

Cost-Effectiveness Analyses of the 21-Gene Assay in Breast Cancer: Systematic Review and Critical Appraisal

Shi-Yi Wang, Weixiong Dang, Ilana Richman, Sarah S. Mougalian, Suzanne B. Evans, and Cary P. Gross

Author affiliations and support information (if applicable) appear at the end of this article.

Published at jco.org on April 16, 2018.

Corresponding author: Shi-Yi Wang, MD, PhD, Department of Chronic Disease Epidemiology, Yale School of Public Health, 60 College St, P.O. Box 208034, New Haven, CT 06520; e-mail: shiyi.wang@yale.edu.

© 2018 by American Society of Clinical Oncology

0732-183X/18/3616w-1619w/\$20.00

A B S T R A C T

Purpose

Prior studies examining cost effectiveness of the 21-gene assay (Oncotype DX [ODX]) for women with hormone receptor-positive, early-stage breast cancer have yielded disparate results. We aimed to explore why these analyses may have yielded different conclusions.

Methods

We conducted a systematic literature review of cost-effectiveness analyses (CEAs) of ODX. We examined the extent to which the structure of CEA modeling, the assumptions of the models, and the selection of input parameters influenced cost-effectiveness estimates. We also explored the prevalence of industry funding and whether industry funding was associated with study designs favoring ODX.

Results

We identified 27 analyses, 15 of which received industry funding. In 18 studies, the clinical characteristics (eg, tumor size and grade) commonly used to make chemotherapy decisions were not incorporated into simulation modeling; thus, these studies would favor ODX being cost effective and might not reflect clinical practice. Most studies ignored the heterogeneous effect of ODX on chemotherapy use; only five studies assumed that ODX would increase chemotherapy use for clinically low-risk patients but decrease chemotherapy use for clinically high-risk patients. No study used population-based joint distributions of ODX recurrence score and tumor characteristics, and 12 studies inappropriately assumed that chemotherapy would increase distant recurrence for the low recurrence score group; both approaches overestimated the benefits of ODX. Industry-funded studies tended to favor ODX; all five studies that reported ODX as being cost saving were industry funded. In contrast, two studies that reported an incremental cost-effectiveness ratio > \$50,000 per quality-adjusted life-year were not funded by industry.

Conclusion

Although a majority of published analyses indicated that ODX is cost effective, they incorporated study designs that can increase the risk of bias.

J Clin Oncol 36:1619-1627. © 2018 by American Society of Clinical Oncology

INTRODUCTION

The use of gene-expression profiling and next-generation sequencing is increasing in clinical care,^{1,2} and assessing the relationship between cost and benefit is important. A widely used 21-gene assay known as Oncotype DX (ODX; Genomic Health, Redwood City, CA) illustrates the challenges of economic evaluation of newer genomic testing. ODX provides clinical treatment guidance for patients with hormone receptor-positive, early-stage breast cancer.³ For these patients, the decision to use adjuvant chemotherapy is difficult.⁴ Although some women may benefit, others may experience chemotherapy-related adverse effects and incur substantial health care costs.⁵

Clinicopathologic assessment, such as tumor size and pathologic grade, provides some indication of recurrence risk; however, such information is still suboptimal for prognostication.

ODX quantifies the expression of 21 genes in breast cancer tissue with a recurrence score (RS), which predicts long-term risk of distant recurrence⁶ and outperforms clinicopathologic assessment.^{7,8} Indeed, ODX has been widely used in oncology practice to guide chemotherapy decision making.⁹ However, ODX is expensive, having cost approximately \$3,400 per test since its introduction.^{10,11} Furthermore, National Comprehensive Cancer Network guidelines suggest ODX testing for select rather than for all patients.¹² Because unrestricted use of ODX could lead to considerable economic burden from both

ASSOCIATED CONTENT

Appendix
DOI: <https://doi.org/10.1200/JCO.2017.76.5941>

DOI: <https://doi.org/10.1200/JCO.2017.76.5941>

individual and societal perspectives, economic analyses are critical. For instance, the UK National Institute for Health and Care Excellence (NICE)^{13,14} considers ODX to be cost effective for patients at intermediate risk of distant recurrence, such as patients with a Nottingham Prognostic Index (NPI) score > 3.4 (detailed NPI information¹⁵ is available in Appendix, online only). However, several economic analyses^{10,16,17} and systematic reviews^{18,19} of ODX cost effectiveness have concluded that ODX is cost effective for a much larger, more diverse group. The resulting discrepancy that seems to exist between guidelines and most published cost-effectiveness analyses (CEAs) has created uncertainty regarding the subgroups for which ODX is cost effective. To our knowledge, no systematic review has investigated the potential reasons why prior CEAs have reached different results. Understanding the underlying mechanism leading to this discrepancy could advance our knowledge in the economic evaluation of precision medicine.

Accordingly, we conducted a systematic review of CEAs of ODX to identify specific study characteristics that are important aspects of genetic and molecular testing economic evaluations. We then assessed how these study characteristics might affect results, as well as the magnitude of the influence. Additionally, literature has suggested that funding sources could affect CEA findings²⁰⁻²²; however, no study has explored the relationship between ODX CEA results and study sponsorship. Thus, we examined the frequency with which published analyses were funded by industry and whether the funding sources were associated with study designs that might lead to different conclusions. Given health care budget constraints, our results could provide critical insights for value-based frameworks through appropriate targeting of patient populations for whom ODX may be most beneficial.

METHODS

Systematic Review

We conducted a literature search for published economic evaluations of ODX using medical subject headings and text keywords in the PubMed database (search term details are listed in Appendix Table A1, online only). We excluded reviews, editorial letters, articles not in English, and CEAs that did not compare ODX with no ODX. Two investigators (S.-Y.W. and W.D.) abstracted data from the studies. A third reviewer (I.R.) checked data accuracy. We abstracted country of study, funding source, population (estrogen, human epidermal growth factor receptor 2 [HER2], and lymph node status), age of population, perspectives of CEA, and time horizon. We also abstracted the comparison groups, sources of key input parameters, model assumptions, and the manner in which a sensitivity analysis was performed. Any discrepancies were discussed, and consensus among the reviewers was reached. We used the Quality of Health Economic Studies (QHES) instrument to evaluate the quality of each CEA.^{23,24} The QHES instrument, consisting of 16 criteria, has been validated to assess methodologic characteristics and transparency of reporting in economic evaluations. We did not conduct a meta-analysis, because we expected there would be large heterogeneity across studies.

Appraisal of Existing CEAs

To understand the potential mechanisms that could explain the discrepancy across published CEAs, we first examined whether the model in the published article reflected the clinical use of ODX. We conceptualized a model structure for each study. The simulated chemotherapy decision-making processes were compared with real-world clinical practices. Second, we evaluated model assumptions for each study, focusing on those

without strong supporting evidence. Third, we identified problematic input parameters within each study. Finally, we examined the magnitude to which the model structure, assumptions, and input parameters in the CEA models could influence results. To do so, we built a simulation model that adopted the NICE analytic model¹⁴ to examine ODX cost effectiveness (Appendix; Appendix Table A2, online only). Using the incremental cost-effectiveness ratios (ICERs) estimated from our model as the benchmark, we evaluated the effect of each concerning issue on ICERs. That is, we conducted a series of simulations where each simulation applied one problematic assumption or input parameter in the model. The corresponding ICER estimates were compared with our benchmark estimates.

Sponsorship and Study Outcome

We compared incremental costs and incremental effectiveness gained by whether the study was financially associated with Genomic Health (the proprietor of ODX), including Genomic Health as the funding source or any author affiliated with Genomic Health. We extracted the incremental costs and effectiveness reported by the investigators from all studies identified. All costs were converted to US dollars, using the year-specific currency exchange rates, and converted to 2015 US dollars for comparison. We then calculated the average incremental cost and effectiveness according to the relationship with Genomic Health.

RESULTS

After screening abstracts and reviewing potentially relevant articles, we identified 27 eligible studies from 26 articles (Appendix Fig A1, online only).^{10,14,16,17,25-46} The manuscript published by Ward et al¹⁴ consisted of two studies: one by Genomic Health (hereafter as Ward-industry) and one by NICE (hereafter as Ward-government). Among the 27 studies, 10 came from Europe or the United Kingdom, six from the United States, and seven from Canada (Table 1). For the analyses of base-case scenarios, all studies concluded that ODX has an ICER of $< \$100,000$ per quality-adjusted life-year (QALY). Most studies used ≥ 25 years as their timeframe, and differences in time horizon were not significantly associated with different ICERs (Appendix Fig A2, online only). The quality of the identified analyses, measured by QHES scores, was generally high; the mean QHES score was approximately 88 (best score, 100; 95% CI, 85.2 to 90.4; Appendix Table A3, online only).

However, when we assessed simulation models, we identified eight issues that might compromise the accuracy and validity of the results (Table 2). Some issues were common across the 27 studies (Table 3); in their base-case analyses, all studies had at least four concerning issues, and three studies had all eight concerning issues. We categorized them as model structure, model assumptions, and model input parameters.

Model Structure

Ignoring clinicopathologic information. In the absence of ODX, clinicians frequently use risk stratification schema, such as Adjuvant! Online (AO) and PREDICT,^{47,48} to guide their chemotherapy decisions.⁴⁹⁻⁵¹ These tools use readily available clinical characteristics to stratify patients according to recurrence risk. Hence, when evaluating ODX cost effectiveness, investigators should assess the degree to which ODX adds important information beyond these widely used tools. To construct a model that reflects clinical practice, investigators should compare clinicopathologic information plus ODX versus clinicopathologic

Table 1. Report Characteristics (N = 27)

Characteristic	No.	%*
Year of publication		
≤ 2010	7	25.9
≥ 2011	20	74.1
Country		
Canada	7	25.9
United States	6	22.2
European Union	6	22.2
United Kingdom	4	14.8
Japan	3	11.1
Other	1	3.7
Funding		
Genomic Health	15	55.6
Not related to Genomic Health	12	44.4
Time horizon, years		
< 25	7	25.9
≥ 25	20	64.1
LN status		
Negative	20	74.1
Positive	3	11.1
Both	4	14.8
Age, years		
≤ 55	11	40.7
≥ 55.1	10	37.0
Unclear	6	22.2
Clinicopathologic information considered		
Yes	9	33.3
No	18	66.7
Short-term chemotherapy adverse effects considered		
Yes	21	77.8
No	6	22.2
Long-term chemotherapy adverse effects considered		
Yes	9	33.3
No	18	66.7

*Percentages may not sum to 100 because of rounding.

Table 2. Concerning Issues in Existing ODX CEAs

ODX CEAs	Issue
Model structure	
Ignoring clinicopathologic information	Available risk classification models, such as Adjuvant! Online and PREDICT risk calculator, could help estimate risks of distant recurrence/breast cancer mortality; ignoring clinicopathologic information would make ODX more cost effective
Combining low-, intermediate-, and high-risk groups	If one risk group has an ICER > \$100,000 per QALY, and other groups have low ICERs, combining low-, intermediate-, and high-risk groups may not reveal the cost-ineffective group
Model assumptions	
ODX decreases chemotherapy use	Limited evidence suggests that ODX increases chemotherapy use among the clinicopathology-based low-risk group; assuming decrease in chemotherapy use would make ODX cost effective
Predictive value of ODX	Limited evidence supports this assumption, which would favor ODX cost effectiveness
Ignoring chemotherapy toxicity	Models that did not include short- or long-term adverse effects attributed to chemotherapy would favor ODX if ODX were to increase chemotherapy use but would be against ODX if ODX were to decrease chemotherapy use
Input parameter selection	
Not real-world RS distributions	Existing models generally used data based on 668 patients enrolled in the NSABP B-14 study; HER2 information is not available in this series, and the distributions are not population based; distributions of high-risk RS group would be overestimated, resulting in bias favoring ODX
Implausible estimates of chemotherapy effectiveness	Some studies selected parameters that chemotherapy increases distant recurrence for the RS low-risk group; these parameters are biologically implausible and would lead to bias favoring ODX
Young patient age	Some studies assumed that the age at breast cancer diagnosis is younger than that in the actual population, which could make ODX cost effective

Abbreviations: CEA, cost-effectiveness analysis; HER2, human epidermal growth factor receptor 2; ICER, incremental cost-effectiveness ratio; NSABP, National Surgical Adjuvant Breast and Bowel Project; ODX, Oncotype DX; QALY, quality-adjusted life-year; RS, recurrence score.

recommendation will vary according to patient characteristics. For instance, among patients in the AO low-risk group, an ODX result that indicates high risk might be quite unlikely, whereas among patients with a higher AO risk, the ODX score is likely to be higher risk. Hence, it is possible that ODX could be cost ineffective in one clinicopathology-based risk group but could be cost effective or cost saving in other groups. Only two studies reported ICERs separately by risk subgroup in their base-case analyses.^{14,16} In one report, the ICER of ODX varied substantially, ranging from \$22,330 per QALY for the AO low-risk group to \$1,111 per QALY for the AO high-risk group.¹⁶ A different analysis (Wald-government) reported that the ICER of ODX was £26,940 (equal to \$45,878) per QALY for all patients combined, but only £9,007 (equal to \$15,338)

Table 3. Quality Assessment of Model Structure, Model Assumptions, and Input Parameter Selection Across Analyses (N = 27)

First Author	Year	Model Structure*		Model Assumption†			Input Parameter Selection‡		
		Item 1	Item 2	Item 1	Item 2	Item 3	Item 1	Item 2	Item 3
Hornberger ¹⁰	2005	Yes	No§	Yes	No	No	No	Yes	No
Lyman ²⁵	2007	No	No	—	No	No	No	No	No
Kondo ²⁶	2008	No	No	—	No	No	No	Yes	No¶
Klang ²⁷	2010	No	No	—	No	No	No	No§	No¶
Tsoi ²⁸	2010	Yes	No	Yes	Yes	No	No	Yes	No¶
Hornberger ²⁹	2010	Yes	No	Unclear	No	No	No	Yes	No
OHTA ³⁸	2010	Yes	No	Unclear	No	No	No	No	No¶
Kondo ³⁰	2011	No	No	—	No	No	No	Yes	No¶
Vanderlaan ³¹	2011	No	No	—	No	Yes	No	Yes	Yes
Hornberger ³²	2011	No	No	—	No	Yes	No	No	No¶
Retel ³³	2011	No	No	—	Unclear	Yes	No	Unclear	No
Hal ³⁴	2012	No	No	—	No	Yes	No	No	Yes
Lamond ³⁵	2012	No	No	—	No	Yes	No	Yes	No¶
Blohmér ³⁶	2013	No	No	—	No	No	No	Yes	No¶
Reed ¹⁷	2013	Yes	No§	Yes	No	No	No	No	No¶
Hannouf ³⁷	2012	No	No	—	No	Yes	No	No	No¶
Davidson ³⁹	2013	No	No	—	No	No	No	Yes	No¶
Holt ⁴⁰	2013	No	No	—	No	No	No	Yes	Yes
Paulden ¹⁶	2013	Yes	Yes	No	No	No	No	No	No
Ward-I ¹⁴	2013	No	No	—	No	No	No	Yes	Yes
Ward-G ¹⁴	2013	Yes	Yes	No	No§	Yes	No	No§	No¶
Hannouf ⁴¹	2014	No	No	—	No	Yes	No	No	No¶
Yamauchi ⁴²	2014	No	No	—	No	No	No	Yes	No¶
Segui ⁴³	2014	Yes	No	Yes	No	No	No	Yes	Yes
Kip ⁴⁴	2015	No	No	—	No	No	No	Yes	Yes
Katz ⁴⁵	2015	No	No	—	No	No	No	No	No¶
Jahn ⁴⁶	2015	Yes	No	Yes	No	Yes	No	No	No¶

Abbreviations: OHTA, Ontario Health Technology Assessment; Ward-G, Ward-government; Ward-I, Ward-industry.

*Model structure: Item 1: Did the study consider clinicopathologic information in the Oncotype DX (ODX) and non-ODX arms? Item 2: Did the study report results by clinically low-, intermediate-, and high-risk groups separately?

†Model assumption: Item 1: Did the study assume that ODX increases chemotherapy use in the clinically low-risk group? Item 2: Did the study not assume that ODX has predictive value? Item 3: Did the study assume that chemotherapy has long-term toxicity?

‡Parameter selection: Item 1: Did the study use real-world ODX recurrence score (RS) distributions? Item 2: Did the study not assume that chemotherapy increases distant recurrence in patients whose RS is < 18? Item 3: Did the study use age ≥ 60 years as its starting age?

§The sensitivity analysis, but not the base-case analysis, addressed the concerning issue.

||Not applicable, because the analysis ignored clinicopathologic information in the model.

¶With some justification or varying age in its sensitivity analysis.

per QALY for patients with an NPI score > 3.4.¹⁴ Two other studies only reported the ICER of ODX across different risk subgroups in sensitivity analyses.^{10,17}

Model Assumptions

We identified several assumptions that were not supported by strong evidence or departed from commonly accepted knowledge. The assumptions could have an effect on the ICER of ODX.

ODX decreases chemotherapy use. At the population level, the decrease in chemotherapy use associated with ODX is believed to be the driver leading ODX to be cost effective or even cost saving.^{18,52} Although ODX may decrease overall chemotherapy use,⁹ evidence suggests that ODX increases chemotherapy use among patients who had clinicopathology-based low-risk breast cancer.^{53,54} For instance, one study examining 7,375 women with estrogen receptor-positive breast cancer at 17 comprehensive and community-based cancer centers showed that ODX increases chemotherapy use among the clinicopathology-based low-risk group.⁵³ However, only five of the 27 analyses we identified assumed an increase in chemotherapy use as a result of ODX in the clinically low-risk group.^{10,17,28,43,46} Therefore, the financial consequences of adopting ODX, when

estimated by a budget impact analysis,⁵⁵ would be underestimated in all except the five analyses.

Predictive value of ODX. Most studies assumed that the relative risk reduction (RRR) of distant recurrence attributed to chemotherapy varied across different RS strata. That is, the ODX results are predictive of a small RRR associated with chemotherapy in the RS low-risk group and a larger RRR for the RS high-risk group. Such an assumption makes ODX seem more cost effective, because ODX can differentiate the magnitude of chemotherapy benefits beyond traditional clinicopathologic information. However, evidence of the predictive value of ODX is limited to retrospective analyses of a subset of patients (n = 651) enrolled in the NSABP (National Surgical Adjuvant Breast and Bowel Project) B-20 trial (N = 2,306).⁶ In fact, the NICE Diagnostics Advisory Committee recommended that it was more appropriate to evaluate ICERs of ODX based on an assumption of equal RRR attributed to chemotherapy for all ODX risk categories.¹³ For instance, several studies assumed an RRR of 0 for the RS low-risk group (no benefits attributed to chemotherapy) and an RRR of 0.74 for the RS high-risk group.^{36,40,42,44} Only the Ward-government study conducted sensitivity analyses and found that assuming no predictive value has an ICER two times higher than that assuming a predictive value.¹⁴

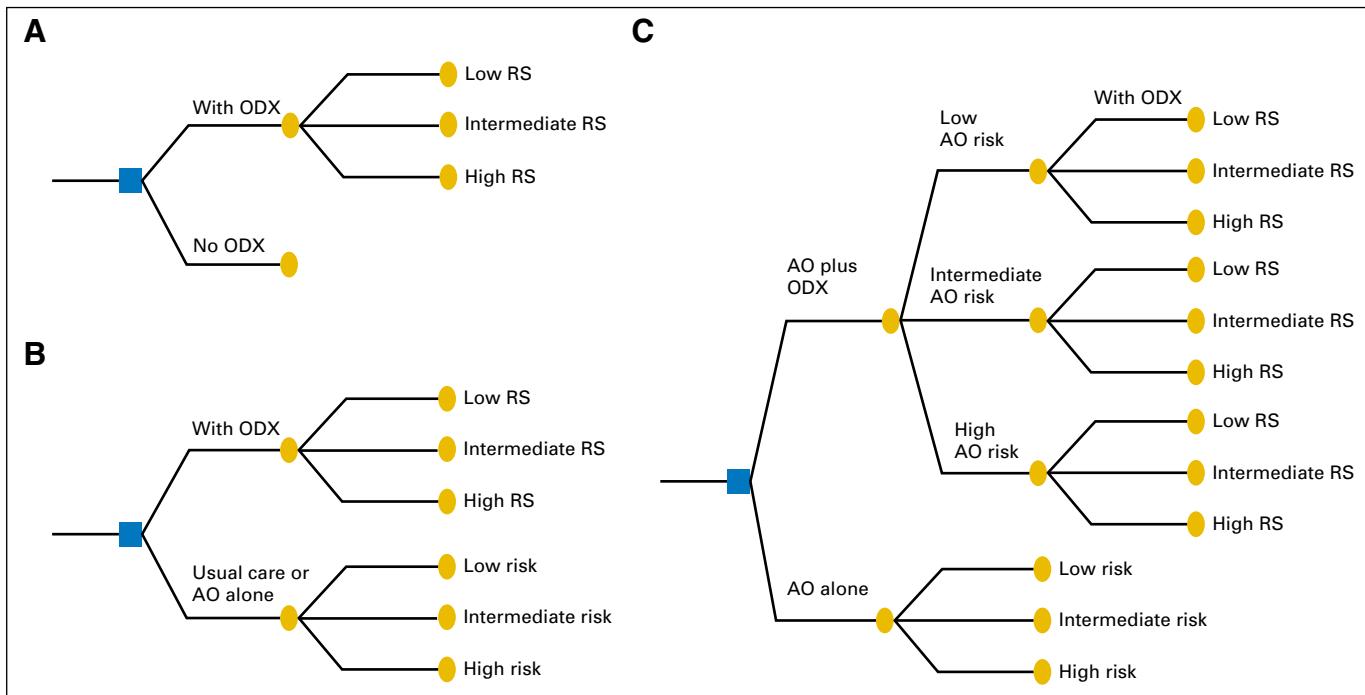


Fig 1. Model structure in 27 cost-effectiveness analyses this study identified: (A) did not consider clinicopathologic information at all ($n = 2$); (B) clinicopathologic information alone versus Oncotype DX (ODX) alone ($n = 16$); and (C) clinicopathologic information plus ODX versus clinicopathologic information alone ($n = 9$). AO, Adjuvant! Online; RS, recurrence score.

Chemotherapy toxicity. It is well known that chemotherapy has short- and long-term adverse effects.⁵⁶⁻⁵⁸ Although 21 studies considered short-term toxicity in their models, only nine reports incorporated long-term adverse effects. Few reports included both cardiotoxicity and hematologic toxicity. Furthermore, although several studies included taxanes in their chemotherapy regimens, none of them considered long-term neuropathy as an adverse effect in their models. Whether ignoring chemotherapy toxicity would bias a model to favor ODX depends on the model structure and whether ODX increases or decreases chemotherapy use.

Input Parameter Selection

We identified several input parameters—including the RS distributions, predictive value of ODX, and patient age—that could potentially bias the results.

Not real-world RS distributions. To appropriately assess the effect of ODX use, it is imperative to have estimates of RS distributions conditional on clinicopathology-based risk groups. This risk reclassification information enables investigators to estimate the proportion of patients who could benefit from ODX. Among the nine studies that used strata-specific data comparing ODX plus clinicopathologic information versus clinicopathologic information alone, all but the Ward-government study referred to the distributions from the NSABP B-14 trial.⁵⁹⁻⁶¹ These distribution estimates were inappropriate in two important ways. First, this trial, conducted in the 1990s when HER2 information was not routinely available, included a mix of patients with HER2-positive and -negative disease. More importantly, these estimates were not population based; they were derived from a randomized trial subset (668 of 2,299 patients) of patients whose cancer specimens were

sufficient for ODX testing. Therefore, patients with a small tumor (ie, low-risk participants) are likely to be excluded. Both would shift toward a higher RS, making ODX seem more cost effective for the clinicopathology-based low-risk group, because the number of patients in the RS high-risk group who could benefit was overestimated. These issues raise concerns about the generalizability of the corresponding CEA results to the modern population.

The RS distributions of the Ward-government study were derived from the UK TransATAC (Translational Substudy of Arimidex, Tamoxifen, Alone or in Combination) trial.¹⁴ This trial, although not population based, is contemporary and more clinically appropriate than the NSABP B-14 trial because it included HER2 information. The Ward-government study reclassified eligible patients of the TransATAC trial by NPI group; only 3.8% of the clinically NPI low-risk group were in the RS high-risk group; however, this was reported as 14.8% in the NSABP B-14 trial, reinforcing our concerns.

Implausible estimates of chemotherapy effectiveness. The estimated risk reduction attributed to chemotherapy among the RS low-risk group was problematic in several CEAs. For instance, 11 analyses cited a study that reported the relative risk of distant recurrence after chemotherapy as 1.31 (95% CI, 0.46 to 3.78) for the RS low-risk group.^{14,16,17,25,27,32,35,37,38,45,46} Because the 95% CI crossed 1, the primary study found no significant effect of chemotherapy for the RS low-risk group. Studies that simply applied a relative risk of 1.31 would assume that chemotherapy increases distant recurrence for the RS low-risk group, which is biologically implausible and would lead to bias favoring ODX.

Young patient age. When conducting a CEA, investigators select a starting age at which a hypothetical group of women is deciding on ODX use. Because younger patients have longer life expectancy, selecting a younger starting age would increase the

QALY gained and would thus favor ODX cost effectiveness. For instance, the median age of patients with breast cancer in the United States is 62 years⁴¹, however, only one of the six US studies in our sample used 62 years as the starting age in its base-case analysis.³¹ Of the 27 studies identified, 11 assumed that all patients were age \leq 55 years, and six studies did not specify the starting age. Only 10 studies assumed a starting age $>$ 55 years. Sixteen studies conducted sensitivity analyses and generally found a substantial effect of starting age on ODX cost effectiveness. For instance, the Ward-government study reported the ICER of ODX for women age 70 years to be two times higher than that for women age 50 years.¹⁴

Impact of Potential Sources of Bias on ICER Results

Our simulation found that the problematic issues we identified do not change the conclusion that ODX is cost effective for the clinically intermediate- or high-risk subgroup, using the willing-to-pay threshold of \$50,000 per QALY (Table 4). However, the problematic issues have great impact on ICERs of the clinically low-risk subgroup, particularly favoring ODX as being cost effective. Among these issues, the assumption that ODX decreases chemotherapy use, the non-population-based RS distributions, and the assumption that ODX has predictive value seemed to have the largest effect on ICER estimates. For instance, incorrectly assuming that ODX decreases chemotherapy use, our simulation indicated that the ICER of ODX would decrease from \$167,600 per QALY to \$39,000 per QALY in the clinically low-risk group. Similarly, when we used data derived from the NSABP B-14 study instead of population-based estimates, the ICER of ODX was \$42,000 per QALY in the low-risk group.

Funding Source

Among the 27 studies, 15 (55.6%) were funded directly by or associated with Genomic Health. We did not find any significant difference in terms of funding source and timing of publication (Appendix Fig A3, online only). The base-case results in industry-funded studies were associated with more favorable ICERs than non-industry-funded studies (Fig 2A). Compared with non-industry-funded studies, industry-funded studies tended to have a lower incremental cost (2015 US\$900 v \$3,100; $P < .001$) but

a similar incremental QALY gained ($P = .10$). The US studies had a mean ICER of -4,300 per QALY (cost saving), whereas the non-US studies had a mean ICER of 15,100 per QALY ($P = .04$; Fig 2B). Notably, all five US studies were associated with Genomic Health. Furthermore, the only two studies that had an ICER $>$ \$50,000 per QALY were not funded by Genomic Health,^{28,37} whereas all five studies that reported ODX as being cost saving were funded by Genomic Health.^{10,29,31,32,45} However, we did not find any significant association between the funding source and each concerning issue (Appendix Table A4, online only). We acknowledge these analyses might be underpowered, given the small number of studies.

DISCUSSION

The results of this structure review indicate that a majority of existing CEAAs of ODX for women with early-stage breast cancer have problematic issues that may result in misleading conclusions. Furthermore, more than half of the 27 articles published were funded by Genomic Health. These industry-funded studies tended to report lower incremental cost and similar incremental QALY gained, favoring ODX cost effectiveness. Furthermore, many of these studies used similar methodologies with little effort to improve the quality of their modeling, raising questions about the value of these studies in advancing our understanding of ODX cost effectiveness.

Prior systematic reviews of CEAAs of ODX generally concluded that ODX is cost effective.^{14,18,19} However, these reviews rarely investigated the methodology used in these CEAAs. For instance, one review assessed the quality of the included CEAAs using the QHES instrument.¹⁸ The reviewers concluded that the quality ratings in existing CEAAs were high, although they acknowledged that a generic instrument such as QHES might have shortcomings for the assessment of CEAAs on diagnostic tests. Although the NICE review appraised the methodology the CEAAs used,¹⁴ it was limited in scope, because it only reviewed two ODX analyses; it excluded CEAAs in which the comparators were not applicable to the UK context. To the best of our knowledge, our systematic review is the first to focus on identifying specific methodologic concerns pertaining to CEA studies of genetic risk classification schema. Our

Table 4. Impact of Concerning Issues on ICER

Issue	ICER of ODX (\$/QALY)		
	Clinically Low-Risk Group	Clinically Intermediate-Risk Group	Clinically High-Risk Group
Benchmark estimate (age, 60 years)*	167,600	37,800	20,400
Model structure			
Ignoring clinicopathologic information†	—	—	—
Model assumption			
Decrease in chemotherapy use in clinically low-risk group	39,000	NA	NA
If assuming no predictive value of ODX*	307,200	43,400	22,700
Ignoring long-term chemotherapy toxicity	144,300	31,200	17,300
Parameter selection			
Not real-world population-based RS distributions	42,000	28,200	25,800
Increasing distant recurrence by chemotherapy if RS $<$ 18	123,700	34,500	19,000
Young patient age (50 years)	122,400	27,500	14,600

Abbreviations: ICER, incremental cost-effectiveness ratio; NA, not applicable; ODX, Oncotype DX; QALY, quality-adjusted life year; RS, recurrence score.

*We conservatively assumed that chemotherapy does not decrease distant recurrence if RS $<$ 18 and has a similar effect in decreasing distant recurrence if RS \geq 18, an assumption favoring ODX cost effectiveness.

†Analyses in which model structure ignored clinicopathologic information cannot provide ICER estimates for each risk group separately.

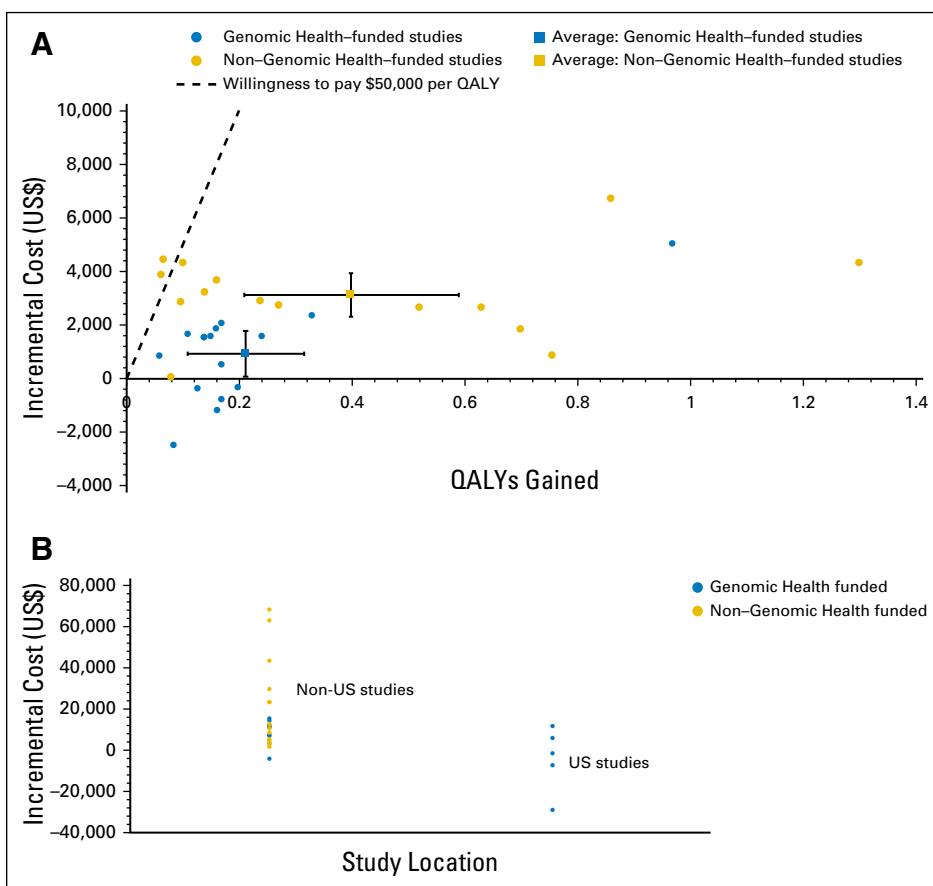


Fig 2. Incremental cost-effectiveness estimates of Oncotype DX (ODX; node-negative breast cancer) (A) according to industry relationship and (B) in US versus non-US analyses. (For comparison purposes, we chose only studies that reported ODX cost effectiveness for estrogen receptor-positive, node-negative breast cancer in their base-case analyses.)

approach could be used for systematic reviews on the economic evaluation of other genetic and biomarker tests. Specifically, by combining a systematic review with simulation modeling, we were able to demonstrate that the ICER estimates in prior CEs were generally underestimated. Three factors—the assumption that ODX has predictive value, the assumption that ODX decreases chemotherapy use across all clinical risk groups, and the non-population-based RS distributions—seemed to have the largest effect on ICER estimates. Recent publications^{53,54} and this structure review provided the updated evidence about the last two factors, and future research examining the predictive value of ODX is warranted.

We also created a conceptual framework to assess the validity of studies examining ODX cost effectiveness. Our study highlighted three key components: model structure, model assumption, and input parameter selection. First, model structure in a CEA of new genetic testing is crucial. It represents the choice of strategies for the decision problem; thus, it should reflect clinical practice. Medical oncologists use clinicopathologic information to predict prognosis and then use individual prognosis to make chemotherapy recommendations. To conduct an accurate economic evaluation of biomarkers in oncology, existing prognostic information routinely used in clinical practice must be incorporated into the model. Model structure that ignores clinicopathologic information does not appropriately address the complexity of the diagnostic testing decision, is unable to report ICERs across different risk subgroups, and would favor new testing as being cost effective. We acknowledge that our review did not account for the fact that oncologists might consider other prognostic factors,

such as mammography-detected lesions or the Ki67 biomarker, in chemotherapy decision making. Future assessment is warranted.

Second, although CEs generally described their assumptions, evaluating the validity of these assumptions requires clinical knowledge, accurate data, and the ability to assess the effect of different assumptions on the resulting CEA inferences. Third, professional judgment is required in reviewing the selection of input parameters. For instance, we were surprised that approximately 40% of the 27 analyses assumed that chemotherapy (compared with no chemotherapy) increases distant recurrence for patients whose RS < 18, and yet, only two studies explicitly conducted one-way sensitivity analyses to test this assumption.^{14,27} Because systematic reviews are a key tool for evidence-based medicine, researchers conducting systematic reviews need to critically appraise the methodology in the CEs they identify.

Our results also suggested a need for QHES revision. For instance, the QHES instrument used yes or no responses rather than a continuous scale for all 16 criteria (one to eight points per criterion; 100 points total).^{23,24} When CEs partially met a criterion, there was no standard to convey the quality score associated with the criterion. Furthermore, the QHES instrument recommends estimates derived from randomized controlled trials. Although randomized controlled trials provide the best estimates regarding the effect size of an intervention, population-based studies would be preferable for prevalence information.

CEA results have important policy implications. Governments outside of the United States have used CEA results to help with coverage and reimbursement decision making. The CEA results of

ODX differed because of differences not only in health care costs and practice patterns but also model structure, assumptions, and input parameter selection. For instance, findings from Ontario, Canada, indicated that ODX is cost effective for all patients with estrogen receptor-positive, HER2-negative, lymph node-negative breast cancer, and ODX is covered for this population.³⁸ In contrast, the UK NICE concluded that ODX was not cost effective for the general population but had an acceptable ICER for patients at a higher risk of distant metastasis, if ODX was provided by the manufacturer at a revised price (undisclosed).^{14,44} For instance, one third of eligible patients in the United Kingdom had NPI score > 3.4 (5-year survival, < 70%), and NICE considered ODX to be cost effective only for this high-risk population. Our study suggests that ODX is cost effective for the clinically intermediate- or high-risk subgroup but not for the clinically low-risk subgroup, which could help to inform ODX coverage or reimbursement decision making.

Our study has some limitations. First, our search strategy was limited to one database, and we included only studies written in English. Any publication bias would limit our ability to review unpublished work, particularly those suggesting ODX is not cost effective. However, the purpose of this review was to identify the weaknesses of existing CEAAs rather than to estimate the ICER for ODX. Although we have conceptualized the model structure of each study, variation in comprehensiveness of information provided in the existing studies precludes actual replication of these analyses. Efforts to share simulation programs among researchers could facilitate CEA transparency in the future. Finally, we acknowledge that the correlation between funding source and CEA results does not imply causality.

In conclusion, a majority of published CEAAs have serious methodologic weaknesses, resulting in potentially inaccurate or

biased conclusions. Several efforts to set standards for CEAAs have been improving the quality of economic evaluations.^{62,63} Some standards, such as model transparency and uncertainty evaluation, should be applied directly to the CEAAs of genetic diagnostic testing. However, additional scrutiny of other standards is warranted. Specifically, to adequately address decision problems regarding diagnostic testing, the scope and structure of the model should account for available prognostic information. Building on the QHES instrument, we have identified several concerning issues in existing economic evaluations of ODX. Because CEAAs play a pivotal role in value-based practice and resource allocation decisions, future CEAAs of these genetic tests should include the additional factors we have identified.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at jco.org.

AUTHOR CONTRIBUTIONS

Conception and design: Shi-Yi Wang, Cary P. Gross

Administrative support: Cary P. Gross

Collection and assembly of data: Shi-Yi Wang, Weixiong Dang, Ilana Richman

Data analysis and interpretation: All authors

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

REFERENCES

- Marchionni L, Wilson RF, Wolff AC, et al: Systematic review: gene expression profiling assays in early-stage breast cancer. *Ann Intern Med* 148: 358-369, 2008
- Yang Y, Muzny DM, Reid JG, et al: Clinical whole-exome sequencing for the diagnosis of Mendelian disorders. *N Engl J Med* 369:1502-1511, 2013
- Paik S, Shak S, Tang G, et al: A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 351:2817-2826, 2004
- Hayes DF: Targeting adjuvant chemotherapy: A good idea that needs to be proven! *J Clin Oncol* 30: 1264-1267, 2012
- Hassett MJ, O'Malley AJ, Pakes JR, et al: Frequency and cost of chemotherapy-related serious adverse effects in a population sample of women with breast cancer. *J Natl Cancer Inst* 98:1108-1117, 2006
- Paik S, Tang G, Shak S, et al: Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 24:3726-3734, 2006
- Dowsett M, Cuzick J, Ware C, et al: Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: A TransATAC study. *J Clin Oncol* 28:1829-1834, 2010
- Sparano JA, Gray RJ, Makower DF, et al: Prospective validation of a 21-gene expression assay in breast cancer. *N Engl J Med* 373:2005-2014, 2015
- Hornberger J, Alvarado MD, Rebecca C, et al: Clinical validity/utility, change in practice patterns, and economic implications of risk stratifiers to predict outcomes for early-stage breast cancer: A systematic review. *J Natl Cancer Inst* 104:1068-1079, 2012
- Hornberger J, Cosler LE, Lyman GH: 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, 2005
- Centers for Medicare and Medicaid Services: License for use of current procedural terminology (ed 4). <https://www.cms.gov/apps/ama/license.asp?file=/ClinicalLabFeeSched/downloads/2015-CLFS-Gapfill-Final-Determinations.zip>
- National Comprehensive Cancer Network: NCCN clinical practice guidelines in oncology. <http://www.nccn.org/patients>
- UK National Institute for Health and Care Excellence: Gene expression profiling and expanded immunohistochemistry tests for guiding adjuvant chemotherapy decisions in early breast cancer management: Mammaprint, Oncotype DX, IHC4 and Mammastrat. <https://www.nice.org.uk/guidance/dg10/resources/gene-expression-profiling-and-expanded-immunohistochemistry-tests-for-guiding-adjuvant-chemotherapy-decisions-in-early-breast-cancer-management-mammaprint-oncotyedihc4-and-mammostrat-pdf-1053623071429>
- Ward S, Scope A, Rafia R, et al: Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in breast cancer management: A systematic review and cost-effectiveness analysis. *Health Technol Assess* 17:1-302, 2013
- Lee AH, Ellis IO: The Nottingham prognostic index for invasive carcinoma of the breast. *Pathol Oncol Res* 14:113-115, 2008
- Paudlen M, Franek J, Pham B, et al: Cost-effectiveness of the 21-gene assay for guiding adjuvant chemotherapy decisions in early breast cancer. *Value Health* 16:729-739, 2013
- Reed SD, Dinan MA, Schulman KA, et al: Cost-effectiveness of the 21-gene recurrence score assay in the context of multifactorial decision making to guide chemotherapy for early-stage breast cancer. *Genet Med* 15:203-211, 2013
- Rouzier R, Pronzato P, Chéreau E, et al: Multigene assays and molecular markers in breast cancer: Systematic review of health economic analyses. *Breast Cancer Res Treat* 139:621-637, 2013
- Marrone M, Stewart A, Dotson WD: Clinical utility of gene-expression profiling in women with early breast cancer: An overview of systematic reviews. *Genet Med* 17:519-532, 2015
- Friedberg M, Saffran B, Stinson TJ, et al: Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA* 282:1453-1457, 1999
- Bell CM, Urbach DR, Ray JG, et al: Bias in published cost effectiveness studies: Systematic review. *BMJ* 332:699-703, 2006
- Valachis A, Polyzos NP, Nearchou A, et al: Financial relationships in economic analyses of targeted therapies in oncology. *J Clin Oncol* 30:1316-1320, 2012
- Ofman JJ, Sullivan SD, Neumann PJ, et al: Examining the value and quality of health economic

- analyses: Implications of utilizing the QHES. *J Manag Care Pharm* 9:53-61, 2003
- 24.** Chiou CF, Hay JW, Wallace JF, et al: Development and validation of a grading system for the quality of cost-effectiveness studies. *Med Care* 41:32-44, 2003
- 25.** Lyman GH, Cosler LE, Kuderer NM, et al: 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, 2007
- 26.** Kondo M, Hoshi SL, Ishiguro H, et al: 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, 2008
- 27.** Klang SH, Hammerman A, Liebermann N, et al: Economic implications of 21-gene breast cancer risk assay from the perspective of an Israeli-managed health-care organization. *Value Health* 13:381-387, 2010
- 28.** Tsui DT, Inoue M, Kelly CM, et al: Cost-effectiveness analysis of recurrence score-guided treatment using a 21-gene assay in early breast cancer. *Oncologist* 15:457-465, 2010
- 29.** Hornberger J, Lyman GH, Chien R: Economic implications of 21-gene recurrence score assay: US multicenter experience. *J Clin Oncol* 28:e382; author reply e383, 2010
- 30.** Kondo M, Hoshi SL, Yamanaka T, et al: Economic evaluation of the 21-gene signature (Oncotype DX) in lymph node-negative/positive, hormone receptor-positive early-stage breast cancer based on Japanese validation study (JBCRG-TR03). *Breast Cancer Res Treat* 127:739-749, 2011
- 31.** Vanderlaan BF, Broder MS, Chang EY, et al: Cost-effectiveness of 21-gene assay in node-positive, early-stage breast cancer. *Am J Manag Care* 17:455-464, 2011
- 32.** Hornberger J, Chien R, Krebs K, et al: US insurance program's experience with a multigene assay for early-stage breast cancer. *J Oncol Pract* 7:e38s-e45s, 2011
- 33.** Retèl VP, Joore MA, van Harten WH: Head-to-head comparison of the 70-gene signature versus the 21-gene assay: Cost-effectiveness and the effect of compliance. *Breast Cancer Res Treat* 131:627-636, 2012
- 34.** Hall PS, McCabe C, Stein RC, et al: Economic evaluation of genomic test-directed chemotherapy for early-stage lymph node-positive breast cancer. *J Natl Cancer Inst* 104:56-66, 2012
- 35.** Lamond NW, Skedgel C, Rayson D, et al: Cost-utility of the 21-gene recurrence score assay in node-negative and node-positive breast cancer. *Breast Cancer Res Treat* 133:1115-1123, 2012
- 36.** Blohmer JU, Rezai M, Kümmel S, et al: Using the 21-gene assay to guide adjuvant chemotherapy decision-making in early-stage breast cancer: A cost-effectiveness evaluation in the German setting. *J Med Econ* 16:30-40, 2013
- 37.** Hannouf MB, Xie B, Brackstone M, et al: Cost-effectiveness of a 21-gene recurrence score assay versus Canadian clinical practice in women with early-stage estrogen- or progesterone-receptor-positive, axillary lymph-node negative breast cancer. *BMC Cancer* 12:447, 2012
- 38.** Health Quality Ontario: Gene expression profiling for guiding adjuvant chemotherapy decisions in women with early breast cancer: An evidence-based and economic analysis. *Ont Health Technol Assess Ser* 10:1-57, 2010
- 39.** Davidson JA, Cromwell I, Ellard SL, et al: A prospective clinical utility and pharmacoeconomic study of the impact of the 21-gene Recurrence Score® assay in oestrogen receptor positive node negative breast cancer. *Eur J Cancer* 49:2469-2475, 2013
- 40.** Holt S, Bertelli G, Humphreys I, et al: A decision impact, decision conflict and economic assessment of routine Oncotype DX testing of 146 women with node-negative or pN1mi, ER-positive breast cancer in the U.K. *Br J Cancer* 108:2250-2258, 2013
- 41.** Howlader N, Noone AM, Krapcho M, et al: SEER Cancer Statistics Review, 1975-2014. Table 1.12. National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2014/, 2017
- 42.** Yamauchi H, Nakagawa C, Yamashige S, et al: Societal cost-effectiveness analysis of the 21-gene assay in estrogen-receptor-positive, lymph-node-negative early-stage breast cancer in Japan. *BMC Health Serv Res* 14:372, 2014
- 43.** Seguí MA, Crespo C, Cortés J, et al: Genomic profile of breast cancer: Cost-effectiveness analysis from the Spanish National Healthcare System perspective. *Expert Rev Pharmacoecon Outcomes Res* 14:889-899, 2014
- 44.** Kip M, Monteban H, Steuten L: Long-term cost-effectiveness of Oncotype DX® versus current clinical practice from a Dutch cost perspective. *J Comp Eff Res* 4:433-445, 2015
- 45.** Katz G, Romano O, Foa C, et al: Economic impact of gene expression profiling in patients with early-stage breast cancer in France. *PLoS One* 10:e0128880, 2015
- 46.** Jahn B, Rochau U, Kurzthaler C, et al: Cost effectiveness of personalized treatment in women with early breast cancer: The application of OncotypeDX and Adjuvant! Online to guide adjuvant chemotherapy in Austria. *Springerplus* 4:752, 2015
- 47.** Ravdin PM, Siminoff LA, Davis GJ, et al: Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 19:980-991, 2001
- 48.** PREDICT: PREDICT tool version 2.0: Breast cancer overall survival. http://www.predict.nhs.uk/predict_v2.0.html
- 49.** Agarwal V, O'Neill P: Adjuvant! Online as a decision-making tool in early breast cancer: A UK national survey. *Clin Oncol (R Coll Radiol)* 23:159-160, 2011
- 50.** Engelhardt EG, Pieterse AH, van Duijn-Bakker N, et al: Breast cancer specialists' views on and use of risk prediction models in clinical practice: A mixed methods approach. *Acta Oncol* 54:361-367, 2015
- 51.** Love N: Management of breast cancer in the adjuvant and metastatic settings. *Patterns of Care in Medical Oncology* 2:11-24, 2005
- 52.** Lamond NW, Skedgel C, Younis T: Is the 21-gene recurrence score a cost-effective assay in endocrine-sensitive node-negative breast cancer? *Expert Rev Pharmacoecon Outcomes Res* 13:243-250, 2013
- 53.** Hassett MJ, Silver SM, Hughes ME, et al: Adoption of gene expression profile testing and association with use of chemotherapy among women with breast cancer. *J Clin Oncol* 30:2218-2226, 2012
- 54.** Dinan MA, Mi X, Reed SD, et al: Association between use of the 21-gene recurrence score assay and receipt of chemotherapy among Medicare beneficiaries with early-stage breast cancer, 2005-2009. *JAMA Oncol* 1:1098-1109, 2015
- 55.** Sullivan SD, Mauskopf JA, Augustovski F, et al: Budget impact analysis-principles of good practice: Report of the ISPOR 2012 Budget Impact Analysis Good Practice II Task Force. *Value Health* 17:5-14, 2014
- 56.** Crump M, Tu D, Shepherd L, et al: Risk of acute leukemia following epirubicin-based adjuvant chemotherapy: A report from the National Cancer Institute of Canada Clinical Trials Group. *J Clin Oncol* 21:3066-3071, 2003
- 57.** Calip GS, Malmgren JA, Lee WJ, et al: Myelodysplastic syndrome and acute myeloid leukemia following adjuvant chemotherapy with and without granulocyte colony-stimulating factors for breast cancer. *Breast Cancer Res Treat* 154:133-143, 2015
- 58.** Bowles EJ, Wellman R, Feigelson HS, et al: Risk of heart failure in breast cancer patients after anthracycline and trastuzumab treatment: A retrospective cohort study. *J Natl Cancer Inst* 104:1293-1305, 2012
- 59.** Paik S, Shak S, Tang G: Risk classification of breast cancer patients by the recurrence score assay: Comparison to guidelines based on patient age, tumor size, and tumor grade. Presented at the 28th Annual San Antonio Breast Cancer Symposium, San Antonio, TX, December 8-11, 2005
- 60.** Bryant J: Toward a more rational selection of tailored adjuvant therapy. Presented at the 9th International St. Gallen Oncology Conference on Primary Therapy of Early Breast Cancer, St Gallen, Switzerland, January 26-29, 2005
- 61.** Tang G, Shak S, Paik S, et al: Comparison of the prognostic and predictive utilities of the 21-gene recurrence score assay and Adjuvant! for women with node-negative, ER-positive breast cancer: Results from NSABP B-14 and NSABP B-20. *Breast Cancer Res Treat* 127:133-142, 2011
- 62.** Sanders GD, Neumann PJ, Basu A, et al: Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: Second panel on cost-effectiveness in health and medicine. *JAMA* 316:1093-1103, 2016
- 63.** Caro JJ, Briggs AH, Siebert U, et al: Modeling good research practices: Overview—A report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. *Med Decis Making* 32:667-677, 2012

Affiliations

Shi-Yi Wang and Weixiong Dang, Yale University School of Public Health; **Shi-Yi Wang, Ilana Richman, Sarah S. Mougalian, Suzanne B. Evans, and Cary P. Gross**, Yale Cancer Center; and **Shi-Yi Wang, Ilana Richman, Sarah S. Mougalian, Suzanne B. Evans, and Cary P. Gross**, Yale University School of Medicine, New Haven, CT.

Support

Supported by Research Scholar Grant No. RSG-14-029-01-CPHPS from the American Cancer Society (C.P.G.) and a P30 Cancer Center Support Grant from Yale Comprehensive Cancer Center through National Cancer Institute (Grant No. P30 CA01635933).



AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST**Cost-Effectiveness Analyses of the 21-Gene Assay in Breast Cancer: Systematic Review and Critical Appraisal**

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/jco/site/ifc.

Shi-Yi Wang**Research Funding:** Genentech**Weixiong Dang**

No relationship to disclose

Ilana Richman

No relationship to disclose

Sarah S. Moughalian**Stock or Other Ownership:** Gilead Sciences, Coronado Biosciences, Roche**Consulting or Advisory Role:** Eisai, Cardinal Health**Research Funding:** Genentech, Pfizer**Suzanne B. Evans****Other Relationship:** 21st Century Oncology**Cary P. Gross****Research Funding:** 21st Century Oncology, Johnson & Johnson, Pfizer**Travel, Accommodations, Expenses:** Flatiron Health

Acknowledgment

We thank David Paltiel for his comments on the early version of this manuscript.

Appendix

Nottingham Prognostic Index

Nottingham Prognostic Index = tumor size \times 0.2 + lymph node stage + histologic grade, where tumor size is the index lesion in centimeters, lymph node stage is nodal status (zero nodes, 1; one to three nodes, 2; and more than three nodes, 3), and histologic grade is grade of tumor (grade I, 1; grade II, 2; and grade III, 3).

Simulation Study to Evaluate Impact of Concerning Issues on ICER Results

We combined classic systematic reviews with simulation modeling to evaluate the magnitude of how the model structure, model assumptions, and input parameters of prior cost-effectiveness analyses (CEAs) might influence results. To evaluate the impact of concerning issues on the incremental cost-effectiveness ratios (ICERs), we first built a simulation model that adopted the National Institute for Health and Care Excellence (NICE) analytic model¹⁴ to examine Oncotype DX (ODX) cost effectiveness (details regarding model assumptions and parameters of the base-case scenario provided in Methodology of Base-Case Analyses). The ICERs estimated from the base-case analyses served as the benchmark. We then conducted a series of simulations where each simulation examined one concerning issue we had identified from our systematic review (as described in Evaluation of Impact of Concerning Issues on ICER Results). The corresponding ICER estimates were then compared with the benchmark estimates; the comparisons provided insight into the extent to which the concerning issues might bias the results.

Methodology of Base-Case Analyses

Overview. We developed state transition models to compare expected costs and quality-adjusted life-years (QALYs) gained over a lifetime horizon between clinicopathologic information alone (risk classification provided by the PREDICT tool⁴⁸) versus clinicopathologic information plus ODX. The population assessed was women with hormone receptor-positive, human epidermal growth factor receptor 2-negative, lymph node-negative early breast cancer. We assumed that oncologists had information about age, tumor grade, and tumor size and used the PREDICT online tool to categorize patients into low-, intermediate-, and high-risk groups (defined as per PREDICT; 10-year breast cancer mortality reduction by chemotherapy as < 3%, 3% to 5%, or > 5%, respectively).⁴⁸ If ODX was provided, patients were further categorized into low, intermediate, and high recurrence score (RS) subgroups, and we allowed either the low RS group or the low and intermediate RS groups to spare chemotherapy. If ODX testing was not provided, we assumed that patients, either all or none, were receiving chemotherapy. The RS distributions, conditional on three PREDICT risk groups, were derived from the Connecticut Tumor Registry. The starting age of the cohort was 60 years, given that the median age at diagnosis of breast cancer in the United States is 62 years (https://seer.cancer.gov/csr/1975_2014/browse_csr.php?sectionSEL=1&pageSEL=sect_01_table.12.html). The model adopted the perspective of the US payer. Costs and QALYs were discounted at 5% annually. Detailed information regarding all model parameters is listed in Appendix Table A2.

Proportion of patients assigned to RS risk category. The proportion of patients assigned to the three RS categories, conditional on the three PREDICT risk groups, was estimated from a population-based study. Among women with estrogen receptor-positive, human epidermal growth factor receptor 2-negative, lymph node-negative early-stage breast cancer diagnosed in the years 2011 to 2013 in Connecticut, 2,245 patients received ODX testing with RS information. We categorized them into five 10-year age groups (30-39, 40-49, 50-59, 60-69, and \geq 70 years), three tumor grade groups (1, 2, and 3), and four tumor size groups (0 to 10, 11 to 20, 21 to 30, and > 30 mm). The 10-year breast cancer mortality reduction attributed to anthracycline plus taxane-based chemotherapy for these 60 categories was abstracted from the PREDICT Web site. Patients were grouped by the expected benefit of chemotherapy, as per PREDICT. For each PREDICT risk group, we then calculated the proportion of patients assigned to RS risk category.

Benefits and harms of chemotherapy. Among our cohort of 2,245 patients, the average 10-year breast cancer mortality reduction from chemotherapy for PREDICT low-, intermediate-, and high-risk groups was 1.32%, 3.75%, and 6.11%, respectively. Using our transition model, we estimated the 10-year distant recurrence rate without chemotherapy to be 4.6% in the PREDICT low-risk group, 14.8% in the PREDICT intermediate-risk group, and 28.4% in the PREDICT high-risk group. There were no published studies reporting distant recurrence without receiving chemotherapy conditional upon ODX plus PREDICT risk. One study simulated a subset of the NSABP (National Surgical Adjuvant Breast and Bowel Project) B-14 study⁶ and estimated 10-year distant

recurrence rates for patients without receiving chemotherapy, conditional on Adjuvant! Online and RS score.¹⁶ Assuming similar magnitudes in predicting distant recurrence across RS groups each stratum, we calibrated the 10-year distant recurrence rates for patients without receiving chemotherapy, according to the RS distributions in our three PREDICT risk groups.

If patients received chemotherapy, we assumed they would receive AC-T (four cycles of doxorubicin and cyclophosphamide followed by 12 cycles of paclitaxel, once every week). For patients who did not receive ODX testing, we assumed that the relative risk of distant recurrence attributed to AC-T was 0.58 compared with those receiving no adjuvant chemotherapy (Early Breast Cancer Trialists' Collaborative Group (EBCTCG): Lancet 379:432-44, 2012). For patients who received ODX testing, we allowed the relative risk reduction to be greater and identical in the intermediate and high RS groups, with no risk reduction in the low RS group.⁶ This approach assumed that ODX has some prediction value but did not assume that chemotherapy increases distant recurrence among patients whose RS < 18.

Chemotherapy can reduce distant breast cancer recurrence but may lead to short- and long-term adverse effects. Following the model developed by researchers in Canada,¹⁶ we assumed a proportion of patients undergoing chemotherapy would have adverse effects resulting from toxicity, which may require hospitalization or lead to death. We assumed patients who received chemotherapy would have an excess risk of congestive heart failure (CHF); the annual probability was 0.2%.⁵⁸ CHF annual mortality was derived from the Seattle model, which reported 3-year survival of 68% from six studies (Levy WC, et al: Circulation 113:1424-1433, 2006). The 8-year cumulative probability of acute myeloid leukemia (AML) or myelodysplastic syndrome after chemotherapy was 0.37%, based on prior meta-analysis (Praga C, et al: J Clin Oncol 23:4179-4191, 2005). The medium survival for patients who developed AML was assumed to be 8 months (Edlin R, et al: Health Technol Assess 14:69-74, 2010). We assumed that patients in distant recurrence state would not transition to CHF state, nor would patients in AML state transition to states of CHF and distant recurrence.

Other probability parameters. Breast cancer annual mortality rate was 24.3%, based on the 5-year survival for patients age > 50 years with advanced-stage breast cancer (https://seer.cancer.gov/archive/csr/1975_2005/results_merged/sect_04_breast.pdf). Annual mortality rate resulting from other causes was based on the 2012 US life table (https://www.cdc.gov/nchs/data/nvsr/nvsr65/nvsr65_08.pdf). Short-term adverse effects because of chemotherapy were based on prior literature and were used for cost calculation. We included short-term mortality resulting from chemotherapy toxicity, assuming a 0.35% mortality rate in the first year if receiving chemotherapy (Goldhirsch A, et al: N Engl J Med 320:491-496, 1989). Following the model developed by researchers in the United Kingdom,¹⁴ we assumed the risk of distant recurrence was constant in the first 10 years, was halved between 10 and 15 years, and was zero after 15 years; 10.5% of distant recurrences were preceded by a local recurrence, and only the local recurrence that led to a distant recurrence was included in the model (both cost and QALY decrement); 25% of patients have an echocardiogram before starting chemotherapy; and 25% of all patients treated with chemotherapy receive granulocyte colony-stimulating factor for the secondary prevention of febrile neutropenia.

Costs. The cost of performing an ODX was assumed to be \$3,416.¹¹ AC-T cost \$4,343 per patient (Nadeem H, et al: J Oncol Pract 12:e299-e307, e251, 2016). Granulocyte colony-stimulating factor was assumed to consist of pegfilgrastim (cost of \$3,773) for two cycles (<https://www.goodrx.com/neupogen>). Ongoing care cost for recurrence-free state was time dependent, derived from prior literature (Will BP, et al: Eur J Cancer 36:724-735, 2000). Costs for treating short-term adverse effects among patients who received chemotherapy were also based on prior literature.¹⁶ Initial cost for treating distant recurrence was assumed to be \$8,710, and ongoing care for this state was \$748 per month (Wong FL, et al: Ann Intern Med 160:672-683, 2014; Mandelblatt JS, et al: J Gen Intern Med 20:487-496, 2005). We assumed 10.5% of distant recurrences had local recurrence, and local recurrence cost \$24,045.¹⁴ Treatment cost for CHF was \$19,567 (Dunlay SM, et al: Circ Cardiovasc Qual Outcomes 4:68-75, 2011), and ongoing care for this state was \$523 per month (Wong FL, et al: Ann Intern Med 160:672-683, 2014). Except AML, cost for the last 1 year of life was \$54,220 for death resulting from CHF, \$57,690 for death resulting from breast cancer, and \$53,040 for death resulting from other causes (Levinsky NG, et al: JAMA 286:1349-1355, 2001; Reed SD, et al: Am J Cardiol 110:1150-1155, 2012). We assumed lifetime cost for AML was \$19,567.¹⁴ All costs were measured in 2015 US dollars.

Utilities. Utility weights were age adjusted at 5-year increments using previously reported trends (Fryback DG, et al: Med Decis Making 17:276-284, 1997; Stout NK, et al: J Natl Cancer Inst 98:774-782, 2006). Quality-of-life utility weights were identified from prior literature. Utility values for patients in the recurrence-free and distant recurrence health states were 0.824 and 0.685, respectively (Campbell HE, et al: Eur J Cancer 47:2517-2530, 2011). We assumed in the first year of breast cancer diagnosis, patients who received chemotherapy had disutility of 0.07, and patients who did not receive chemotherapy had disutility of 0.04 (Wang SY, et al: Value Health 19:631-638, 2016).^{14,16} We assumed the decrement in utility per patient experiencing local recurrence to be -0.108 (Campbell HE, et al: Eur J Cancer 47:2517-2530, 2011).¹⁴ Patients with CHF or AML had utility of 0.71 or 0.29,

respectively (Fryback DG, et al: Med Decis Making 13:89-102, 1993).¹⁴ Utility weights for each health state were age adjusted at 5-year increments using previously reported trends.

Evaluation of Impact of Concerning Issues on ICER Results

Our base-case analyses showed that the ICER of ODX was \$167,000 per QALY for the clinically low-risk group, \$37,800 per QALY for the clinically intermediate-risk group, and \$20,400 per QALY for the clinically high-risk group (Table 4). We then conducted a series of simulations, where each simulation examined one concerning issue we identified from our systematic review.

First, analyses where the model structure ignored clinicopathologic information could not provide ICER estimates for each risk group separately.

Second, for the concerning issue regarding a decrease in chemotherapy use in the clinically low-risk group, we used the parameters derived from prior literature¹⁶; among the clinicopathologic low-risk group, 9.8% received chemotherapy in the ODX low-risk group, 17.6% received chemotherapy in the ODX intermediate-risk group, and 63.4% received chemotherapy in the ODX high-risk group; if ODX is not provided, 46.1% received chemotherapy. Applying these parameters, our simulation showed that the ICER would be \$39,000 per QALY, indicating that this assumption could have substantial effect on the ICER estimate in the clinically low-risk group.

Third, our base-case analyses conservatively assumed that chemotherapy does not decrease distant recurrence if RS < 18 and has a similar effect in decreasing distant recurrence if RS ≥ 18, an assumption favoring ODX cost effectiveness. If we assumed that ODX does not have predictive value (ie, the relative risk reduction attributed to chemotherapy is 0.58 for the RS low-, intermediate-, and high-risk groups), the ICER of ODX would be \$307,200 per QALY for the clinically low-risk group, \$43,400 per QALY for the clinically intermediate-risk group, and \$22,700 per QALY for the clinically high-risk group. The assumption that ODX has predictive value has substantial effect on the ICER estimate in the clinically low-risk group but not in the clinically intermediate- or high-risk group.

Fourth, for the concerning issue of ignoring long-term adverse effects attributed to chemotherapy, we changed the transition probability of developing AML and CHF to 0. Under this assumption, the ICER of ODX would be \$144,300 per QALY for the clinically low-risk group, \$31,200 per QALY for the clinically intermediate-risk group, and \$17,300 per QALY for the clinically high-risk group, indicating that this concerning issue might favor ODX.

Fifth, to examine the effect of the model not using real-world population-based RS distributions, we used the Adjuvant! Online-RS joint distribution from the NSABP B-14 trial rather than the data collected from the Connecticut Tumor Registry. The ICER of ODX was \$42,000 per QALY for the clinically low-risk group, \$28,200 per QALY for the clinically intermediate-risk group, and \$25,300 per QALY for the clinically high-risk group, indicating that this assumption could lead to bias favoring ODX being cost effective in the clinically low-risk group.

If we assumed that chemotherapy increases distant recurrence for patients whose RS < 18, the ICER of ODX would be \$123,700 per QALY for the clinically low-risk group, \$34,500 per QALY for the clinically intermediate-risk group, and \$19,000 per QALY for the clinically high-risk group. Also, when we changed the starting age from 60 to 50 years, the ICER of ODX would be \$122,400 per QALY for the clinically low-risk group, \$27,500 per QALY for the clinically intermediate-risk group, and \$14,600 per QALY for the clinically high-risk group. These results suggested that these two concerning issues would favor ODX being cost effective.

Table A1. Search Terms Used in PubMed

MeSH Term Used	Other MeSH Term Used
Cost-effectiveness analysis	21-gene recurrence score OR recurrence score OR Oncotype DX
Oncotype DX 21-gene assay	Breast Cost effectiveness OR cost utility OR economic

NOTE. Search was performed on April 13, 2016.
Abbreviation: MeSH, medical subject heading.

Table A2. Full List of Parameters Used in Model for Base-Case Analyses

Parameter	Value (%)	Source
Probabilities		
Proportion of patients assigned to each risk category		
PREDICT low risk		
21-gene assay low risk	64.90	
21-gene assay intermediate risk	31.93	
21-gene assay high risk	3.17	
PREDICT intermediate risk		
21-gene assay low risk	38.74	
21-gene assay intermediate risk	39.07	
21-gene assay high risk	22.19	
PREDICT high risk		
21-gene assay low risk	33.33	
21-gene assay intermediate risk	35.96	
21-gene assay high risk	30.70	
Risk of 10-year distant recurrence without chemotherapy		
PREDICT low risk		
21-gene assay low risk	2.78	
21-gene assay intermediate risk	6.15	
21-gene assay high risk	26.36	
21-gene assay not provided	4.60	
PREDICT intermediate risk		
21-gene assay low risk	3.14	
21-gene assay intermediate risk	9.92	
21-gene assay high risk	33.84	
21-gene assay not provided	12.60	
PREDICT high risk		
21-gene assay low risk	4.16	
21-gene assay intermediate risk	13.16	
21-gene assay high risk	44.89	
21-gene assay not provided	19.90	
Risk of 10-year distant recurrence with chemotherapy		
PREDICT low risk		Calibration
21-gene assay low risk	2.78	
21-gene assay intermediate risk	1.90	
21-gene assay high risk	8.16	
21-gene assay not provided	2.67	
PREDICT intermediate risk		Calibration
21-gene assay low risk	3.14	
21-gene assay intermediate risk	5.31	
21-gene assay high risk	18.11	
21-gene assay not provided	7.31	
PREDICT high risk		Calibration
21-gene assay low risk	4.16	
21-gene assay intermediate risk	7.22	
21-gene assay high risk	24.62	
21-gene assay not provided	11.54	
Proportion of distant recurrence patients with local recurrence	10.5	Early Breast Cancer Trialists' Collaborative Group (EBCTCG): Lancet 379(9814):432-44, 2012
Adverse effects resulting from chemotherapy		Calibration
Risk of hospital visit because of toxicity	17.04	Early Breast Cancer Trialists' Collaborative Group (EBCTCG): Lancet 379(9814):432-44, 2012
Cause of hospital visits because of toxicity		Ward ¹⁴
Neutropenia/fever/infections	53.56	
Injuries and trauma	11.48	
Malignant neoplasm	10.89	
Pain and pain management	7.51	
Nausea/vomiting/dehydration	6.02	
GI tract	5.64	
Chest pain	4.89	
Mortality resulting from chemotherapy toxicity (first year)	0.35	Goldhirsch A, et al: N Engl J Med 320:491-496, 1989
8-year risk of AML resulting from chemotherapy	0.37	Praga C, et al: J Clin Oncol 23:4179-4191, 2005
Annual risk of heart failure resulting from chemotherapy	0.2	Bowles ⁵⁸
Annual survival after heart failure	86.5	Levy WC, et al: Circulation 113:1424-1433, 2006

(continued on following page)

Table A2. Full List of Parameters Used in Model for Base-Case Analyses (continued)

Parameter	Value (%)	Source
Median survival after distant recurrence, months	21	http://seer.cancer.gov/csr/1975_2005/results_merged/sect_04_breast.pdf
Median survival after AML, months	8	Edlin R, et al: Health Technol Assess 14:69-74, 2010
Mortality resulting from other causes	Life table	https://www.cdc.gov/nchs/data/hvsr65/nvsr65_08.pdf
Costs, 2015 US\$		
Cost of 21-gene assay (per patient)	3,416	Centers for Medicare and Medicaid Services ¹¹
Cost of chemotherapy regimen (AC-T)	4,343	Nadeem H, et al: J Oncol Pract 12:e299-e307, e251, 2016
G-CSF (per infusion)	3,773	https://www.goodrx.com/neupogen
Cost for recurrence treatment		
Initial cost of treatment of distant recurrence (one time)	15,710	Ward ¹⁴ ; Wong FL, et al: Ann Intern Med 160:672-683, 2014; Mandelblatt JS, et al: J Gen Intern Med 20:487-496, 2005
Ongoing care for distant recurrence per month	445	
Cost to treat local recurrence (one time)	24,045	
Cost for CHF management		
Initial cost of treatment of CHF	32,783	Dunlay SM, et al: Circ Cardiovasc Qual Outcomes 4:68-75, 2011
Ongoing care for CHF (per month)	523	Wong FL, et al: Ann Intern Med 160:672-683, 2014
Lifetime cost for AML	19,567	Ward ¹⁴
End-of-life care in addition to ongoing costs		
Death resulting from breast cancer (one time)	28,764	Levinsky NG, et al: JAMA 286:1349-1355, 2001
CHF (one time)	19,722	Bowles ⁵⁸ ; Levinsky NG, et al: JAMA 286:1349-1355, 2001
Other cause (one time)	37,131	Levinsky NG, et al: JAMA 286:1349-1355, 2001
Tamoxifen 20mg (per day)	0.3	Paulden ¹⁶
Laboratory test (per visit)	33.83	Edlin R, et al: Health Technol Assess 14:69-74, 2010
Costs, 2015 US\$		
OPD (per visit)	160.56	Edlin R, et al: Health Technol Assess 14:69-74, 2010
Echocardiogram (per visit)	251.25	Edlin R, et al: Health Technol Assess 14:69-74, 2010
Ongoing care for recurrence-free patients (per month)		Will BP, et al: Eur J Cancer 36:724-735, 2000
First year	58	
Second year	52	
Third year	46	
Fourth year	41	
≥ Fifth year	4	
Treatment of nonfatal chemotherapy toxicity		
Neutropenia/fever/infections	7,117	Paulden ¹⁶
Injuries and trauma	9,100	
Malignant neoplasm	7,041	
Pain and pain management	4,536	
Nausea/vomiting/dehydration	4,317	
GI tract	7,061	
Chest pain	3,151	
Treatment of fatal chemotherapy toxicity	48,934	Hornberger ¹⁰
Utilities		
First year after diagnosis (no chemotherapy)	0.784	Paulden ¹⁶ ; Wang SY, et al: Value Health 19:631-638, 2016
First year after diagnosis (with chemotherapy)	0.754	Ward ¹⁴
Second and following years before distant recurrence	0.824	Campbell HE, et al: Eur J Cancer 47:2517-2530, 2011
Distant recurrence	0.685	Campbell HE, et al: Eur J Cancer 47:2517-2530, 2011
Local recurrence (per person)	-0.108	Campbell HE, et al: Eur J Cancer 47:2517-2530, 2011
CHF	0.71	Fryback DG, et al: Med Decis Making 13:89-102, 1993
AML	0.26	Ward ¹⁴
Death	0	

Abbreviations: AC-T, four cycles of doxorubicin and cyclophosphamide followed by 12 cycles of paclitaxel, once every week; AML, acute myeloid leukemia; CHF, chronic heart failure; G-CSF, granulocyte colony-stimulating factor; OPD, oncology outpatient department; RS, recurrence score.

Table A3. QHES Evaluation of Analyses (N = 27)

First Author	Year	QHES															Score	
		1 ^a	2 ^b	3 ^c	4 ^d	5 ^e	6 ^f	7 ^g	8 ^h	9 ⁱ	10 ^j	11 ^k	12 ^l	13 ^m	14 ⁿ	15 ^o		
Hornberger ¹⁰	2005	7	4	8	1	9	6	5	5	8	6	0	8	7	5	8	3	90
Lyman ²⁵	2007	7	4	8	0	9	6	0	0	4	6	0	8	7	5	8	3	75
Kondo ²⁶	2008	7	4	8	1	9	6	0	5	4	6	7	8	7	5	8	3	88
Klang ²⁷	2010	7	4	8	1	9	6	0	5	4	6	0	8	7	5	8	3	81
Tsoi ²⁸	2010	7	4	8	1	9	6	5	5	8	6	0	8	7	5	8	3	90
Hornberger ²⁹	2010	7	4	8	1	0	6	0	5	4	6	0	4	7	0	8	3	63
Kondo ³⁰	2011	7	4	8	1	9	6	0	5	4	6	7	8	7	5	8	3	88
OHTA ³⁸	2010	7	4	8	1	9	6	0	5	8	6	3	8	7	5	8	3	88
Vanderlaan ³¹	2011	7	4	0	1	9	6	5	5	8	6	0	8	7	5	8	3	82
Hornberger ³²	2011	7	4	8	1	9	6	5	5	4	6	0	8	7	5	8	3	86
Retèl ³³	2011	7	4	0	0	9	6	0	7	8	6	3	8	7	5	8	3	81
Hall ³⁴	2012	7	4	8	1	9	6	5	5	8	6	3	8	7	5	8	3	93
Lamond ³⁵	2012	7	4	8	1	9	6	5	5	8	6	0	8	7	5	8	3	90
Blohmér ³⁶	2013	7	4	8	1	9	6	5	5	8	6	0	4	7	5	8	3	86
Reed ¹⁷	2013	7	4	8	1	9	6	5	5	8	6	0	8	7	5	8	3	90
Hannou ³⁷	2012	7	4	8	1	9	6	5	5	8	6	0	8	7	5	8	3	90
Davidson ³⁹	2013	7	4	8	1	9	6	5	5	8	3	3	8	7	5	8	3	90
Holt ⁴⁰	2013	7	4	8	1	9	6	5	5	8	6	3	8	7	5	8	3	93
Paulden ¹⁶	2013	7	4	8	1	9	6	5	5	8	6	7	8	7	5	8	3	97
Ward-I ¹⁴	2013	7	4	8	1	9	6	5	5	8	6	3	8	7	5	8	3	93
Ward-G ¹⁴	2013	7	4	8	1	9	6	5	5	8	6	7	8	7	5	4	3	93
Hannou ⁴¹	2014	7	4	8	1	9	6	5	5	8	6	0	8	7	5	8	3	90
Yamauchi ⁴²	2014	7	4	8	1	9	6	5	5	8	3	3	8	7	5	8	3	90
Segu ⁴³	2014	7	4	8	1	9	6	5	5	8	6	3	8	7	5	8	3	93
Kip ⁴⁴	2015	7	4	8	1	9	6	5	7	8	6	3	8	7	5	8	3	95
Katz ⁴⁵	2015	7	4	8	0	9	6	5	5	8	3	0	8	7	5	8	3	86
Jahn ⁴⁶	2015	7	4	8	1	9	6	5	5	8	6	0	8	7	5	8	3	90

Abbreviations: OHTA, Ontario Health Technology Assessment; QHES, Quality of Health Economic Studies; Ward-I, Ward-industry; Ward-G, Ward-government.

^aQHES 1. Was the study objective presented in a clear, specific, and measurable manner?

^bQHES 2. Were the perspective of the analysis (societal, third-party payer, and so on) and the reasons for its selection stated?

^cQHES 3. Were variable estimates used in the analysis from the best available source (ie, randomized controlled trial, best; expert opinion, worst)?

^dQHES 4. If estimates came from a subgroup analysis, were the groups prespecified at the beginning of the study?

^eQHES 5. Was uncertainty handled by (1) statistical analysis to address random events or (2) sensitivity analysis to cover a range of assumptions?

^fQHES 6. Was incremental analysis performed between alternatives for resources and costs?

^gQHES 7. Was the methodology for data abstraction (including the value of health states and other benefits) stated?

^hQHES 8. Did the analytic horizon allow time for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted (3% to 5%) and justification given for the discount rate?

ⁱQHES 9. Was the measurement of costs appropriate and the methodology for the estimation of quantities and unit costs clearly described?

^jQHES 10. Were the primary outcome measures for the economic evaluation clearly stated, and were the major short-term, long-term, and negative outcomes included?

^kQHES 11. Were the health outcome measures/scales valid and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?

^lQHES 12. Were the economic model (including structure), study methods and analysis, and components of the numerator and denominator displayed in a clear, transparent manner?

^mQHES 13. Were the choice of economic model, main assumptions, and limitations of the study stated and justified?

ⁿQHES 14. Did the author(s) explicitly discuss direction and magnitude of potential biases?

^oQHES 15. Were the conclusions/recommendations of the study justified and based on the study results?

^pQHES 16. Was there a statement disclosing the source of funding for the study?

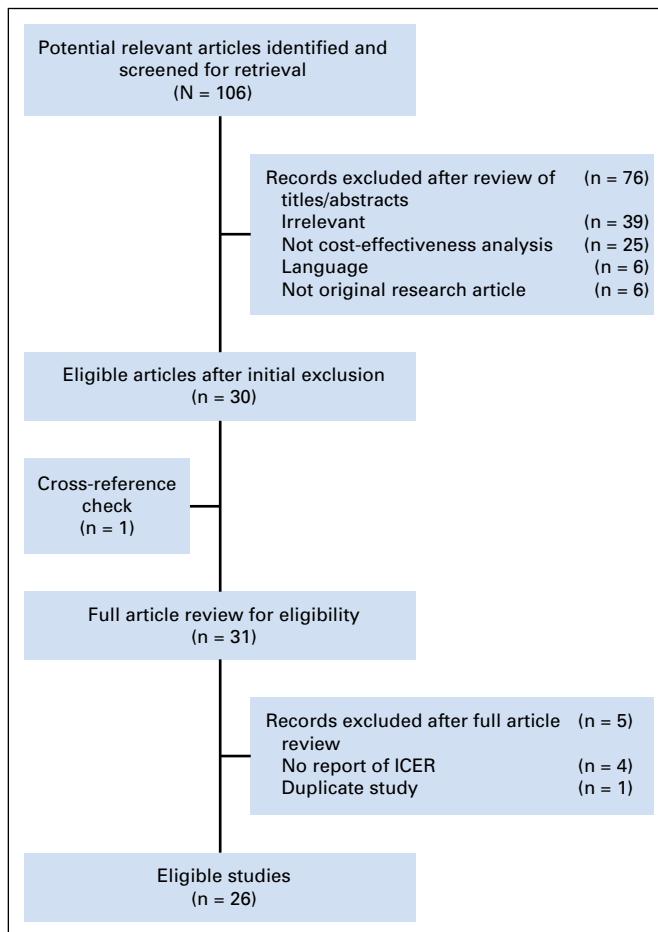
Table A4. Concerning Issues by Funding Source

Issue	No. (%)			Bias Direction	P
	Genomic Health Funded (n = 15)	Non-Genomic Health Funded (n = 12)			
Model structure					
Ignoring clinicopathologic information	12 (80.0)	6 (50.0)		Favor ODX	.10
Combining low-, intermediate-, and high-risk groups	15 (100)	10 (83.3)		Favor ODX	.18
Model assumption					
Decrease in chemotherapy use*	12 (80.0)	10 (83.3)		Favor ODX	.82
Predictive value of ODX	15 (100)	11 (91.7)		Favor ODX	.44
Ignoring chemotherapy toxicity					
Ignoring short-term toxicity	4 (26.7)	1 (8.3)		Varied†	.22
Ignoring long-term toxicity	12 (80)	6 (50)		Varied†	.11
Parameter selection					
Not real-world population-based RS distributions	15 (100)	12 (100)		Favor ODX	1
Increasing distant metastasis by chemotherapy if RS < 18	6 (40)	6 (50)		Favor ODX	.45
Young starting age	10 (66.7)	8 (66.7)		Favor ODX	1

Abbreviations: ODX, Oncotype DX; RS, recurrence score.

*Only five studies (three Genomic Health funded and two non-Genomic Health funded) assumed that ODX increases chemotherapy use in the clinicopathology-based low-risk group.

†Depending on both model structure and whether ODX increases chemotherapy use.

**Fig A1.** Flowchart for study selection for systematic review of 21-gene assay cost-effectiveness analysis. ICER, incremental cost-effectiveness ratio.

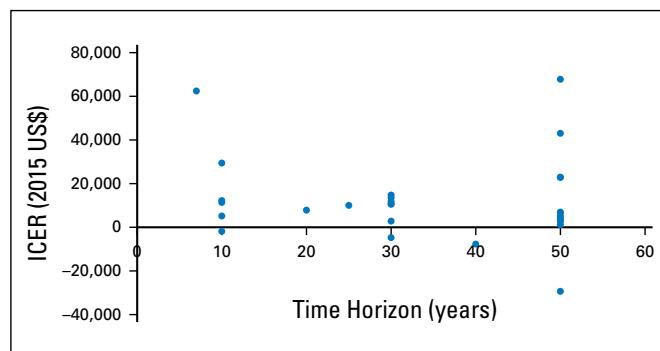


Fig A2. Incremental cost-effectiveness ratios (ICERs) by time horizon. We assumed 50 years for lifetime horizon. There was no significant association between ICERs and time horizon ($P = .84$).

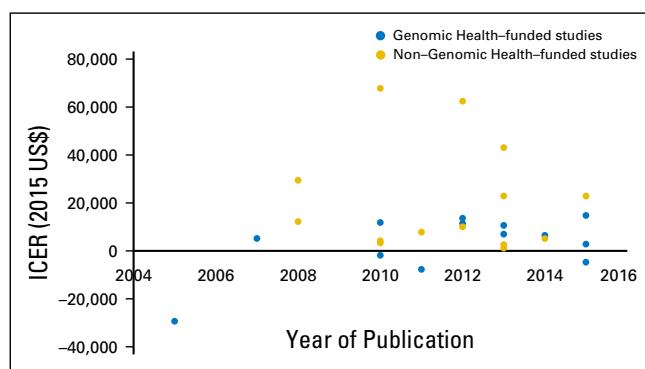


Fig A3. Incremental cost-effectiveness ratios (ICERs) by year of publication and funding source.