Published in final edited form as:

Genet Med. 2015 July ; 17(7): 519–532. doi:10.1038/gim.2014.140.

Clinical utility of gene-expression profiling in women with early breast cancer: an overview of systematic reviews

Michael Marrone, MPH[1,2], **Alison Stewart, PhD**[1,2], and **W. David Dotson, PhD**[2]

[1]McKing Consulting Corporation, Atlanta, Georgia, USA

[2]Office of Public Health Genomics, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

Abstract

Purpose—This overview systematically evaluates the clinical utility of using Oncotype DX and MammaPrint gene-expression profiling tests to direct treatment decisions in women with breast cancer. The findings are intended to inform an updated recommendation from the Evaluation of Genomic Applications in Practice and Prevention Working Group.

Methods—Evidence reported in systematic reviews evaluating the clinical utility of Oncotype DX and MammaPrint, as well as the ability to predict treatment outcomes, change in treatment decisions, and cost-effectiveness, was qualitatively synthesized.

Results—Five systematic reviews found no direct evidence of clinical utility for either test. Indirect evidence showed Oncotype DX was able to predict treatment effects of adjuvant chemotherapy, whereas no evidence of predictive value was found for MammaPrint. Both tests influenced a change in treatment recommendations in 21 to 74% of participants. The cost-effectiveness of Oncotype DX varied with the alternative compared. For MammaPrint, lack of evidence of the predictive value led to uncertainty in the cost-effectiveness.

Conclusion—No studies were identified that provided direct evidence that using gene-expression profiling tests to direct treatment decisions improved outcomes in women with breast cancer. Three ongoing studies may provide direct evidence for determining the clinical utility of gene-expression profiling testing.

Keywords

breast cancer; clinical utility; evidence-based medicine; gene-expression profiling; prognosis

Correspondence: Michael Marrone (mmarron6@jhu.edu).

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at http://www.nature.com/gim.

DISCLOSURE

The authors declare no conflict of interest.

INTRODUCTION

Breast cancer is now known to be a heterogeneous matrix of molecular subtypes characterized by significant variability in overall prognosis.[1–3] The most common subtype, luminal A, tends to be estrogen receptor (ER) positive and/or progester-one receptor (PR) positive, and HER2/neu negative, and has the best overall prognosis.[2,3] Each subtype has also been shown to be associated with a differential response to treatment with adjuvant therapy.[3–5] It has now become routine practice to tailor treatment regimens to a tumor's molecular signature (i.e., ER and PR status and HER2/neu status) while also considering traditional clinical characteristics (age, tumor size, and tumor grade).[6]

Gene-expression profiling (GEP) tests measure the expression or biological activity of several genes in tumor samples.[7,8] GEP testing may provide a more complete picture of a tumor's molecular signature and enable a better estimate of the risk of distant recurrence when considered along with other molecular signatures and clinical characteristics. This has led clinicians to speculate that GEP testing can distinguish between those patients who would benefit from adjuvant chemotherapy (i.e., high-risk GEP signatures) and those who would not (i.e., low-risk GEP signatures).[6,9]

Test description

Oncotype DX (Genomic Health) is a commercially available GEP test that uses reverse-transcriptase polymerase chain reaction to measure the expression of 21 genes from fixed, paraffin-embedded tumor samples. Based on the expression of these genes, a recurrence score (RS) on a scale from 0 to 100 quantifies the risk of 10-year distant recurrence. The Oncotype DX RS was initially validated within a cohort of lymph node–negative (LN−), estrogen receptor–positive (ER+) breast cancer patients treated with tamoxifen. Individuals with an RS less than 18 are considered to have a low risk of recurrence. Those with an RS between 18 and 31 are classified as being at intermediate risk for recurrence, and those with an RS higher than 31 are at high risk for recurrence.[7] In 2009, the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group (EWG) found adequate evidence of the clinical validity of the prognostic performance of Oncotype DX, that is, an association between the RS and disease recurrence.[10]

MammaPrint (Agendia) is built on a microarray platform that measures the expression of 70 genes, none of which overlap with Oncotype DX. MammaPrint stratifies patients into two groups—low risk or high risk of distant recurrence at 5 years.[8] In the same 2009 recommendation, the EWG found adequate evidence of the clinical validity of the prognostic ability of MammaPrint, that is, an association with future metastases.[10] Currently, MammaPrint is the only GEP test approved by the US Food and Drug Administration for determining the risk of distant recurrence at 5 and 10 years in women less than 61 years of age with stage I or II lymph node–negative early breast cancer.

With respect to clinical utility, the EWG found no direct evidence that using Oncotype DX or MammaPrint to guide treatment decisions improved health outcomes. The EWG did note that Oncotype DX held promise for the future possibility of demonstrating clinical utility.[10] Our objective was to provide an updated summary of the clinical utility of Oncotype and

MammaPrint GEP tests in women with breast cancer to support an updated recommendation from the EWG.

MATERIALS AND METHODS

Search strategy

We adapted the search strategy used by Marchionni et al.[11] to include additional text terms specific to Oncotype DX ("oncotype," "21 gene," "recurrence score") and MammaPrint ("mammaprint," "70 gene") (**Supplementary Table S1** online). We searched PubMed, EMBASE, the Cochrane Library, and CINAHL from 2007 to December 2013, and we did not limit by language or publication status.

Types of included studies

We included only systematic reviews evaluating the clinical utility of Oncotype DX or MammaPrint in women with early-stage breast cancer. We required that each systematic review report a detailed search strategy, the dates that the searches were conducted, and search terms. Systematic reviews were also required to include specific aims and objectives or key questions considered, eligibility criteria for selecting studies, and methods for assessing methodological quality of included studies. We did not include narrative reviews, editorials, or commentaries. We did not apply any restrictions on the primary studies evaluated by eligible systematic reviews. Some studies evaluated in the included systematic reviews were previously considered by Marchionni et al.[11]

Key questions

We defined clinical utility as the overall improvement in treatment outcomes using GEP to guide treatment decisions compared with outcomes in which treatment decisions did not incorporate findings from GEP tests. In addition to the primary objective of this review, in consultation with members of the EWG, we developed a list of three key questions (KQs) and outcomes based on the KQs specified by Marchionni et al.[11] Potentially eligible systematic reviews must have reviewed and summarized results from primary studies that provide direct evidence of clinical utility or studies pertaining to at least one of the following KQs:

1. Does GEP testing predict treatment effect with adjuvant chemotherapy?
    - Disease-free survival at 5 and 10 years
    - Distant recurrence (i.e., metastatic breast cancer) at 5 and 10 years
    - Overall survival at 5 and 10 years
2. To what extent are GEP test results utilized in treatment decisions?
    - Change in pre-GEP treatment decision or recommendation compared with post-GEP treatment decisions or recommendations
3. What is the cost-effectiveness and budgetary impact of using GEP testing to guide treatment decisions with adjuvant chemotherapy?

- Quality-adjusted life years (QALYs)
- Incremental cost-effectiveness ratio (ICER)
- Health-care costs

Although they may not provide direct evidence of clinical utility (i.e., improvement in treatment outcomes using GEP tests), studies of these questions form part of the evidence base on the potential impact of the clinical use of GEP and constitute the bulk of the currently available evidence regarding GEP in breast cancer.

Evidence synthesis

One review author (M.M.) extracted information describing the methodological characteristics and overall design of each included systematic review, which was verified by a second reviewer (A.S.). We only reviewed the primary publication to confirm discrepancies in data reported by the included systematic reviews or to verify pertinent data (e.g., outcome time-points). We used the AMSTAR criteria[12] to assess the methodological quality of the included systematic reviews. We performed a qualitative synthesis of the results from the primary studies reported by the systematic reviews for each KQ. If a primary study was evaluated by more than one systematic review, then we used all information reported by the included systematic reviews. We included a summary of the quality assessment of the primary studies and conclusions of each included systematic review for each KQ. We did not evaluate methodological quality of the primary studies included in the systematic reviews. We found significant clinical and methodological heterogeneity in the primary studies evaluated by the included systematic reviews, precluding meaningful meta-analyses for any of the KQs.

We incorporated findings from the included systematic reviews into a summary of findings table and followed the GRADE approach[13] to evaluate the quality of the body of evidence with respect to clinical utility. We considered the following factors that could potentially decrease the quality of evidence for each KQ: (i) study design and quality (i.e., risk of bias); (ii) indirectness in the way test results are utilized in making treatment decisions or study designs preventing direct comparison of treatment outcomes based on GEP testing (i.e., indirect evidence); (iii) inconsistency or heterogeneity in reported results; (iv) imprecision in effect estimates; and (v) potential publication bias. It could be argued that the domain for directness of evidence is not appropriate in this context, given that the KQs were known from the outset to be only indirectly related to clinical utility. However, we believe it is important to assess the overall quality of the current evidence base in relation to clinical utility, given that clinical implementation is already under way.

RESULTS

The electronic searches were last conducted on 18 December 2013, and 3,603 titles and abstracts were identified. After fulltext review of 347 studies, we included five systematic reviews evaluating Oncotype DX and MammaPrint in women with early-stage breast cancer (**Supplementary Table S2** online). The five systematic reviews, taken together, provided a chronological evidence-based assessment of both Oncotype DX and MammaPrint. After the

original review by Marchionni et al.[11] in 2007, Smartt et al.[14] published the first update in 2009 evaluating KQ 1 and KQ 2 for both Oncotype DX and MammaPrint. A 2010 review commissioned by the Ontario Health Technology Advisory Committee (OHTAC) evaluated all three KQs for Oncotype DX only.[15] The most recent systematic reviews, published in 2011, 2012, and 2013, considered multiple GEP tests in addition to Oncotype DX and MammaPrint.[16–18] One review, prepared by Ward et al.,[16] was commissioned by the National Institute for Health Research Health Technology Assessment Program on behalf of the UK National Institute for Health and Care Excellence, and evaluated all three KQs. The fourth review published in 2012 was supported by Genomic Health and evaluated all three KQs.[17] The most recent review evaluated cost-effectiveness analyses for both Oncotype DX and MammaPrint.[18] The authors of this review also reported receiving financial remuneration from Genomic Health.

In total, the five included systematic reviews evaluated seven studies related to KQ 1, 12 studies for KQ 2, and 22 studies for KQ 3. We did not consider any of the studies cited in the review by Hornberger et al.[17] addressing KQ 3 because the methodological quality of these studies was not evaluated. We excluded 13 reviews and have provided our reason for excluding each review in **Supplementary Table S4** online.

We found some overlap in the primary studies evaluated in the included systematic reviews (**Supplementary Table S3** online). For Oncotype DX, three of seven (43%) primary studies considered for KQ 1 were represented in more than one systematic review, with one study present in three systematic reviews. For Oncotype DX KQ 2, 5 of the 12 (42%) studies considered were covered in more than one systematic review, and two studies were covered in three systematic reviews. For MammaPrint, overlap was limited to KQ 2, where one of the two studies (50%) considered was covered in two systematic reviews.

Our assessment of the five systematic reviews according to the AMSTAR criteria is provided in **Supplementary Table S5** online. All five reviews reported criteria for judging the methodological quality of the studies included. Two systematic reviews used a modified version of the REMARK reporting guidelines[14,16] as well as the STARD guidelines.[14] One review used GRADE to evaluate the overall body of evidence and did not report assessments of each included study.[15] Hornberger et al.[17] applied the level-of-evidence criteria described by Simon et al.[19] for studies reporting the predictive ability of GEP tests but did not assess the methodological quality of cost-effectiveness studies. The Quality of Health Economic Studies tool was used by Rouzier et al.[18] to evaluate economic analyses.

EVIDENCE OVERVIEW

None of the included systematic reviews identified studies providing direct evidence of clinical utility for either test. The following is a summary of the evidence identified by the systematic reviews as it applies to each KQ.

Prediction of treatment effect with adjuvant chemotherapy

Oncotype DX—Four systematic reviews evaluated studies reporting the ability of Oncotype to predict treatment outcomes from adjuvant chemotherapy.[14–17] Three studies

included in the review by Smartt et al.[14] reported the association between RS and treatment outcomes in participants treated with endocrine therapy only[20,21] and in participants treated with chemotherapy only.[22] The lack of an appropriate treatment comparison limits the ability to draw conclusions regarding the predictive ability of Oncotype DX from these study reports. The OHTAC review[15] included two retrospective analyses reporting the ability of the RS to predict the effect of adjuvant chemotherapy compared with hormone therapy—the NSABP B-20 trial,[23] with 651 LN− early breast cancer participants, and the SWOG-8814 trial,[24] with 367 LN+ breast cancer participants (**Table 1**). The same retrospective analysis of LN+ breast cancer participants included in the OHTAC review, Albain et al.[24] was also included in the review by Ward et al.[16] Two additional retrospective analyses including 651 LN− participants from the NSABP B-14 and B-20 trials[25] and a conference abstract[26] that reported the analysis 651 LN− participants from NSABP B-20 data only showed a statistically significant interaction between RS and adjuvant chemotherapy (**Table 1**). Ward et al.[16] found the studies by Albain et al.[24] and Tang et al.[26] to be of moderate quality and found the study by Tang et al.[25] to be of high quality. However, they noted the absence of a reliable quality assessment tool to evaluate such studies.

The studies by Paik et al.,[23] Albain et al.,[24] and Tang et al.[25,26] investigated the predictive ability of Oncotype DX using archived tissue samples collected from randomized controlled trials comparing tamoxifen with tamoxifen plus chemotherapy. A potential limitation in the retrospective analyses is that only a subset of the participants in the original trial were included in the predictive analyses—30% and 40% in Paik et al.[23] and Albain et al.,[24] respectively. The retrospective analyses reported by Paik et al.[23] and Tang et al.[25,26] used overlapping data from participants enrolled in NSABP B-14 and NSABP-20 and included both HER2− and HER2+, which may confound the results given the correlation of HER2 status with RS.[15] Additional limitations noted in the review by Ward et al.[16] include insufficient reporting of the clinical and demographic characteristics of trial participants with tumor samples eligible for GEP testing and those included in the predictive analyses, as well as the dates or range of dates on which tumor samples were collected from the participants in two studies.[24,26] One study was reported as a conference abstract,[26] limiting the ability to complete a full assessment of the study. The corresponding full-text publication for this study has yet to be evaluated in a systematic review and is not considered in this overview.

Applying the domain-based GRADE approach, we determined the body of evidence addressing KQ 1 to be very-low-quality evidence of clinical utility (**Table 2**). This determination was due to (i) risk of bias because of the incomplete outcome data included in the retrospective analyses; (ii) the indirect nature of these study designs in addressing clinical utility; and (iii) imprecision in treatment effect estimates across all three RS strata. Three reviews included in this overview reached similar conclusions. In 2010, the Ontario Health Technology Advisory Committee concluded that “there is very low quality evidence that Oncotype-DX can predict which women will benefit from adjuvant CMF/MF chemotherapy in women being treated with adjuvant tamoxifen.”[15] Smartt et al.[14] had also concluded in 2009 that the quality of evidence demonstrating clinical utility for Oncotype

DX was low. The overall conclusion by Ward et al.[16] noted the presence of a large evidence base for Oncotype DX, but they pointed out that methodological characteristics such as heterogeneity in study participants and study design suggested the need for additional evidence to determine the clinical utility of the test and that prospective studies would improve the quality of the evidence base.

Using the level-of-evidence criteria described by Simon et al.,[19] Hornberger et al.[17] determined there is level 1, category B evidence for the Onocotype DX RS in predicting treatment effects of adjuvant chemotherapy. This designation indicates the existence of "prospective studies using archived tissue samples" that have been validated in at least two studies meeting the level 1 criteria.

MammaPrint—Four included systematic reviews did not find evidence that MammaPrint can predict treatment effects of adjuvant chemotherapy.[14,16,17] Therefore, we could not apply the GRADE criteria to define the quality of evidence for MammaPrint (**Table 3**).

Change in clinical decisions based on gene-expression profile testing

Oncotype DX—Four systematic reviews[14–17] summarized the impact of Oncotype DX on treatment recommendations (**Table 4**). Three small studies summarized by Smartt et al. showed that Oncotype DX led to a reduction in the number of LN− breast cancer patients recommended for treatment with chemotherapy.[27–29] Two additional studies described in the OHTAC review,[30,31] followed by one published report and one conference abstract summarized by Ward et al.,[32,33] showed that Oncotype DX had a similar impact on treatment recommendations. Five additional retrospective studies described by Hornberger et al. also showed that use of the RS led to an overall reduction in recommendations for chemotherapy.[34–38]

In total, 11 published reports and one conference abstract reported that the RS led to a change in treatment recommendations in 21–74% of participants. Six studies also reported that 13–34% fewer patients received chemotherapy, and one study reported 27% of patients changed their own treatment decisions after GEP testing. Heterogeneity in the molecular and clinical profiles of participants between studies and in the criteria used to determine pre-GEP treatment recommendation limits the ability to generalize the impact of Oncotype DX on treatment recommendations or decisions. Using the same domain-based approach for grading the body of evidence, we determined the studies reporting the change in treatment recommendations or decisions based on Oncotype DX results to be low-quality evidence. This designation results partly from the fact that these types of study cannot provide direct evidence of clinical utility; however, it also reflects problems with the quality of evidence provided by these studies, particularly due to heterogeneity in the study participants (**Table 2**). The OHTAC review found insufficient evidence for the impact of Oncotype DX on clinical decision making.

MammaPrint—One study described in two reviews[14,17] reported 61% (259/427) of patients were recommended to receive adjuvant treatment (chemotherapy or hormone therapy or both) based on current clinical guidelines, prognosis signature, and patient preference compared with 47% who had been recommended for adjuvant systemic therapy

based on clinician interpretation of current Dutch clinical guidelines[39] (**Table 4**). Among the participants who were recommended to receive adjuvant therapy after consideration of clinical guidelines, patient preference, and prognosis signature, 18, 29, and 13% of participants were recommended for chemotherapy alone, chemotherapy plus hormone therapy, and hormone therapy only, respectively. One additional study summarized by Ward et al.[40] retrospectively compared actual treatment received with participants' MammaPrint risk classifications. Among 59 participants classified as high risk, 19 (32%) did not receive chemotherapy, whereas 35 out of 77 participants classified as low risk did receive chemotherapy. It was suggested that MammaPrint would affect the treatment 40% of the time based on the assumption that all high-risk participants would receive chemotherapy and all low-risk participants would not.[40] Heterogeneity in the study populations and in the initial criteria used to make treatment decisions limits the ability to generalize the impact of MammaPrint on clinical decision making. Indirectness in the interpretation of results, participant characteristics, and the lack of a reliable control groups led us to determine there is low-quality evidence for MammaPrint to lead to change in treatment decisions or recommendations (**Table 3**).

Cost-effectiveness of gene-expression profile testing

Two systematic reviews conducted cost-effectiveness analyses of Oncotype DX[15,16] and MammaPrint[16] in which model parameters were based on results reported in studies included in their respective systematic review when available. Because of the lack of published prospective data, additional unpublished clinical data were also used in the cost-effectiveness analyses. One systematic review summarized both cost-effective analyses and budget impact analyses conducted in a variety of settings with different societal perspectives[18] (**Table 5**).

Oncotype DX—The OHTAC review found, in a Canadian setting, that at a cost of 4,191 Canadian dollars (CAD) for Oncotype DX in August 2010, the total testing and treatment cost per patient if all patients received the test was CAD 17,466. This resulted in an ICER of $23,983 per QALY gained when compared with testing only patients at intermediate/high risk based on Adjuvant! Online. Based on this model, OHTAC concluded "it is cost-effective to provide Oncotype-DX to all patients at any typical willingness-to-pay for a QALY."[15]

In the UK National Health Service setting, Ward et al. found that, when compared with current clinical practice, offering Oncotype DX to all women with ER+, LN−, HER2− early breast cancer at a cost of £2,580 (2010) resulted in an ICER of £26,940 per QALY gained. In a sensitivity analysis in which only women who were ER+, LN−, and HER2−, and who also had a Nottingham Prognostic Index of 3.4 or more were administered the test, the ICER decreased to £9,007 per QALY gained as compared with current practice.[16]

Across 22 studies[35,36,41–60] reviewed by Rouzier et al.,[18] cost-effectiveness ranged from ¥1,239,005 to CAD 63,054 per QALY gained when compared with a variety of alternatives (National Comprehensive Cancer Network guidelines, St. Gallen guidelines, Adjuvant! Online, or current clinical practice). Use of Oncotype DX led, in some studies, to decreases

in direct costs ($2,256 to more than $1 million), whereas in others it led to increased direct costs ($1,828 to $4,272) (**Table 5**). Rouzier et al.[18] concluded Oncotype DX would likely improve outcomes, reduce the number of patients treated with chemotherapy, and be cost-effective from the perspective of the payer. Cost-saving and improvements in quality of life were likely attributable to less frequent use of chemotherapy. Because of uncertainty in the cost-effectiveness and the variation in the magnitude and direction of costs (increase or decreases) from budget impact assessments of Oncotype DX, we found the economic analyses to be evidence of moderate quality (**Table 2**).

MammaPrint—There was considerable uncertainty in the evidence with respect to the benefit of chemotherapy in each risk group, resulting in less reliable estimates for parameters in the cost-effectiveness model.[15] In the analysis by Ward et al.[16] for the National Institute for Health and Care Excellence in a UK setting, at a cost of £2,657 in 2010, the ICER for using MammaPrint ranged between £12,240 and £53,058 per QALY gained (compared with current clinical practice) when the test was administered to all ER+, LN−, HER2− early breast cancer patients. In the subsequent sensitivity analysis in which the test was only administered to women with a Nottingham Prognostic Index of 3.4 or more, the ICER ranged from £6,053 to £29,569 per QALY gained compared with current practice. This uncertainty was also observed across the studies described by Rouzier et al.[18] One study reported a decrease in direct costs of €7,430 compared with St. Gallen guidelines,[61] whereas two studies reported increases in direct cost of ¥231,385[62] and from $401 to $1,811[63] per patient compared with St. Gallen guidelines. Against Adjuvant! Online, one study reported an increase of €1,130 in direct costs with €4,614 per QALY gained.[61] Two other studies reported a $2,882 decrease in direct and indirect costs when compared with National Institutes of Health clinical practice guidelines[64] and €9,043 cost-savings per 100 patients per year compared with current practice.[65] The five cost-effectiveness analyses described by Rouzier et al.[18] had an average Quality of Health Economic Studies score of 86. Imprecision in the estimates of cost-effectiveness due to the lack of data on the estimated treatment benefits led us to determine there is low-quality evidence for MammaPrint (**Table 3**).

DISCUSSION

We identified five systematic reviews[14–18] that provide a chronological assessment of Oncotype DX and MammaPrint from 2007 through 2013. None of the included systematic reviews reported direct evidence of clinical utility for either test, and deficiencies were also found in the quality of the indirect evidence relating to clinical utility.

The included systematic reviews identified four studies[23–26] that reported results suggesting the Oncotype DX RS can predict treatment effects of adjuvant chemotherapy in LN−/+, ER+, HER2− breast cancer. In addition to the indirect nature of the study designs, which did not compare outcomes in women whose treatment was guided by GEP test results with those whose treatment was based on standard practice, we also found potential risk of bias from incomplete outcome data and imprecision in effect estimates, leading to the overall conclusion that these studies represent very-low-quality evidence of clinical utility.

Oncotype DX was also shown to provide additional clinical information to influence a change in the treatment recommendations in up to 74% of women. However, reports of a net reduction in chemotherapy recommendations or decisions in the study group as a whole sometimes obscured a more complex situation in which it was not always clear in what direction GEP testing influenced a change in treatment decisions (from less aggressive/toxic treatment to more aggressive, or vice versa).

For MammaPrint, none of the reviews identified evidence of the ability to predict treatment effects of adjuvant chemotherapy. MammaPrint influenced treatment decisions, but uncertainty remains in how to interpret the MammaPrint risk classification to determine the most appropriate treatment regimen.

Four systematic reviews evaluated the cost-effectiveness of both tests. Independent cost-effectiveness models developed in two included reviews[15,16] showed that in the Canadian and UK health-care settings, most of the additional costs incurred by introducing Oncotype DX testing were associated with the cost of the test itself, followed by the cost of additional chemotherapy. Another important finding from sensitivity analyses was that using Adjuvant! Online to triage patients may provide a cost-savings by reserving GEP testing for those patients whose prognosis is most uncertain and whose care will most likely be influenced by GEP testing.[16] However, the variability in the cost-effectiveness for both tests illustrates the underlying uncertainty in the clinical effectiveness and overall utility of these tests.

A potential limitation in our overview is that it is only as up to date as the most recent search of the included systematic reviews (May 2011 and January 2012). Despite this limitation, we believe that the inclusion of additional studies, related to any of the three KQs, would be unlikely to change the conclusions of the current overview or our judgment of the overall quality of evidence. Another potential limitation is that secondary publications are the unit of analysis in this overview, rather than the primary studies themselves. It is possible that latent biases from systematic reviews, compounded with those of the individual studies, could be inadvertently carried forward in, and propagated by, an overview of reviews such as this.[66] This concern may be at least partially mitigated by the relatively high degree of study overlap between systematic reviews, although this may also reduce the power of our overview and limit the strength of our conclusions.

Two ongoing clinical trials—TAILORx[67] and MINDACT[68]— will provide direct evidence for using Oncotype DX and MammaPrint to guide treatment decisions in women with early-stage, ER+, HER2−, LN− breast cancer. Both trials were designed to compare outcomes in women whose treatment was based on GEP test results with outcomes in women treated according to standard practices. In TAILORx, women with an intermediate RS were randomized to receive hormone therapy plus adjuvant chemotherapy or hormone therapy alone, whereas women with a low RS received hormone therapy only and women with high RS received hormone therapy plus adjuvant chemotherapy. Women in MINDACT were evaluated with both MammaPrint and Adjuvant! Online. Women with discordant results were then randomized to the adjuvant chemotherapy group or no chemotherapy group. Another recently launched trial, RxPONDER,[69] is evaluating Oncotype DX in participants with LN+ breast cancer.

Despite the lack of direct evidence of clinical utility, Oncotype DX is currently recommended by the American Society of Clinical Oncology[9] and the National Comprehensive Cancer Network[6] based on the reasoning that studies designed to address our KQ 1 can provide sufficient evidence to support clinical implementation. The National Comprehensive Cancer Network used a similar level of evidence framework as Hornberger et al.[17] and rated it at as category IIA, concluding that use of the intervention is appropriate but recommending that the RS should be interpreted "in the context of other elements of risk stratification for an individual."[6] The UK's National Institute for Health and Care Excellence released a provisional recommendation in 2012 in which the use of Oncotype DX was not recommended for guiding the use of adjuvant chemotherapy in ER+, LN−, HER2− breast cancer.[70] In early 2013, the National Institute for Health and Care Excellence released an updated provisional recommendation for public comment in which Oncotype DX was recommended for ER+, LN−, HER2− early breast cancer that is intermediate risk (on the basis of the Nottingham Prognostic Index) and when the decision to use chemotherapy is uncertain. The updated recommendation would only apply if Genomic Health provided Oncotype DX to the National Health Service at an undisclosed price.[71]

Another important question concerns the most appropriate cut-points for defining the risk of recurrence using Oncotype DX. The RS cut-point in the TAILORx trial is more conservative than that used in previous studies (low risk <11 compared with <18; intermediate risk 11 to 25 compared with 18 to 31; and high risk >25 compared with >31). One of the objectives of the RxPONDER trial is to determine an RS cut-point above which chemotherapy may be most beneficial.[69] Researchers have also begun to compare Oncotype DX with standard methods for measuring other molecular targets already included in clinical decision making (e.g., ER, PR, and HER2).[72,73] The analytic and clinical validity and utility of incorporating these established molecular markers within a GEP test should be evaluated before a single GEP platform is used for all molecular profiling.

As evidence accumulates of the prognostic and predictive ability of both tests in patients at low risk on the basis of clinical factors, investigators are now applying these tests in higher-risk groups. Recent findings have shown those with a good prognosis on the basis of MammaPrint results were more likely to have pathological complete response and recurrence-free survival when treated in the neoadjuvant setting.[74] Additional validation of the prognostic and predictive ability of these tests in high-risk patients is another avenue of further research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We acknowledge the assistance provided by Andy Freedman, Sheri Schully, Pamela Smartt, Sue Ward, and Alison Scope, who critically reviewed the manuscript and provided thoughtful detailed comments to improve this report. We also thank the members of the EGAPP Working Group, who provided assistance and guidance throughout the review process. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

REFERENCES

1. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012; 490:61–70. [PubMed: 23000897]
2. Voduc KD, Cheang MC, Tyldesley S, Gelmon K, Nielsen TO, Kennecke H. Breast cancer subtypes and the risk of local and regional relapse. J Clin Oncol. 2010; 28:1684–1691. [PubMed: 20194857]
3. Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA. 2006; 295:2492–2502. [PubMed: 16757721]
4. Dawood S, Hu R, Homes MD, et al. Defining breast cancer prognosis based on molecular phenotypes: results from a large cohort study. Breast Cancer Res Treat. 2011; 126:185–192. [PubMed: 20711652]
5. Kim SR, Paik S. Genomics of adjuvant therapy for breast cancer. Cancer J. 2011; 17:500–504. [PubMed: 22157294]
6. Network NCC. NCCN Clinical Practice Guidelines in Oncology - Breast Cancer. 2013 v2.2013.
7. 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. 2004; 351:2817–2826. [PubMed: 15591335]
8. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature. 2002; 415:530–536. [PubMed: 11823860]
9. Harris L, Fritsche H, Mennel R, et al. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. J Clin Oncol. 2007; 25:5287–5312. [PubMed: 17954709]
10. Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group. Recommendations from the EGAPP Working Group: can tumor gene expression profiling improve outcomes in patients with breast cancer? Genet Med. 2009; 11:66–73. [PubMed: 19125125]
11. Marchionni L, Wilson RF, Marinopoulos SS, et al. Impact of gene expression profiling tests on breast cancer outcomes. Evid Rep Technol Assess (Full Rep). 2007:1–105. [PubMed: 18457476]
12. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol. 2007; 7:10. [PubMed: 17302989]
13. Guyatt G, Oxman A, Vist G, Kiun R, Falck-Ytter Y, Schunemann H. GRADE: what is "quality of evidence" and why is it important to clinicians? BMJ. 2008; 336:995–998. [PubMed: 18456631]
14. Smartt P. A Comparison of Gene Expression Profiling Tests for Breast Cancer. HSAC Report. 3(16):2009.
15. Gene expression profiling for guiding adjuvant chemotherapy decisions in women with early breast cancer: an evidence-based and economic analysis. Ontario Health Technology Assessment Series [Internet]. 2010; 10:1–57.
16. Ward S, Scope A, Rafia R, Pandor A, Harnan S, Evans P. Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in breast cancer management. Health Technology Assessment. 2011
17. Hornberger J, Alvarado MD, Rebecca C, Gutierrez HR, Yu TM, Gradishar WJ. 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. 2012; 104:1068–1079. [PubMed: 22767204]
18. Rouzier R, Pronzato P, Chéreau E, Carlson J, Hunt B, Valentine WJ. Multigene assays and molecular markers in breast cancer: systematic review of health economic analyses. Breast Cancer Res Treat. 2013; 139:621–637. [PubMed: 23722312]
19. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. J Natl Cancer Inst. 2009; 101:1446–1452. [PubMed: 19815849]
20. Akashi-Tanaka S, Shimizu C, Ando M, et al. 21-Gene expression profile assay on core needle biopsies predicts responses to neoadjuvant endocrine therapy in breast cancer patients. Breast. 2009; 18:171–174. [PubMed: 19410462]

21. Kok M, Linn SC, Van Laar RK, et al. Comparison of gene expression profiles predicting progression in breast cancer patients treated with tamoxifen. Breast Cancer Res Treat. 2009; 113:275–283. [PubMed: 18311582]
22. Li LF, Xu XJ, Zhao Y, et al. Integrated gene expression profile predicts prognosis of breast cancer patients. Breast Cancer Res Treat. 2009; 113:231–237. [PubMed: 18278552]
23. 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. 2006; 24:3726–3734. [PubMed: 16720680]
24. Albain KS, Barlow WE, Shak S, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol. 2010; 11:55–65. [PubMed: 20005174]
25. 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. 2011; 127:133–142. [PubMed: 21221771]
26. Tang, CJ.; Crager, M.; Shak, S.; Wolmark, N. Comparing the prediction of chemotherapy benefit in patient with node-geative, ER-postive breast cancer using the recurrence score and new measure that integrates clinical and pathological factors with the recurrence score. 33rd Annual San Antonio Breast Cancer Symposium; San Antonio, TX. 8-12 December 2010;
27. Asad J, Jacobson AF, Estabrook A, et al. Does oncotype DX recurrence score affect the management of patients with early-stage breast cancer? Am J Surg. 2008; 196:527–529. [PubMed: 18809056]
28. Rayhanabad JA, Difronzo LA, Haigh PI, Romero L. Changing paradigms in breast cancer management: introducing molecular genetics into the treatment algorithm. Am Surg. 2008; 74:887–890. [PubMed: 18942607]
29. Henry LR, Stojadinovic A, Swain SM, Prindiville S, Cordes R, Soballe PW. The influence of a gene expression profile on breast cancer decisions. J Surg Oncol. 2009; 99:319–323. [PubMed: 19204954]
30. Geffen DB, Amir N, Sion-Vardy N, et al. Stage I breast cancer in a regional oncology practice in Israel 2002–2006: clinicopathologic features, risk estimation and planned therapy of 328 consecutive patients. Breast. 2009; 18:316–321. [PubMed: 19819143]
31. Lo SS, Mumby PB, Norton J, et al. 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. 2010; 28:1671–1676. [PubMed: 20065191]
32. Ademuyiwa FO, Miller A, O'Connor T, et al. The effects of oncotype DX recurrence scores on chemotherapy utilization in a multi-institutional breast cancer cohort. Breast Cancer Res Treat. 2011; 126:797–802. [PubMed: 21197567]
33. Holt BG, Pudney D, Rolles M, et al. Results from a prospective clinical study on the impact of Oncotype DX on adjuvant treatment decision and risk classification by Nottingham prognostic index (NPI) and recurrence score (RS). Breast. 2011; (suppl 1) s43.
34. Oratz R, Paul D, Cohn AL, Sedlacek SM. Impact of a commercial reference laboratory test recurrence score on decision making in early-stage breast cancer. J Oncol Pract. 2007; 3:182–186. [PubMed: 20859407]
35. Klang SH, Hammerman A, Liebermann N, Efrat N, Doberne J, Hornberger J. Economic implications of 21-gene breast cancer risk assay from the perspective of an Israeli-managed health-care organization. Value Health. 2010; 13:381–387. [PubMed: 20412544]
36. Hornberger J, Chien R, Krebs K, Hochheiser L. US insurance program's experience with a multigene assay for early-stage breast cancer. J Oncol Pract. 2011; 7(suppl 3):e38s–e45s. [PubMed: 21886510]
37. Joh JE, Esposito NN, Kiluk JV, et al. The effect of Oncotype DX recurrence score on treatment recommendations for patients with estrogen receptor-positive early stage breast cancer and correlation with estimation of recurrence risk by breast cancer specialists. Oncologist. 2011; 16:1520–1526. [PubMed: 22016474]

38. Partin JF, Mamounas EP. Impact of the 21-gene recurrence score assay compared with standard clinicopathologic guidelines in adjuvant therapy selection for node-negative, estrogen receptor-positive breast cancer. Ann Surg Oncol. 2011; 18:3399–3406. [PubMed: 21537874]
39. Bueno-de-Mesquita JM, van Harten WH, Retel VP, et al. Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: a prospective community-based feasibility study (RASTER). Lancet Oncol. 2007; 8:1079–1087. [PubMed: 18042430]
40. Gevensleben H, Göhring UJ, Büttner R, et al. Comparison of MammaPrint and TargetPrint results with clinical parameters in German patients with early stage breast cancer. Int J Mol Med. 2010; 26:837–843. [PubMed: 21042777]
41. 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. 2005; 11:313–324. [PubMed: 15898220]
42. Lyman GH, Cosler LE, Kuderer NM, Hornberger J. 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. 2007; 109:1011–1018. [PubMed: 17311307]
43. Kondo M, Hoshi SL, Ishiguro H, Yoshibayashi H, Toi M. 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. 2008; 112:175–187. [PubMed: 18075786]
44. Cosler LE, Lyman GH. Economic analysis of gene expression profile data to guide adjuvant treatment in women with early-stage breast cancer. Cancer Invest. 2009; 27:953–959. [PubMed: 19909009]
45. de Lima Lopes, G.; Chien, R.; Hornberger, J. Cost-benefit analysis of a 21-gene recurrence score for early-stage breast cancer in Singapore. ISPOR 4th Asia-Pacific Conference; Phuket, Thailand. 5-7 September 2010;
46. O'Leary, B.; Yoshizawa, C.; Foteff, C.; Chao, C. Cost-effectiveness of the Oncotype DX assay in Australia: an exploratory analysis. ISPOR 4th Asia-Pacific Conference; Phuket, Thailand. 5-7 September 2010;
47. Tsoi DT, Inoue M, Kelly CM, Verma S, Pritchard KI. Cost-effectiveness analysis of recurrence score-guided treatment using a 21-gene assay in early breast cancer. Oncologist. 2010; 15:457–465. [PubMed: 20421264]
48. de Lima Lopes, G.; Chien, R.; Hornberger, J. Cost-benefit analysis of a 21-gene recurrence score for early-stage breast cancer in Singapore. 12th St. Gallen International Breast Cancer Conference; St Gallen, Switzerland. 16-19 March 2011;
49. Holt, S.; Bennet, H.; Bertelli, G.; Valentine, WJ.; Phillips, C. Cost-effectiveness evaluation of the Oncotype DX breast cancer assay in clinical practice in the UK. 34th Annual San Antonio Breast Cancer Symposium; San Antonio, TX. 6-10 December 2011;
50. Kondo M, Hoshi SL, Yamanaka T, Ishiguro H, Toi M. 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. 2011; 127:739–749. [PubMed: 21082239]
51. Hassan, S.; Mittmann, N. A cost benefit analysis of the 21-gene breast cancer assay within a Canadian healtcare system. ASCO Annual metting; Chicago, IL. 2-7 June 2011;
52. Lacey, L.; Chien, R.; Hornberger, J. Cost-utility of the 21-gene breast cancer assay (Oncotype DX) in the Irish healtcare setting. 34th Annual San Antonio Breast Cancer Symposium; San Antonio, TX. 6-10 December 2011;
53. Lacey, L.; Hornberger, J. Economic evaluation of Oncotype DX to target chemotherapy use in lymph-node-negative, oestrogen-receptor-positive, early-stage breast cancer. ISPOR 13th Annual European Conference; Prague, Czech Republic. 6-9 November 2010;
54. Paulden, M.; Franek, J.; Pham, B.; Krahn, M. Gene expression profiling for guidling adjuvant chemotherapy decisions in women with early breast cancer: a cost-effectiveness analysis of 1000 strategies for the provision of Adjuvant! Online, Oncotype DX and chemotherapy. ISPOR 16th Annual International Meeting; Baltimore, MD. 21-25 May 2011;

55. Ragaz, J.; Wilson, K.; Wong, H.; Muraca, G.; Hryniuk, W.; Bajdik, C. Molecualr classification with 21 gene assay (Oncotype DX) shows in 196,967 ER positive patients high frequency of low reccurence. 34th Annual San Antonio Breast Cancer Symposium; San Antonio, TX. 6-10 December 2011;
56. Vanderlaan B, Broder MS, Chang E, Oratz R, Bentley T. Cost-effectiveness of 21-gene assay in node-positive, early-stage breast cancer. Am J Managed Care. 2011; 17:455–464.
57. Hall PS, McCabe C, Stein RC, Cameron D. Economic evaluation of genomic test-directed chemotherapy for early-stage lymph node-positive breast cancer. J Natl Cancer Inst. 2012; 104:56–66. [PubMed: 22138097]
58. Lamond NW, Skedgel C, Rayson D, Lethbridge L, Younis T. Cost-utility of the 21-gene recurrence score assay in node-negative and node-positive breast cancer. Breast Cancer Res Treat. 2012; 133:1115–1123. [PubMed: 22361999]
59. Madaras, B.; Rozsa, P.; Gerencser, Z. The impact of chemotherapeutic regimes on the cost-utility analysis of Oncotype DX assay. EBCC 8, Vienna; Vienna, Austria. 21–24 March 2012;
60. Wilson, E.; McDonnell, D.; Gullo, G., et al. Economic impact of Oncotype DX assay in axillary node negative breast cancer, (AXN-BC) with postivie hormone receptor (REC+) and normal HER-2 (HER2−). European Society of Medical Oncology; Milan, Italy. 8-12 October 2010;
61. Retèl VP, Joore MA, Knauer M, Linn SC, Hauptmann M, Harten WH. Cost-effectiveness of the 70-gene signature versus St. Gallen guidelines and Adjuvant Online for early breast cancer. Eur J Cancer. 2010; 46:1382–1391. [PubMed: 20359886]
62. Kondo M, Hoshi SL, Ishiguro H, Toi M. Economic evaluation of the 70-gene prognosis-signature (MammaPrint®) in hormone receptor-positive, lymph node-negative, human epidermal growth factor receptor type 2-negative early stage breast cancer in Japan. Breast Cancer Res Treat. 2012; 133:759–768. [PubMed: 22315134]
63. Chen E, Tong KB, Malin JL. Cost-effectiveness of 70-gene MammaPrint signature in node-negative breast cancer. Am J Manag Care. 2010; 16:e333–e342. [PubMed: 21291290]
64. Oestreicher N, Ramsey SD, Linden HM, et al. Gene expression profiling and breast cancer care: what are the potential benefits and policy implications? Genet Med. 2005; 7:380–389. [PubMed: 16024969]
65. Zarca, D.; Stork-Sloots, K.; de Snoo, F.; Brink, G.; Dervaux, Y.; Boubli, L. French cost effectiveness study of the MammaPrint 70-gene signature in early-stage breast cancer patients. European Cancer Organization; Berlin, Germany. 20-24 September 2009;
66. Thomson D, Russell K, Becker L, Klassen T, Hartling L. The evolution of a new publication type: Steps and challenges of producing overviews of reviews. Res Synthesis Methods. 2010; 1:198–211.
67. Zujewski J, Kamin L. Trial assessing individualized options for treatment for breast cancer: the TAILORx trial. Breast Cancer Res Treat. 2008; 133:759–768.
68. Rutgers E, Piccart-Gebhart M, Bogaerts J, Delaloges S, Van 't Veer L, Rubio I. The EORTC 10041/BIG 03-04 MINDACT trial is feasible: results of the pilot phases. Eur J Cancer. 2011; 47:2742–2749. [PubMed: 22051734]
69. Ramsey SD, Barlow WE, Gonzalez-Angulo AM, et al. Integrating comparative effectiveness design elements and endpoints into a phase III, randomized clinical trial (SWOG S1007) evaluating oncotypeDX-guided management for women with breast cancer involving lymph nodes. Contemp Clin Trials. 2013; 34:1–9. [PubMed: 23000081]
70. National Institute for Health and Clinical Excellence. Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in early breast cancer mamanagement: MammaPrint, Oncotype DX, IHC4 and Mammostrat. Diagnostics consultation documents. 2012. http://guidance.nice.org.uk/DT/4/Consultation/DraftGuidance/pdf/English. Accessed 17 February, 2013
71. National Institute for Health and Clinical Excellence. Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in early breast cancer mamanagement: MammaPrint, Oncotype DX, IHC4 and Mammostrat. Diagnostics consultation documents. 2013. http://guidance.nice.org.uk/DT/4/9Jan2013/pdf/English. Accessed 19 Februrary 2013

72. Dvorak L, Dolan M, Fink J, Varghese L, Henriksen J, Gulbahce HE. Correlation between HER2 determined by fluorescence in situ hybridization and reverse transcription-polymerase chain reaction of the oncotype DX test. Appl Immunohistochem Mol Morphol. 2013; 21:196–199. [PubMed: 22914611]

73. Kraus JA, Dabbs DJ, Beriwal S, Bhargava R. Semi-quantitative immunohistochemical assay versus oncotype DX(®) qRT-PCR assay for estrogen and progesterone receptors: an independent quality assurance study. Mod Pathol. 2012; 25:869–876. [PubMed: 22301704]

74. Esserman LJ, Berry DA, Cheang MC, et al. Chemotherapy response and recurrence-free survival in neoadjuvant breast cancer depends on biomarker profiles: results from the I-SPY 1 TRIAL (CALGB 150007/150012; ACRIN 6657). Breast Cancer Res Treat. 2012; 132:1049–1062. [PubMed: 22198468]

Table 1

Ability of Oncotype DX to predict benefit from adjuvant chemotherapy

Study	Design	Population	Treatment	Outcome measure	Treatment outcomes
Paik (2006)[23]	Prospective-retrospective study in NSABP B-20	ER+	HT: tamoxifen (*N* = 227)	10-year distant recurrence	Low RS: CHT = 4.4%/HT = 3.2%; RR = 1.31 (0.46–3.78)
		LN−			
		HER2+/−	CHT: tamoxifen + CMF/MF[a,b] (*N* = 424)		Intermediate RS: CHT = 10.9%/ HT = 9.1%; RR = 0.61 (0.24–1.59)
		N = 651/2,229 (28.9%)			High RS: CHT = 11.9%/HT = 39.5%; RR = 0.26 (0.13–0.53)
Albain (2010)[24]	Prospective-retrospective study in SWOG-8814	ER and/or PR+	HT: tamoxifen (*N* = 148)	10-year disease-free survival	Low RS: HR = 1.02 (0.54–1.93)
		LN+			Intermediate RS: HR: 0.72 (0.39–1.31)
		HER2+/−			High RS: HR = 0.59 (0.35–1.01)
		Mean age: 60.4 years; *N* = 367/927 (39.6%)	CHT: CAF >tamoxifen[c] (*N* = 219)	10-year overall survival	Low RS: HR = 1.18 (0.55–2.54)
					Intermediate RS: HR = 0.84 (0.40–1.78)
					High RS: HR = 0.56 (0.31–1.02)
Tang (2011)[25]	Prospective-retrospective study in NSABP B-14 and NSABP B-20	ER+	HT: tamoxifen	Distance recurrence-free interval	*P* = 0.031 for RS × treatment interaction
		LN−		Overall survival	*P* = 0.011 for RS × treatment interaction
		N = 651 (B20 cohort)	CHT: tamoxifen + chemotherapy	Disease-free survival	*P* = 0.082 for RS × treatment interaction
Tang (2010)[d][26]	Prospective-retrospective study in NSABP B-20	ER+	HT: tamoxifen	Distant recurrence	HR = 0.84 (*P* = 0.037 for RS × treatment interaction)
		LN− *N* = 625	CHT: NR		

CHT, chemotherapy; ER, estrogen receptor; HT, hormone therapy; LN, lymph node; PR, progesterone receptor; RR, relative risk; RS, recurrence score.

[a]CMF, cyclophosphamide, methotrexate, fluorouracil 5-FY.

[b]MF, methotrexate and fluorouracil 5-FU.

[c] CAF>tamoxifen—cyclophosphaminde, doxorubicin, fluorouracil 5-FU followed by tamoxifen.

[d] Results from conference abstract.

Table 2

Summary of findings table for Oncotype DX

Quality assessment							Summary of findings					
							Events (%)				Anticipated absolute effects	
Participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Quality of evidence	RS	CHT	Non-CHT	Relative effect (95% CI)	Risk in controls (non-CHT)	Risk difference (95% CI)
Disease-free survival—5-years												
0 (0)						No evidence	Low:	NR	NR			
							Int.:	NR	NR			
							High:	NR	NR			
Disease-free survival—10 years												
367 (1 study)	Serious[a]	Undetected[b]	Serious[c]	Serious[d]	Undetected	Very low	Low:	NR	NR	HR = 1.02 (0.54–1.93)	Not estimable	Not estimable
							Int.:	NR	NR	HR = 0.72 (0.39–1.31)	Not estimable	Not estimable
							High:	NR	NR	HR = 0.59 (0.35–1.01)	Not estimable	Not estimable
Distant recurrence—5 years												
0 (0)						No evidence	Low:	NR	NR			
							Int.:	NR	NR			
							High:	NR	NR			
Distant recurrence—10 years												
651 (1 study)	Serious[e]	Undetected[b]	Serious[c]	Serious[d]	Undetected	Very low	Low:	4.4%	3.2%	RR = 1.31 (0.46–3.78)	3 per 100	2 fewer to 8 more
							Int.:	10.9%	9.1%	RR = 0.61 (0.24–1.59)	9 per 100	7 fewer to 5 more
							High:	11.9%	39.5%	RR = 0.26 (0.13–0.53)	40 per 100	18 fewer to 35 fewer
Overall survival—5 years												

Quality assessment							Summary of findings					
							Events (%)				Anticipated absolute effects	
Participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Quality of evidence	RS	CHT	Non-CHT	Relative effect (95% CI)	Risk in controls (non-CHT)	Risk difference (95% CI)
0 (0)						No evidence	Low:	NR	NR			
							Int.:	NR	NR			
							High:	NR	NR			
Overall survival—10 years												
367 (1 study)	Serious[a]	Undetected[b]	Serious[c]	Serious[d]	Undetected	Very low	Low:	NR	NR	HR = 1.18 (0.55–2.54)	Not estimable	Not estimable
							Int.:	NR	NR	HR = 0.84 (0.4–1.78)	Not estimable	Not estimable
							High:	NR	NR	HR = 0.56 (0.31–1.02)	Not estimable	Not estimable
Change in treatment decisions												
2,330 (12 studies)	Undetected	No serious inconsistency	Very serious[c,f]	No serious imprecision	Undetected	Low	Range: 21–74%			Not estimable	Not estimable	Not estimable
Economic evaluation												
24 Studies	Undetected	Serious inconsistency[g]	No serious indirectness	No serious imprecision	Undetected	Moderate	Not estimable			ICER/QALY gained[f]: CAD 518 to CAD 63,054	Costs: SGD 5,138 decrease to $4,272 increase	QALYs gained[f]: −0.31 to 1.71

CAD, Canadian dollars; CHT, chemotherapy; CI, confidence interval; HR, hazard ratio; ICER, incremental cost-effectiveness ratio; NR, not reported; QALY, quality-adjusted life year; RR, relative risk; RS, recurrence score; SGD, Singapore dollars.

[a] Forty percent of participants from prospective randomized trial contributed to retrospective gene-expression profiling (GEP) analysis.

[b] Only one study contributed data for this outcome which limits our ability to determine overall consistency of results.

[c] Indirectness in the overall study design which did not have a non-GEP test control group and did not make treatment decisions according to GEP test result.

[d] Confidence intervals around relative effect estimates indicate significant uncertainty in estimated treatment effects between CHT treated participants and non-CHT treated participants for all three RS groups.

[e] Thirty percent of participants from prospective randomized trial contributed to retrospective GEP analysis.

[f] Indirectness in the participant populations within and between studies as well as in criteria used to determine initial treatment compared to GEP-test based treatment decisions.

[g] Inconsistency in the comparator used to estimate the ICER for Oncotype DX.

Table 3

Summary of findings table for MammaPrint

Quality assessment							Summary of findings					
							Events (%)				Anticipated absolute effects	
Participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Quality of evidence	Risk	CHT	Non-CHT	Relative effect (95% CI)	Risk in controls (non-CHT)	Risk difference (95% CI)
Disease-free survival—10 years												
0 (0)						No evidence	Low:	NR	NR			
							High:	NR	NR			
Disease-free survival—5 years												
0 (0)						No evidence	Low:	NR	NR			
							High:	NR	NR			
Distant recurrence—10 years												
0 (0)						No evidence	Low:	NR	NR			
							High:	NR	NR			
Distant recurrence—5 years												
0 (0)						No evidence	Low:	NR	NR			
							High:	NR	NR			
Overall survival—10 years												
0 (0)						No evidence	Low:	NR	NR			
							High:	NR	NR			
Overall survival—5 years												

Quality assessment							Summary of findings					
							Events (%)				Anticipated absolute effects	
Participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Quality of evidence	Risk	CHT	Non-CHT	Relative effect (95% CI)	Risk in controls (non-CHT)	Risk difference (95% CI)
0 (0)						No evidence	Low:	NR	NR			
							High:	NR	NR			
Change in treatment decisions												
563 (2 studies)	Undetected	No serious inconsistency	Serious[a,b]	No serious imprecision	Undetected	Low	Not estimable			Not estimable	Not estimable	Not estimable
Economic evaluation												
Six studies	Undetected	Serious inconsistency[c]	No serious indirectness	Serious imprecision[d]	Undetected	Low	Not estimable			ICER/QALY gained[c]: $716 to £53,058	Costs: €7,430 decrease to €1,130 increase	QALYs gained[c]: −0.21 to 1.20

CHT, chemotherapy; CI, confidence interval; ICER, incremental cost-effectiveness ratio; NR, not reported; QALY, quality-adjusted life year.

[a] Indirectness in the overall study design which did not have a non-gene-expression profiling (GEP) test control group and did not make treatment decisions according to GEP test result.

[b] Indirectness in the participant populations within and between studies as well as in criteria used to determine initial treatment compared to GEP-test based treatment decisions.

[c] Inconsistency in the comparator used to estimate the ICER for MammaPrint.

[d] Imprecision in estimates of treatment effects across risk groups determined by MammaPrint add to variability to the economic impact.

Table 4

Change in treatment recommendation based on Oncotype DX and MammaPrint

Study	Design	Population	Treatment	Results
Oncotype DX				
Asad (2008)[27]	Retrospective chart review	ER+ LN− Mean age: 54 years *N* = 85	CHT for high risk based on NCCN guidelines; and HT for low risk	GEP testing influenced CHT treatment decision in 37 (44%) of patients; 34% reduction in CHT recommendations
Rayhanabad (2008)[28]	Retrospective chart review	ER+ LN− Tumors ≤ 5 cm Mean age: 54 years (range: 26–78) *N* = 58	CHT for high risk based on NCCN guidelines; and HT for low risk	Results from GEP testing led to change in treatment decisions in 15 (26%) patients
Geffen (2009)[30]	Prospective study	LN− *N* = 25	Not reported	Nine patients (36%) had their treatment recommendations changed following GEP testing, including six patients from CHT to no CHT
Henry (2009)[29]	Retrospective study	ER+ LN− *N* = 29	Discretion of medical oncologist; clinical data including Adjuvant! Online (AOL) risk estimates followed by RS	Results from GEP testing altered CHT decisions in 9/29 (31%) patients—seven from CHT to no CHT and two from no CHT to CHT with low RS
Lo (2010)[31]	Prospective, multicenter study	ER+ LN− Mean age: 55 years (range: 35–77) *N* = 89	Clinician treatment recommendation before and after GEP testing	Oncologist treatment decision changed in 28 (32%) of patients following GEP testing. 20 of these were from CHT to HT. 24 (27%) patients changed their own treatment decision, of whom 9 changed from CHT to HT, 7 from HT to CHT, 2 from undecided to HT, and 2 from undecided to CHT
Ademuyiwa (2011)[32]	Retrospective, consecutive series	ER+ LN− HER− Mean age: 54.8 years (range: 29–82) *N* = 276	CHT recommendations based on clinicopathological characteristics	RS led to change in treatment in 38% of patients, with 37 (13%) fewer patients receiving CHT
Holt (2011)[49]*	Prospective cohort	ER+	Nottingham Prognostic Index	RS led to change in treatment recommendations in 35 (33%)

Study	Design	Population	Treatment	Results
				patients, with 25 (23.5%) changing from CHT to no CHT
		LN− or N1 *N* = 106		
Oratz (2007)[34]	Retrospective study	*N* = 74	Clinician treatment recommendation before and after GEP testing	RS led to change in treatment recommendations in 21% of patients, and in actual treatment received for 25% of patients
Klang (2010)[35]	Retrospective study	*N* = 313	Clinician treatment recommendation before and after GEP testing	RS led to change in treatment recommendations in 40% of patients; 27% reduction in CHT recommendations
Hornberger (2011)[36]	Retrospective study	*N* = 952	Treatment based on guideline recommendations	RS led to a 27% reduction in CHT
Joh (2011)[37]	Retrospective study	*N* = 154	Clinician panel	RS led to a 25% change in treatment recommendations
Partin (2011)[38]	Retrospective study	*N* = 169	Treatment recommendations based on AOL and St. Gallen	RS led to a change in treatment recommendation in 27–74% of patients depending on comparator guideline
MammaPrint				
Bueno-de-Mesquita (2007)[39]	Prospective multicenter study	ER+/−	Initial CHT recommendations based on Dutch Institute for Healthcare Improvement (CBO) guidelines	Use of guidelines + prognosis signature + patients' preferences led to an actual change of treatment for 19% of patients, with a 14% overall increase in adjuvant treatment (2% more CHT, 6% more HT, and 6% more CHT + HT)
		LN+/− Mean age: 48 years *N* = 427		
Gevensleben (2010)[40]	Consecutive cohort	ER+/−	Not reported	GEP testing showed 40% of patients were either over- or undertreated
		LN+/− *N* = 136/140 had clinical treatment recorded		

CHT, chemotherapy; ER, estrogen receptor; GEP, gene-expression profiling; HT, hormone therapy; LN, lymph node; NCCN, National Comprehensive Cancer Network; RS, recurrence score.

*Results from conference abstract.

Table 5

Economic analyses of Oncotype DX and MammaPrint

Study	Population	Country	Comparator	QALYs gained	ICER (per QALY gained)	Cost results
Oncotype DX						
Homberger (2005)[41]	ER+	USA	NCCN guidelines	8.6 per 100 patients	Oncotype DX dominates	$2,028 decrease per patient
	LN−					
Lyman (2007)[42]	ER+	USA	Tamoxifen	0.97	$4,432	$4,272 increase in direct costs
	LN−		Tamoxifen + chemotherapy	1.71	Oncotype DX dominates	$2,256 decrease in direct costs
Kondo (2008)[43]	ER+	Japan	NCCN guidelines	0.097	¥2,997,495	¥289,355 increase in direct costs
	LN−		St. Gallen guidelines	0.237	¥1,239,005	¥293,211increase in direct costs
Cosler (2009)[44]	ER+	USA	Tamoxifen	2.2 life years gained	$1,944 per life year gained	$4,272 increase in direct costs
	LN−		Tamoxifen + chemotherapy	No difference	Oncotype DX dominates	$2,256 decrease in direct costs
de Lima Lopes (2010)[45]***	ER+	Singapore	Current practice	—	—	SGD 5,528 decrease in direct costs
	LN−					
Klang (2010)[35]	ER+	Israel	Traditional treatment	0.17	$10,770	$1,828 increase per patient
	LN−					
OHTA (2010)[15]	ER+	Canada	Clinical practice	13.34	—	—
	LN−					
	HER2−		Oncotype DX – AOL high-risk patients	13.03	CAD 518	—
	50 years		Oncotype DX – AOL int./ high-risk patients	14.42	CAD 795	—
			Oncotype DX – all patients	14.64	CAD 23,983	—
O'Leary (2010)[46]***	LN−/+	Australia	Conventional treatment	0.098	AUD 9,986	AUD 974 increase in direct costs

Study	Population	Country	Comparator	QALYs gained	ICER (per QALY gained)	Cost results
Tsoi (2010)[47]	ER+ LN− HER2−	Canada	AOL	0.065	CAD 63,054	CAD 4,102 increase in direct costs
de Lima Lopes (2011)[48]***	ER+ LN−	Singapore	Current practice	0.12*; 0.15**	—	SGD 5,138 decrease in direct costs
Holt (2011b)[49]***	ER+ LN 0–3	UK	Current practice	0.14	£6,232	£888 increase in direct costs
Hornberger (2011)[36]	ER+ LN−	USA	NCCN guidelines	0.162	Oncotype DX dominates	>$1 million decrease in costs to health plans
Kondo (2011)[50]	ER+ LN−/+ ER+ LN−	Japan	St. Gallen guidelines	0.47 0.63	¥568,533 ¥384,828	¥270,035 increase in direct costs ¥240,683 increase in direct costs
Ward (2011)[16]	ER+ LN− HER2− < 75 years	UK	Current practice (based on cancer registry data)	13.54	£26,940	—
Hassan (2011)[51]***	ER+ LN− HER2−	Canada	—	—	—	CAD 34.5 million cost savings
Lacey (2011a)[52]***	ER+ LN−	Ireland	Current practice	0.12	€9,462	€1,139 increase in direct costs
Lacey (2011b)[53]***	LN−	Ireland	Current practice	—	—	0.4% increase in direct costs

Study	Population	Country	Comparator	QALYs gained	ICER (per QALY gained)	Cost results
Paulden (2011)[54]***	ER+	Canada	AOL low risk	—	CAD 29,000	—
	LN−		AOL high risk	—	Oncotype DX dominates	—
Ragaz (2011)[55]***	ER+ LN−/+	Canada and USA	—	—	—	$330.8 million cost savings in USA; $46.2 million cost savings in Canada
Vanderlaan (2011)[56]	ER+	USA	Current practice	0.127	—	$384 decrease in direct costs
	LN 1–3					
Hall (2012)[57]	ER+	UK	Current practice	0.16	£5,529	£860 increase in direct costs
	LN+					
Lamond (2012)[58]	LN−/+	Canada	Current practice	0.18	CAD 10,316	CAD 1,852 increase in direct costs
	LN−			0.27	CAD 9,591	CAD 2,585 increase in direct costs
	LN+			0.06	CAD 14,844	CAD 864 increase in direct costs
Madaras (2012)[59]***	ER+	Hungary	Current practice	—	€6,871	—
	LN−					
Wilson (2012)[60]***	ER+	Ireland	Current practice	—	—	€666,844 cost savings if chemotherapy given to high-risk patients only
	LN−					
	HER2−					
MammaPrint						
Oestreicher (2005)[64]	ER+	Netherlands	National Institutes of Health guidelines	−0.21	—	$2,882 decrease in direct and indirect costs
	LN−/+					
Zarca (2009)[65]***	LN 1–2	France	Current practice	—	—	€9,043 cost savings per 100 patients per year
Chen (2010)[63]	ER+	USA	St. Gallen guidelines	0.23	$6,167	$1,332 increase in direct costs per

Study	Population	Country	Comparator	QALYs gained	ICER (per QALY gained)	Cost results
						patient
	LN−					
	HER2−					
	ER−			−0.098	—	$1,811 increase in direct costs per patient
	LN−					
	HER2−					
	SEER registry			0.571	$716	$401 increase in direct costs per patient
	Overall population			0.15	$9,428	$1,440 increase in direct costs per patient
Retel (2010)[61]	ER+	Netherlands	St. Gallen guidelines	1.20	MammaPrint dominates	€7,430 decrease in direct costs
	LN−		AOL	0.24	€4,614	€1,130 increase in direct costs
Kondo (2012)[62]	ER+	Japan	St. Gallen guidelines	0.048 years gained	¥4,820,813 per life year gained	¥231,385 increase in societal costs per patient
	LN−					
	HER2−					
Ward (2011)[16]	ER+	UK	Current clinical practice	13.47–13.78	£12,240–£53,058	
	LN−					
	HER2−					
	< 75 years					

AOL, Adjuvant! Online; AUD, Australian dollars; CAD, Canadian dollars; ER, estrogen receptor; ICER, incremental cost-effectiveness ratio; LN, lymph node; NCCN, National Comprehensive Cancer Network; SGD, Singapore dollars; QALY, quality-adjusted life year.

*QALYs gained from avoiding chemotherapy.

**QALYs gained by preventing future recurrence.

***Results from conference abstract.