real world compliance as they are based on a randomized setting. Two other published articles, which were used in the SA, were available regarding compliance in a non-randomized setting. The study of Bueno-de-Mesquita was the first who published compliance data and based on an early adoption phase trial, in which a suboptimal compliance can be expected upfront. The Lo-series have been commented regarding their way of presenting the compliance rates [31]. On the other hand, both were feasibility studies (no randomization effect) and both were performed in the same time span; from 01/2004 till 12/2006 for the Bueno-de-Mesquita-series [26] and from 12/2005 till 08/2006 for the Lo-series [27]. The St. Gallen guidelines of 2003 were used for the comparison in the Thomassen-series, ideally we should use the more current guidelines of 2009. It would be interesting to look closer into the mechanisms behind non-compliance as they are of great influence on the CE in daily practice; why do physicians decide whether or not to follow the guideline or the genomic test result? Apparently, it seems that the compliance increases over time as we can see in the MINDACT-pilot phase, where the compliance to treatment according to the different categories is much higher (95%) [19]. This issue also appeared to be a driver for outcomes; if a policy decision must be made based on the analyses incorporating compliance, the results using the compliance rates of the feasibility studies show that the results on CE are different.

A last driver for deciding based on CEA outcomes is the question what is more important; costs per LY, or costs per

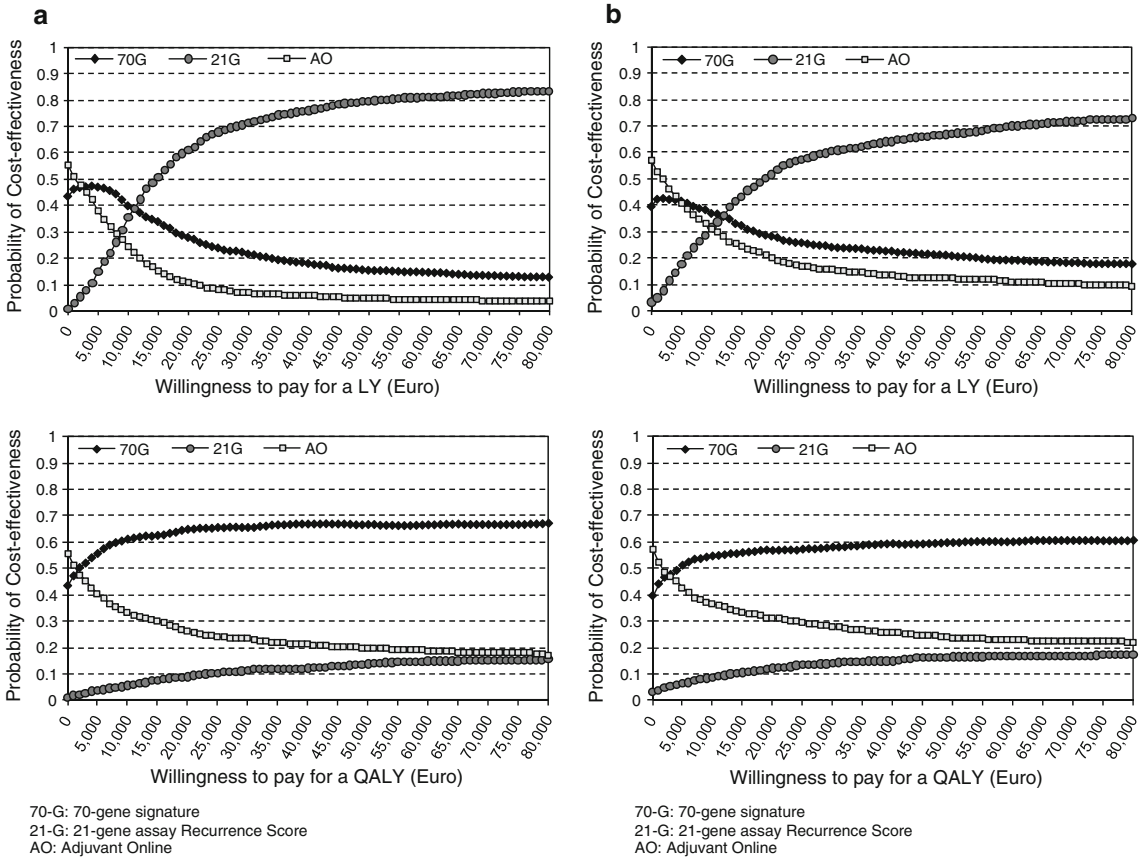

Fig. 3 Cost-effectiveness acceptability curves (LY and QALY) based on the Fan-series, for the base case (**a**) and including compliance (**b**); presenting the probability of cost-effectiveness for a range of values of thresholds (ceiling ratios, willingness to pay for one QALY)

QALYs? The measurements of utilities are debated, as it has proven to be difficult to estimate health state utilities, especially among cancer patients [28], however, the side effects of for example chemotherapy are impossible to ignore. In this case, decision making based on CE is different when only focusing on survival, or taking also the quality of those LYs into account.

Conclusion

The results of the previous performed CEAs all showed that both the 21-gene assay and the 70-gene signature are cost-saving and/or cost-effective strategies as compared to clinicopathological guidelines. However, one has to be careful in such a comparison because of the different settings in the reported trials. This study, however, indicates that the CEA performances of the 70-gene signature and the 21-gene assay based on reported studies are close, and the uncertainty is high. The 70-gene signature seems to have the highest probability to be cost-effective when focusing on cost/QALY, while the 21-gene assay seems to have the highest probability to be cost-effective when focusing on cost/LY. The level of compliance can have serious impact on the CE. With additional data, preferably from head-to-head outcome studies and especially on compliance concerning discordant test results, calculations can be made with higher degrees of certainty. Therefore, it is recommended to invest on knowledge transfer regarding the clinical value of the gene expression profiles.

Acknowledgments We would like to thank Prof. Dr. EJT Rutgers and Dr. SC Linn for their contribution to this manuscript. This study was funded by the Dutch Health Care Insurance Board (DHCIB, CVZ), the Netherlands. The DHCIB had no role in the study design and the collection, analysis, and interpretation of data and the writing of the article and the decision to submit it for publication.

Conflict of interest W.H. van Harten is a non-remunerated, non-stake holding member of the supervisory board of Agendia BV. All other authors declared no conflicts of interest.

References

1. Van't Veer LJ, Dai H, van de Vijver MJ et al (2002) Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530–536

2. Paik S, Shak S, Tang G et al (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 351:2817–2826
3. Cardoso F, Van't Veer L, Rutgers E et al (2008) Clinical application of the 70-gene profile: the MINDACT trial. J Clin Oncol 26:729–735
4. Sparano JA, Paik S (2008) Development of the 21-gene assay and its application in clinical practice and clinical trials. J Clin Oncol 26:721–728
5. Ravdin PM, Siminoff LA, Davis GJ et al (2001) Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. J Clin Oncol 19:980–991
6. Goldhirsch A, Ingle JN, Gelber RD et al (2009) Threshold for therapies: highlights of the St Gallen international expert consensus on the primary therapy of early breast cancer 2009. Ann Oncol 20:1319–1329
7. Hornberger J, Cosler LE, Lyman GH (2005) Economic analysis of targeting chemotherapy using a 21-gene RT-PCR assay in lymph-node-negative, estrogen-receptor-positive, early-stage breast cancer. Am J Manag Care 11:313–324
8. Lyman GH, Cosler LE, Kuderer NM et al (2007) Impact of a 21-gene RT-PCR assay on treatment decisions in early-stage breast cancer: an economic analysis based on prognostic and predictive validation studies. Cancer 109:1011–1018
9. Kondo M, Hoshi SL, Ishiguro H et al (2008) Economic evaluation of 21-gene reverse transcriptase-polymerase chain reaction assay in lymph-node-negative, estrogen-receptor-positive, early-stage breast cancer in Japan. Breast Cancer Res Treat 112:175–187
10. Tsoi DT, Inoue M, Kelly CM et al (2010) Cost-effectiveness analysis of recurrence score-guided treatment using a 21-gene assay in early breast cancer. Oncologist 15:457–465
11. Oestreicher N, Ramsey SD, Linden HM et al (2005) Gene expression profiling and breast cancer care: what are the potential benefits and policy implications? Genet Med 7:380–389
12. Retèl VP, Joore MA, Knauer M et al (2010) Cost-effectiveness of the 70-gene signature versus St. Gallen guidelines and Adjuvant Online for early breast cancer. Eur J Cancer 46:1382–1391
13. Drummond MF, Sculpher MJ, Torrance GW et al (2005) Methods for the economic evaluation of health care programmes. Oxford University Press, Oxford
14. Marchionni L, Wilson RF, Wolff AC et al (2008) Systematic review: gene expression profiling assays in early-stage breast cancer. Ann Intern Med 148:358–369
15. Thomassen M, Tan Q, Eiriksdottir F et al (2007) Comparison of gene sets for expression profiling: prediction of metastasis from low-malignant breast cancer. Clin Cancer Res 13(18):5355–5360
16. Fan C, Oh DS, Wessels L et al (2006) Concordance among gene-expression-based predictors for breast cancer. N Engl J Med 355:560–569
17. Grutters JPC, Seferina SC, Tjan-Heijnen VCG et al (2011) Bridging trial and decision: a checklist to frame health technology assessments for resource allocation decisions. Value Health 14(5):777–784
18. Goldhirsch A, Wood WC, Gelber RD et al (2003) Meeting highlights: updated international expert consensus on the primary therapy of early breast cancer. J Clin Oncol 21:3357–3365
19. Rutgers E, Piccart M, Delaloge S et al (2010) The EORTC 10041/BIG 03-04 MINDACT trial is feasible: first results of the pilot phase (Abstract 444). Eur J Cancer Suppl 8(3):188
20. Oostenbrink JB, Koopmanschap MA, Rutten FFH (2000) Handleiding voor kostenonderzoek, methoden en richtlijnprijzen voor economische evaluaties in de gezondheidszorg. College voor zorgverzekeringen, Amstelveen
21. Weinstein MC (2006) Recent developments in decision-analytic modelling for economic evaluation. Pharmacoeconomics 24:1043–1053
22. Fenwick E, Claxton K, Sculpher M (2001) Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ 10:779–787
23. (2006) Raad voor de Volksgezondheid en Zorg: Zinnige en duurzame zorg [Sensible and durable care]. Raad voor de Volksgezondheid en Zorg, Zoetermeer (in Dutch). Available at: http://rvz.net/uploads/docs/Advies_-_Zinnige_en_duurzame_zorg.pdf
24. Lothgren M, Zethraeus N (2000) Definition, interpretation and calculation of cost-effectiveness acceptability curves. Health Econ 9:623–630
25. Haibe-Kains B, Desmedt C, Piette F et al (2008) Comparison of prognostic gene expression signatures for breast cancer. BMC Genomics 9:394
26. Bueno-de-Mesquita JM, van Harten W, Retel V et al (2007) Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: a prospective community-based feasibility study (RASTER). Lancet Oncol 8:1079–1087
27. Lo SS, Mumby PB, Norton J et al (2010) Prospective multicenter study of the impact of the 21-gene recurrence score assay on medical oncologist and patient adjuvant breast cancer treatment selection. J Clin Oncol 28:1671–1676
28. Earle CC, Coyle D, Evans WK (1998) Cost-effectiveness analysis in oncology. Ann Oncol 9:475–482
29. Tengs TO, Wallace A (2000) One thousand health-related quality-of-life estimates. Med Care 38:583–637
30. Jansen JP, Crawford B, Bergman G, Stam W (2008) Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. Value Health 11(5): 956–964
31. Brauchli P, Thürlimann B, Crowe SN, Hermann R (2010) What is the value of the 21-gene recurrence score? J Clin Oncol 28(32):671–672