



This is a repository copy of *Gene expression profiling and expanded immunohistochemistry tests to guide selection of chemotherapy regimens in breast cancer management: a systematic review.*

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/113359/>

Version: Accepted Version

---

**Article:**

Scope, A. [orcid.org/0000-0003-1604-1758](https://orcid.org/0000-0003-1604-1758), Essat, M., Pandor, A. et al. (5 more authors) (2017) Gene expression profiling and expanded immunohistochemistry tests to guide selection of chemotherapy regimens in breast cancer management: a systematic review. *International Journal of Technology Assessment in Health Care*, 33 (1). pp. 32-45. ISSN 0266-4623

<https://doi.org/10.1017/S0266462317000034>

---

This article has been published in a revised form in *International Journal of Technology Assessment in Health Care* [<https://doi.org/10.1017/S0266462317000034>]. This version is free to view and download for private research and study only. Not for re-distribution, re-sale or use in derivative works. © Cambridge University Press 2017.

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

**Title:** Gene expression profiling and expanded immunohistochemistry tests to guide selection of chemotherapy regimens in breast cancer management: a systematic review.

**Authors:** Alison Scope<sup>1\*</sup>, Munira Essat<sup>1</sup>, Abdullah Pandor<sup>1</sup>, Rachid Rafia<sup>1</sup>, Sue. E. Ward<sup>1</sup>, Lynda Wyld<sup>2</sup>, Simon Cross<sup>2</sup>, & Helen Buckley Woods<sup>1</sup>.

School of Health and Related Research (SchARR), The University of Sheffield, Sheffield, UK.<sup>1</sup>

Royal Hallamshire Hospital, Sheffield, UK.<sup>2</sup>

\*Corresponding Author: School of Health and Related Research (SchARR), Regent Court, 30 Regent Street, The University of Sheffield, Sheffield. S1 4DA, UK.

[A.scope@sheffield.ac.uk](mailto:A.scope@sheffield.ac.uk). Telephone: (44)1142220670. Fax: (44)1142724995.

Running Head: GEP and expanded IHC tests in breast cancer management.

## Abstract

**OBJECTIVES:** The aim of this report was to assess the clinical effectiveness of two Gene expression profiling (GEP) and two expanded immunohistochemistry (IHC) tests compared with current prognostic tools in guiding the use of adjuvant chemotherapy in patients with early breast cancer.

**METHODS:** A systematic review of the evidence on clinical effectiveness of OncotypeDX, IHC4, MammaPrint and Mammostrat, compared with current clinical practice using clinicopathological parameters, in women with early breast cancer was conducted. Ten databases were searched to include citations to May 2016.

**RESULTS:** Searches identified 7064 citations, of which 41 citations satisfied the criteria for the review. A narrative synthesis was performed. Evidence for OncotypeDX demonstrated the impact of the test on decision-making and there was some support for OncotypeDX predicting chemotherapy benefit. There were relatively lower levels of evidence for the other three tests included in the analysis. MammaPrint, Mammostrat and IHC4 tests were limited to a small number of studies. Limitations in relation to study design were identified for all tests.

**CONCLUSIONS:** The evidence base for OncotypeDX is considered to be the most robust. Methodological weaknesses relating to heterogeneity of patient cohorts and issues arising from the retrospective nature of the evidence were identified. Further evidence is required for all of the tests using prospective randomised controlled trial data.

**Keywords:** Gene expression profiling; Immunohistochemistry; breast cancer.

## **Acknowledgements**

This is an updated version of a review which was originally undertaken to inform the UK National Institute for Health and Clinical Excellence's (NICE) assessment of GEP (MammaPrint, OncotypeDX, Mammostrat) and IHC (IHC4) tests to guide selection of chemotherapy regimens in breast cancer management. Funding for that study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

## **Introduction**

Breast cancer is not a single disease but rather a group of heterogeneous tumours at the molecular level(1). Based on the knowledge that certain biological features of cancers may indicate an increased likelihood of rapid growth and metastasis (in particular, distant recurrence) gene expression profiling (GEP) and expanded immunohistochemistry (IHC) (or protein expression) tests have been developed. These tests have an aim of improving the targeting of chemotherapy in breast cancer by stratifying patients and identifying those patients who will gain most benefit from adjuvant chemotherapy. These tests either measure the risk of cancer recurrence (by incorporating a wider range of biomarkers with prognostic significance than standard clinico-pathological algorithms), or aim to identify breast cancer sub-types which may influence recurrence risk and guide treatment decisions.

In current practice treatment regimens are tailored according to traditional clinical characteristics such as age, tumour size and grade together with a tumour's molecular signature based on estrogen (ER) and progesterone (PR) receptor status and HER2 receptor status(2), although guidelines may differ slightly from country to country.

The purpose of this systematic review was to evaluate the clinical effectiveness of GEP and expanded IHC tests in guiding the use of adjuvant chemotherapy in women with early breast cancer. A summary of the evaluated gene expression profiling and expanded immunohistochemistry tests is presented in Table 1. This review was originally undertaken to inform the UK National Institute for Health and Care Excellence's (NICE) assessment of GEP (MammaPrint, OncotypeDX) and IHC

(IHC4 and Mammostrat) tests to guide selection of chemotherapy regimens in breast cancer management(3), but has been updated with new evidence up to May 2016.

### **Table 1 here**

#### **Method**

A systematic review of the evidence was undertaken according to the general principles recommended in the Centre for Reviews and Dissemination (CRD)(4) guidance for undertaking systematic reviews, and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement(5), and The NICE Diagnostic Assessment Programme Interim Methods Statement(6).

#### Data sources and searches

Ten electronic databases were searched, these were: Medline and Medline in Process via Ovid SP, Embase via Ovid SP; Cochrane Library databases all via Wiley: Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Database of Systematic Reviews (CDSR), Database of Abstracts of Reviews of Effects (DARE) , Health Technology Assessment Database (HTA), NHS Economic Evaluation Database(NHS-EED); Web of Science databases all via Thomson Reuters: BIOSIS Previews, Science Citation Index Expanded (SCI-Expanded) and the Conference Proceedings Citation Index- Science (CPCI-S).

The search strategy used free text and thesaurus terms and combined breast cancer related synonyms (e.g. breast neoplasm) with terms related to gene expression profiling tests or biomarkers (e.g. MammaPrint or “gene?twentyone”). A publication

date limit of January 2002 was applied. This was the date that the longest standing test used in the review had been devised, as confirmed by manufacturers' submissions to NICE as part of the original review, and therefore it would not be possible for evidence to predate this. For the OncotypeDX and MammaPrint test, the current review used two previous systematic reviews(7,8) to identify included studies, thus the searches were limited from January 2009 (last date from earlier reviews) for these tests. Although a number of other systematic reviews examining GEP tests have been reported, these reviews(7,8) were considered the most appropriate reviews to update. The reviews were assessed as being of high quality, and in particular the search strategies were assessed as being complete. No other limits were applied to the searches. An update search was conducted in Medline and Medline in Process from January 2013 - May 2016.

Supplementary search techniques were also undertaken to augment the topic searches, these included hand searching of relevant journals, citation searches of included papers in the review, searching of conference proceedings, and finally experts in the field were contacted to ask for suggestions for relevant evidence for the project.

### Study selection

The inclusion of potentially relevant articles was undertaken using a two-stage process. First, all titles and abstracts were screened for inclusion, followed by the assessment of full manuscripts. Both stages were undertaken by one reviewer and any uncertainties in the selection process were resolved through discussion with another reviewer. All study designs were included. Eligible studies included adult patients diagnosed with early invasive breast cancer. The index test included

OncotypeDX, MammaPrint, IHC4 or Mammostrat. The comparator was standard care and could include the use of Adjuvant! Online (AoL) and/or the Nottingham Prognostic Index (NPI), to predict the risk of recurrence and survival for patients with early breast cancer. The outcome measure was clinical utility (the test's ability to discriminate between those who will have more or less benefit from a therapeutic intervention) (7,8). Specifically, (i) the ability of the test to predict treatment effect with adjuvant chemotherapy, and (ii) to what extent are test results used in treatment decisions. Studies published in languages other than English (unless no other comparable data existed) were excluded. Abstracts were considered but only included if they represented significant new knowledge, such as prospective RCT evidence.

#### Data extraction and quality assessment

Data relating to study design, methodological quality, and outcomes, were extracted by one reviewer into a standardised data extraction form and independently checked for accuracy by a second. Discrepancies were resolved by discussion. The methodological quality of each included study was assessed by two reviewers according to the criteria recommended by Altman (2001)(9) for assessing the internal validity of prognostic (predictive factor) studies.

#### Data synthesis and analysis

Although a meta-analysis was planned, this was not considered appropriate due to a high degree of heterogeneity e.g. study populations, outcomes and diagnostic



thresholds between and within studies. Therefore, data were tabulated and discussed in a narrative review.

## **Results**

### PRISMA flow

Figure 1 summarises the process of identifying and selecting relevant literature. Of the 7064 citations identified, 29 new studies (30 citations) were identified and were added to the 11 studies from the previous systematic reviews.

### **Figure 1 about here**

#### Study and patient characteristics

Forty studies (41 citations) were included in the review. All studies were published between 2002 and May 2016.

Most of the evidence was related to the OncotypeDX (32 studies). Four studies related to the prediction of treatment effect with adjuvant chemotherapy, with the remaining 28 studies relating to evidence on the test result leading to changes in treatment decisions. Six studies were identified for MammaPrint, all relating to evidence on the test result leading to changes in treatment decisions. Only one relevant study was identified for IHC4, and one for Mammostrat. The IHC4 study provided evidence relating to the test leading to changes in treatment decisions, whereas the Mammostrat study provided evidence on the prediction of treatment effect with adjuvant chemotherapy. Details of the study and patient characteristics, together with key findings of the included studies are provided in Tables 2 and 3.

## **Quality Assessment**

Limitations in the clinical data were identified for all tests. No studies had a prospective, randomised controlled trial (RCT) design and only five studies included a prospective analysis of archived tissue samples from a previous RCT (OncotypeDX n=4; Mammostrat n=1). For the four OncotypeDX studies and the one Mammostrat study providing evidence relating to the prediction of treatment effect with adjuvant chemotherapy, the overall risk of bias was judged to be moderate, although retrospective analysis of archived tissue samples, the evidence was derived from relatively large scale RCTs. The remaining 28 OncotypeDX studies providing evidence relating to changes in treatment recommendations, were in the main, small scale studies (n=25-979). Fifteen were retrospective in study design, and some (n=14) did not provide full details of the patient characteristics. Similarly, of the six studies identified for MammaPrint two were retrospective in study design, and some were lacking full details of patient characteristics. The IHC4 study was prospective in design, however the sample size was relatively small (n=124). Overall, particularly for the studies relating to evidence of the tests leading to changes in treatment decisions, there was a high level of clinical heterogeneity across studies both within each test and across the four tests.

**Table 2 here**

**Table 3 here**

## **Narrative data synthesis**

Prediction of treatment effect with adjuvant chemotherapy

OncotypeDX

Studies by Paik, Tang and Shak et al.(10), Albain, Barlow and Shak et al.(11), Tang, Shak and Paik et al.(12) and Tang, Constantino and Crager et al.(13) assessed the predictive ability of OncotypeDX using archived tissue samples collected during randomised controlled trials comparing tamoxifen with tamoxifen plus chemotherapy. The strongest evidence appeared to be presented by Paik, Tang and Shak et al.(10). The OncotypeDX recurrence score was found to be correlated with chemotherapy benefit, defined in terms of 10-year DRFS, with a significantly increased benefit from the use of chemotherapy in the OncotypeDX high-risk group compared with the low-risk group, in ER+, LN- breast cancer patients. However, in a multivariate analysis the benefit from chemotherapy was unclear due to large confidence intervals in the low and intermediate RS risk groups. Albain, Barlow and Shak et al.(11) demonstrated that the RS was prognostic for tamoxifen-treated patients with positive nodes and predicts significant benefit of chemotherapy in tumours with a high recurrence score. They concluded that a low score could identify women who might not benefit from anthracycline-based chemotherapy, despite positive nodes. It was also reported by Tang, Shak and Paik et al.(12) that both RS and AoL provided strong independent prognostic information in tamoxifen treated patients, and that RS used alone remained the best predictor of chemotherapy benefit in ER+, LN- breast cancer(13).

Of these four studies reporting evidence that OncotypeDX predicts benefit from chemotherapy, only one, on a LN+ population(11) presented that had not come from the NSABP cohorts. However, there were limitations associated with this study. It had only a moderate sample size, and the time over which tumour samples were collected was not reported, therefore there may be differences in diagnostic criteria

being applied. Two other studies(12,13) reported the same trial data as Paik, Tang and Shak et al.(10) from the NSABP cohorts, introducing biases associated with double counting in the evidence base as a whole. It should further be noted that the Paik, Tang and Shak et al.(10) study may also have been subject to bias, as some patients in the validation dataset were also in the training dataset which may partly explain the treatment interaction seen with OncotypeDX.

### Mammostrat

No prospective studies of the impact of Mammostrat on long-term outcomes such as overall survival were identified. Initial evidence for the predictive ability of Mammostrat from one study(14) suggests that low and high-risk groups benefited from chemotherapy, with high-risk patients benefiting more than low- risk. The intermediate-risk group did not appear to benefit.

### Changes in treatment recommendations as a result of testing

#### OncotypeDX

Twenty-eight studies (see table 3) provided evidence on the impact of OncotypeDX on clinical decision-making. These studies indicated that the use of OncotypeDX leads to changes in treatment recommendations for between 21% and 74% of all patients who underwent OncotypeDX testing. Three studies (17,24,25) did not report whether changes led to increased or decreased use of chemotherapy. However, where this was reported the number of patients being recommended chemotherapy after the test was introduced declined in most studies. This change from chemotherapy to no chemotherapy ranged from 6% to 51.4% of all patients tested. However, in one study more chemotherapy was used after the introduction

of OncotypeDX(28). It was not clear in a large number of the studies whether these figures represented actual changes in the treatments patients received.

### MammaPrint

Six studies were identified which provided evidence on changes in treatment recommendations as a result of MammaPrint (see table 3). These studies indicated that the use of MammaPrint in addition to clinicopathological factors led to changes in treatment recommendations for between 18% and 40% of all patients tested, and that the between 2% and 32% of all patients would be recommended to change from chemotherapy to no chemotherapy. One of these studies(45) reported the use of MammaPrint compared to AoL would result in altered treatment advice for 40% of patients. However, this was based on the assumption that all patients classified as high-risk would receive chemotherapy and patients classified as low risk would not receive chemotherapy. Again, in a number of these studies it is not clear if actual treatment changes occurred following introduction of the test.

A prospective observational study(43) showed that adjuvant treatment was recommended for 48% of patients based on, and Dutch Institute for Healthcare Improvement (CBO) guidelines (2004) alone, increasing to 62% when MammaPrint was added. This increased the number of patients receiving adjuvant systemic therapy by 20 (5%). For the other guidelines assessed (St Gallen guidelines, the NPI and AoL), less adjuvant chemotherapy would be given when the data was based on prognostic signature alone are used. A 5 year follow up study (44) showed that 15% of the MammaPrint low risk patients received adjuvant chemotherapy versus 81% of the high-risk patients. The 5 year distant recurrence free interval (DRFI) probabilities for MammaPrint low-risk patients were 97%, and 91.7% for the high-risk

patients. Actual treatment decisions were based on restrictive CBO guidelines, and doctors and patients preferences limiting the generalisability of these findings.

#### IHC4

Evidence from one prospective study (50) demonstrated that the IHC4 test led to changes in treatment recommendations for 34% of the patients, with 25% recommended to switch from chemotherapy to no chemotherapy. As there is only one study available and it has a small sample size (n=124), it is difficult to make generalisations based on this evidence. Again, it is not clear whether actual treatment given was changed.

#### **Discussion**

OncotypeDX currently has the largest body of evidence on clinical utility relative to the other three tests included in this review. Although, no prospective studies reporting the impact of OncotypeDX on long term outcomes, such as overall survival, yet exist. The Paik, Tang and Shak et al.(10) study represented the most robust evidence of clinical utility. The study showed a decreased relative benefit of chemotherapy in the lower-risk groups. However, the specific cancers in the low-risk groups were less likely to respond to chemotherapy, independent of actual survival probability. Other specific limitations include that fact that in one study (32), compared to the study regimens, more effective chemotherapy regimens are currently being used, and more than 44% of patients were aged below 50 years old, limiting the generalisability of the findings.

The evidence base for MammaPrint, is primarily based on small sample sizes ( $n < 427$ ). Some studies were retrospective in design and had heterogeneous patient populations. Some studies included only pre-menopausal women, which may overestimate the benefit of MammaPrint in the early breast cancer population as a whole, given that younger women are likely to be at higher risk of recurrence and are more likely to be classified as poor prognosis using MammaPrint. Further evidence is required to clarify whether using the test will improve the use of adjuvant chemotherapy in the management of breast cancer. It is also unclear to what extent MammaPrint risk groups are predictive of chemotherapy benefit or how the use of MammaPrint will improve patient outcomes through increases in disease-free and overall survival.

One study on Mammostrat(14) provides evidence relating to the benefit of chemotherapy by risk group. However, this indicates that both low and high-risk groups benefit, whilst it is unclear how those in the moderate risk group would be affected. Further evidence is required. In particular there was no published evidence on the impact of the test on decision-making.

One clinical utility study was available for IHC4(50). This study provided evidence on the impact of the test on decision-making leading to reductions in the amount of chemotherapy recommended. Although the design was prospective it included a relatively small sample of patients.

#### Limitations

The varied nature of the evidence base makes comparisons between tests difficult. A characteristic feature of the studies across all tests was their heterogeneity, and a large proportion of the studies were small. Many studies used old archived tumour

samples, and some, retrospective chart review to elicit treatment recommendations before and after testing. There was a lack of standardised decision-making tools both within and between studies and non-standardised methods of patient selection were used. Furthermore, a number of the studies for OncotypeDX and MammaPrint were funded by the manufacturer giving rise to potential issues of conflict of interests and publication bias.

### Conclusion and Implications

One of the tests (OncotypeDX) has a reasonably large evidence base, although there are some methodological weaknesses relating to this evidence, in terms of heterogeneity of patient cohorts, and retrospective study design. The previous systematic reviews(7,8) on which our updates were based reported that OncotypeDX was furthest along the validation pathway, and that recurrence score was significantly correlated with disease-free-survival and overall survival. There was also some evidence that there may be a significant benefit from the use of chemotherapy in the OncotypeDX high-risk group, although it was acknowledged that this study may have been subject to bias. Our previous review (3) and this update demonstrates that further larger studies have now reported, which support the prognostic capability of the OncotypeDX test, and in the evidence base has been extended to include the LN+ population. Also, further studies have presented evidence on the impact of OncotypeDX on clinical decision-making. The previous reviews(7,8) indicated that evidence relating to the clinical validity of MammaPrint was not always conclusive or supportive of the prognostic value of the test, and one study was identified which suggested that MammaPrint had an impact on clinical decision-making. Our previous review (3), together with this update identified studies which showed the MammaPrint score is a strong independent prognostic



factor and may provide additional value to standard clinicopathological measures, although the populations in all of these studies were relatively small. Further studies on the clinical utility of MammaPrint reported on test reclassification against currently used guidelines, reporting that treatment advice for a percentage of patients may change. However, none of the studies provided evidence of actual changes in treatment decisions following introduction of the test.

This update has demonstrated that in comparison to our original review(3) a number of new studies have emerged which assess the effect of the tests on clinical decision making. However, most of these studies are small scale and it remains the case that further robust evidence on the clinical utility of all of these tests is needed. This would include studies investigating predictive ability, and prospective studies investigating how the tests will be used in clinical practice. Two ongoing trials relating to OncotypeDX(51) and MammaPrint(52) have been designed to address some of these issues, specifically relating to the effect of these tests on patient outcomes and their ability to predict treatment response. The TAILORx trial(51) aims to demonstrate that endocrine treatment alone is non-inferior to chemoendocrine treatment in women with an intermediate OncotypeDX score. Patients allocated to an intermediate risk group using the recurrence score will receive endocrine therapy and be randomly assigned to chemotherapy or no chemotherapy. The MINDACT trial(52) aims to assess the value of MammaPrint in predicting which patients would benefit from chemotherapy compared with Adjuvant! Online. Patients assessed as high risk by one method and low risk by the other will then be randomised to follow the treatment indicated by MammaPrint or the treatment indicated by Adjuvant! Online. Two further objectives of the trial relating to the efficacy of different chemotherapy agents and endocrine treatment strategies are

addressed by two further stages of randomisation. These trials will result in direct evidence that these tests in breast cancer patients lead to improvement in outcomes with the use of RCTs comparing the outcomes of patients following standard management to those of patients managed with the aid of the expression-based assays. All tests would benefit from further evidence demonstrating how they will be used in the current decision-making process and, especially, how this will impact on patient management decisions.

## References

- (1) Perou CM, Sørli T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747–52.
- (2) Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E, et al. on behalf of the ESMO Guidelines Committee. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology*. 2015;26 (Supplement 5): v8–v30.
- (3) Ward S, Scope A, Rafia R, Pandor A, Harnan S, Evans P, 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 Technology Assessment*. 2013; 17(44).
- (4) Centre for Reviews and Dissemination, *Systematic reviews: CRD's guidance for undertaking reviews in healthcare*. Centre for Reviews and Dissemination. 2009.
- (5) Liberati A, Altman D, Tetzlaff J, Mulrow C, Gotzsche PC, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of

studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*. 2009, Vol. 339, No. b2700.

- (6) National Institute for Health and Care Excellence, NICE Diagnostic Assessment Programme interim methods statement (pilot), NICE, London, 2010.
- (7) Marchionni L, Wilson R, Marinopoulos S, Wolff A, Parmigiani G, Bass E, et al. Impact of gene expression profiling tests on breast cancer outcomes. Evidence Report/Technology Assessment No. 160. Rockville, MD: Agency for Healthcare Research and Quality; 2008.
- (8) Smartt P. A comparison of gene expression profiling tests for breast cancer. HSAC Report. 2010; 3(16). Christchurch, New Zealand: Health Services Assessment Collaboration.
- (9) Altman D. Systematic reviews in health care: systematic reviews of evaluations of prognostic variables. *BMJ* 2001;323:224–8.
- (10) Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 2006;24:3726–34.
- (11) Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, 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..
- (12) Tang G, Shak S, Paik S, Anderson SJ, Costantino JP, Geyer CE, 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;127:133–42.
- (13) Tang G, Constantino J, Crager M, Shak S, Wolmark N. [S4–9] Comparing the prediction of chemotherapy benefit in patients with node-negative, ER-positive breast cancer using the recurrence score and a new measure that integrates clinical and pathologic factors with the recurrence score. *Proceedings of the 33rd Annual San Antonio Breast Cancer Symposium*, 8–12 December, 2010.
- (14) Ross D, Kim C-Y, Tang G, Bohn O, Beck R, Ring B, et al. Chemosensitivity and stratification by a five monoclonal antibody immunohistochemistry test in the NASBP B14 and B20 trials. *Clin Cancer Res* 2008;14:6602–9.
- (15) Oratz R, Paul D, Cohn A, Sedlacek S. Impact of a commercial reference laboratory test recurrence score on decision making in early-stage breast cancer. *J Oncol Pract* 2007;3:182–6.
- (16) Asad J, Jacobson A, Estabrook A, Smith S, Boolbol S, Feldman S, et al. Does Oncotype DX recurrence score affect the management of patients with early-stage breast cancer? *Am J Surg* 2008;196:527–9
- (17) Rayhanabad J, Difronzo L, Haigh P, Romero L. Changing paradigms in breast cancer management: introducing molecular genetics into the treatment algorithm. *Am Surg* 2008;74:887–90.
- (18) Geffen DB, Amir N, Sion-Vardy N, Ariad S, Kachko L, Bayme M, 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;18:316–21.

- (19) Henry L, Stojadinovic A, Swain S, Prindiville S, Cordes R, Soballe PW. The influence of a gene expression profile on breast cancer decisions. *J Surg Oncol* 2009;99:319–23.
- (20) 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.
- (21) 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.
- (22) 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.
- (23) Geffen DB, Abu-Ghanem S, Sion-Vardy N, Braunstein R, Tokar M, Ariad, S, et al. The impact of the 21-gene recurrence score assay on decision making about adjuvant chemotherapy in early-stage estrogen-receptor-positive breast cancer in an oncology practice with a unified treatment policy. *Ann Oncol.* 2011; 22:2381-2386.
- (24) Joh JE, Esposito NN, Kiluk JV, Laronga C, Lee MC, Loftus L, 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

- (25) 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.
- (26) Albanell J, Gonzalez A, Ruiz-Borrego M, Garcia-Saenz JA, Corminas JM, Burques O, et al. Prospective transGEICAM study of the impact of the 21-gene Recurrence Score assay and traditional clinicopathological factors on adjuvant clinical decision making in women with estrogen receptor positive (ER?) node-negative breast cancer. *Ann Oncol*. 2012; 23(3):625–631
- (27) Bargallo JE, Lara F, Shaw-Dulin R, Perez-Sanchez V, Villarreal-Garza C, Maldonado-Martinez H, et al. A study of the impact of the 21-gene breast cancer assay on the use of adjuvant chemotherapy in women with breast cancer in a Mexican public hospital. *J Surg Oncol*. 2015;111(2):203-7.
- (28) Biroschak JR, Schwartz GF, Palazzo JP, Toll AD, Brill KL, Jaslow RJ, et al. Impact of Oncotype DX on treatment decisions in ER-positive, node-negative breast cancer with histologic correlation. *Breast J*. 2013; 19(3):269-75.
- (29) Davidson JA, Cromwell I, Ellard SL, Lohrisch C, Gelmon KA, Shenkier T, 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*. 2013;49(11):2469-75.
- (30) de Boer RH, Baker C, Speakman D, Chao CY, Yoshizawa C, Mann GB. The impact of a genomic assay (Oncotype DX) on adjuvant treatment recommendations in early breast cancer. *Med J Aust*. 2013;199(3):205-8.

- (31) Eiermann W, Rezai M, Kummel S, Kuhn T, Warm M, Friedrichs K, et al. The 21-gene recurrence score assay impacts adjuvant therapy recommendations for ER-positive, node-negative and node-positive early breast cancer resulting in a risk-adapted change in chemotherapy use. *Ann Oncol.* 2013; 24(3):618–624
- (32) Holt, S., Bertelli, G., Humphreys, I., Valentine, W., Durrani, S., Pudney, D., et al. A decision impact, decision conflict and economic assessment of routine OncotypeDX testing of 146 women with node-negative or pN1mi, ER-positive breast cancer in the UK. *Br J Cancer.* 2013; 108:2250-2258.
- (33) Cheung PS, Tong AC, Leung RC, Kwan WH, Yau TC. Initial experience with the Oncotype DX assay in decision-making for adjuvant therapy of early oestrogen receptor-positive breast cancer in Hong Kong. *Hong Kong Med.* 2014;20(5):401-6.
- (34) Fried G, Moskovitz M. Treatment decisions in estrogen receptor-positive early breast cancer patients with intermediate oncotype DX recurrence score results. *Springerplus.* 2014;3:71.
- (35) Jaafar H, Bashir MA, Taher A, Qawasmeh K, Jaloudi M. Impact of Oncotype DX testing on adjuvant treatment decisions in patients with early breast cancer: a single-center study in the United Arab Emirates. *Asia Pac J Clin Oncol.* 2014;10(4):354-60.
- (36) Yamauchi H, Nakagawa C, Takei H, Chao C, Yoshizawa C, Yagata H, et al. Prospective study of the effect of the 21-gene assay on adjuvant clinical decision-making in Japanese women with estrogen receptor-positive, node-negative, and node-positive breast cancer. *Clin Breast Cancer.* 2014;14(3):191-7.

- (37) Gligorov JG, Pivot XB, Jacot W, Naman HL, Spaeth D, Misset JL, et al. Prospective Clinical Utility Study of the Use of the 21-Gene Assay in Adjuvant Clinical Decision Making in Women With Estrogen Receptor-Positive Early Invasive Breast Cancer: Results From the SWITCH Study. *Oncologist*. 2015;20(8):873-9.
- (38) Lee MH, Han W, Lee JE, Kim KS, Park H, Kim J, et al. The clinical impact of 21-gene recurrence score on treatment decisions for patients with hormone receptor-positive early breast cancer in Korea. *Cancer Res*. 2015;47(2):208-14.
- (39) Zhang YN, Zhou YD, Mao F, Sun Q. Impact of the 21-Gene Recurrence Score Assay in adjuvant chemotherapy selection for node-negative, hormone receptor-positive breast cancer in the Chinese population. *Neoplasma*. 2015;62(4):658-65.
- (40) Kuchel A, Robinson T, Comins C, Shere M, Varughese M, Sparrow G, et al. The impact of the 21-gene assay on adjuvant treatment decisions in oestrogen receptor-positive early breast cancer: a prospective study. *Br J Cancer*. 2016;114(7):731-6.
- (41) Levine MN, Julian JA, Bedard PL, Eisen A, Trudeau ME, Higgins B, et al. Prospective Evaluation of the 21-Gene Recurrence Score Assay for Breast Cancer Decision-Making in Ontario. *J Clin Oncol*. 2016;34(10):1065-71.
- (42) Ozmen V, Atasoy A, Gokmen E, Ozdogan M, Guler N, Uras C, et al. Impact of Oncotype DX Recurrence Score on Treatment Decisions: Results of a Prospective Multicenter Study in Turkey. *Cureus*. 2016;8(3):e522.



- (43) Bueno-De-Mesquita JM. 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(12): 1079-1087.
- (44) Drukker CA, Bueno-De-Mesquita JM, Retel VP, van Harten WH, van Tinteren H, Wesseling J, et al. A prospective evaluation of a breast cancer prognosis signature in the observational RASTER study. *Int J Cancer*, Vol. 2013; 133(4):929-936.
- (45) Gevensleben H, Gohring UJ, Buttner R, Heukamp LC, Kunz G, Dimpfl T, 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(6): 837-843.
- (46) Hartmann S, Gerber B, Elling D, Heintze K, Reimer T. The 70-Gene Signature as Prognostic Factor for Elderly Women with Hormone Receptor-Positive, HER2-Negative Breast Cancer. *Breast Care.* 2012; 7:19-24.
- (47) Cusumano PG, Generali D, Ciruelos E, Manso L, Ghanem I, Lifrange E, et al. European inter-institutional impact study of MammaPrint. *Breast.* 2014;23(4):423-8.
- (48) Drukker CA, van den Hout HC, Sonke GS, Brain E, Bonnefoi H, Cardoso F, et al. Risk estimations and treatment decisions in early stage breast cancer: agreement among oncologists and the impact of the 70-gene signature. *Eur J Cancer.* 2014;50(6):1045-54.
- (49) Exner R, Bago-Horvath Z, Bartsch R, Mittlboeck M, Retel VP, Fitzal F, et al. The multigene signature MammaPrint impacts on multidisciplinary team

decisions in ER+, HER2- early breast cancer. *Br J Cancer*.

2014;111(5):837-42.

- (50) Yeo B, Zabaglo L, Hills M, Dodson A, Smith I, Dowsett M. Clinical utility of the IHC4+C score in oestrogen receptor-positive early breast cancer: a prospective decision impact study. *Br J Cancer*. 2015;113(3):390-5.
- (51) TAILORx trial. Hormone therapy with or without combination chemotherapy in treating women who have undergone surgery for node-negative breast cancer (the TAILORx trial). 2011. URL: <http://clinicaltrials.gov/ct2/show/study/NCT00310180?term=oncotype+dx&rank=3> (last accessed 13 June 2016).
- (52) MINDACT trial. Microarray In Node negative Disease may Avoid ChemoTherapy. A prospective, randomised study comparing the 70-gene expression signature with common clinical-pathological criteria in selecting patients for adjuvant chemotherapy in nodenegative breast cancer. 2011. URL: [http://www.eortc.be/services/unit/mindact/MINDACT\\_websiteii.asp#Current](http://www.eortc.be/services/unit/mindact/MINDACT_websiteii.asp#Current) (last accessed 13 June 2016).

**Table 1: Summary of evaluated gene expression profiling and expanded immunohistochemistry tests.**

	OncotypeDX	MammaPrint	IHC4	Mammostrat
Function	Risk of recurrence	Risk of recurrence	Risk of recurrence	Subtyping and Risk of recurrence
Technology	Reverse transcription polymerase chain reaction  (21 gene)	Microarray  (70 gene)	Combining 4 IHC tests & clinical parameters to derive prognostic score	Uses 5 biomarkers to derive risk score
Location of testing	Central	Central Irvine, USA	Local (but quality assurance issues need to be addressed)	Central
Type of sample	Formulin fixed paraffin embedded	Fresh  (Use of Formulin fixed paraffin embedded has now been introduced)	Formulin fixed paraffin embedded	Formulin fixed paraffin embedded
Population	ER +, LN-, LN 1-3	ER+ (or ER-), LN-, LN 1-3, tumour size <5cm	Post menopausal. ER+, LN-	ER+, LN-, LN 1-3
Presentation of results	RS and risk group (Low <18, Intermediate 18-30, high>=31)	2 categories – low and high risk	IHC4 risk score	Risk groups - (High. >0.7, Moderate, ≤0.7, Low, ≤0.)
Cost	GBP 2,580 (EUR 3263)	GBP 2,675 (EUR 3387)	Approx GBP 100-200 (EUR 126-253)	Approx GBP 1120-1620 (EUR 1417-2049)
RS = recurrence score, ROR = risk of recurrence score, ER+ = oestrogen receptor positive, ER- = oestrogen receptor negative, LN- = lymph node negative, LN 1-3 = 1-3 lymph nodes involved, IHC = Immunohistochemistry.				

**Table 2: Summary of patient characteristics, study characteristics, and key findings relating to the prediction of treatment effect with adjuvant chemotherapy**

Author (year)	Design	Population	Treatment	Outcome measure	Treatment outcomes and Key Findings
OncotypeDX					
Paik, et al. (2006) (10)	Prospective-retrospective study in using existing trial data (NSABP B-20).	ER+, LN-, HER2+/- N=651/2229 (28.9%)	HT: tamoxifen (n=227) CHT: tamoxifen + CMF/MF (N=424).	10-year DRFS	Low RS: CHT = 4.4% / HT =3.2%; RR = 1.3 (0.46-3.78) Intermediate RS: CHT = 10.9%/HT = 9.1%; RR=0.61 (0.24-1.59) High RS:CHT = 11.9%/HT = 39.5%; RR = 0.26 (0.13-0.53) RS was correlated with chemotherapy benefit, (10-year DRFS). Significate benefit of chemotherapy in the high RS group (p = 0.001).
Albain, et al. (2010) (11)	Prospective-retrospective study from the SWOG-8814 trial.	ER+ and/or PR+, LN+ (postmenopausal) N=367/927 (39.6%)	HT: tamoxifen (N=148) CHT: CAF >tamoxifen (N=219)	10-year DFS	Low RS: HR=1.02 (0.54-1.93) Intermediate RS: HR:0.72 (0.39-1.31) High RS: HR = 0.59 (0.35-1.01)
				10-year OS	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) RS is prognostic for tamoxifen-treated patients with positive nodes and predicts significant benefit of CHT in tumours with a high recurrence score.
Tang, et al. (2011) (12)	Prospective-retrospective study from the NSABP-B14 and B20 trial.	ER+, LN- N=651 (B20 cohort)	HT: tamoxifen CHT: tamoxifen + chemotherapy	DRFI	P=0.031 for RS x treatment interaction
				OS	P=0.011 for RS x treatment interaction
				DFS	P=0.082 for RS x treatment interaction RS was significantly predictive of chemotherapy benefit. (for DRFI, for OS, and DFS), but for AoL was not. In the larger B-20 sub-cohort, AoL was significantly predictive of chemotherapy benefit for OS but not for DRFI or DFS.
Tang, et al. (2010) (Abstract only)(13)	Prospective-retrospective study from the NSABP B-20 trial.	ER+, LN- N=625	HT: tamoxifen CHT: NR	DR	HR=0.84 (P=0.037 for RS x treatment interaction). RS used alone remains the best predictor of chemotherapy benefit in ER+, N- breast cancer.
Mammostrat					
Ross,	Prospective	ER+, LN-	HT:	DRFI	Low risk: improved by 5% from 86% to 91%, HR 0.4 (95%CI:

Author (year)	Design	Population	Treatment	Outcome measure	Treatment outcomes and Key Findings
et al. (2008) (14)	retrospective study from the NSABP B14 and B20 trials	N=711	tamoxifen		0.2 – 0.8), High risk: improved by 21% from 64% to 85%, HR 0.4 (95%CI: 0.2 – 0.9), Showing that these groups benefited from chemotherapy, whereas the patients in the intermediate risk group did not.

RS = recurrence score, RR = Relative risk, HR = Hazard ratio, AoL = Adjuvant! Online, ER+ = oestrogen receptor positive, ER- = oestrogen receptor negative, LN- = lymph node negative, LN 1-3 = 1-3 lymph nodes involved, HT = hormone therapy, CHT = chemotherapy, CMF/MF, CAF = specific chemotherapy regimen, DRFS = Distant recurrence free survival, DRFI = Distant recurrence free interval, DFS = disease free survival, OS = Overall survival.

**Table 3: Summary of patient characteristics, study characteristics, and key findings relating to changes in treatment recommendations**

Author (year)	Design	Population	Prior treatment recommendation	Results
OncotypeDX				
Oratz, et al., (2007) (15)	Retrospective study	N=74	Clinician treatment recommendation before and after GEP testing	RS led to change in clinicians' treatment recommendations in 21% of patients, and in actual administered treatment in 25% of patients.
Asad, et al. (2008) (16)	Retrospective chart review.	ER+, LN- Mean age: 54 years N=85	CHT for high risk based on international guidelines; and HT for low risk.	RS led to changes in the decision for chemotherapy in 37 (44%) of patients; 34% reduction in CHT recommendations.
Rayhanabad, et al. (2008) (17)	Retrospective chart review.	ER+, LN- Mean age: 54 years (range:26-78) N=58	CHT for high risk based on international guidelines; and HT for low risk.	RS led to change in management for 15 (26%) patients.
Geffen, et al. (2009) (18)	Prospective study.	LN- N=25	Not reported	RS led to a change in treatment recommendation for nine patients (36%). Six (24%) from chemotherapy to no chemotherapy.
Henry, et al. (2009) (19)	Retrospective study.	ER+, LN- N=29	Medical oncologist opinion; clinical data, AoL risk estimates followed by RS.	RS led to a change in CHT decisions in 9/29 (31%) patients, seven (24%) from CHT to no CHT and two (7%) from no CHT to CHT with low RS.
Klang, et al. (2010) (20)	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.
Lo, et al. (2010)	Prospective multicen	ER+, LN- Mean age: 55 years	Clinician treatment recommendation before and after	RS led to changes in clinician treatment recommendations for 28 patients (31.5%). 20 (22%) of these were from CHT to HT. Twenty-four patients (27%) changed their own

Author (year)	Design	Population	Prior treatment recommendation	Results
(21)	tre study	(range 35-77) N=89	GEP testing.	treatment decision, 9 from CHT to HT, 7 from HT to CHT, 2 from undecided to HT, and 2 from undecided to CHT.
Ademuyiwa, et al. (2011) (22)	Retrospective, consecutive series.	ER+,LN-, HER2- Mean age: 54.8 years (range:29-82) N=276	CHT recommendations based on clinicopathological characteristics	RS led to change in treatment for 38% of patients with 37 (13%) fewer patients receiving CHT.
Geffen, et al. (2011) (23)	Retrospective study	ER+ (N=134), LN-/LN1 Median age, 58 range (33–75) N=135	Treatment recommendations based on AoL.	RS led to a change in treatment recommendation for 34 (25.2%) patients. This change was from CHT to no CHT for 24 (17.8%) patients and from no CHT to CHT for 10 (7.4%).
Joh, et al. (2011) (24)	Retrospective study	ER+ N=154	Clinician panel	RS led to a 25% change in treatment recommendations.
Partin & Mamonunas (2011) (25)	Retrospective study	ER+, LN- N=169	Treatment recommendations based on AoL and St.Gallen	RS led to change in treatment recommendation in 27-74% of patients depending on comparator guideline.
Albaneil, et al. (2012) (26)	Prospective study	ER+, LN-, HER2- Mean age: NR, <50 yrs (n=40), ≥50 yrs (n=67) N=107	Treatment recommendation based on traditional clinicopathological factors.	RS led to changes in treatment recommendations in 32% of 107 patients enrolled: in 21% from CHT to HT and in 11% from HT to CHT.
Bargallo, et al. (2012) (27)	Prospective study	ER+, HER2-, LN-/LN1-3 Mean age: NR (range 32-89 years) N=96	Treatment recommendation based on conventional clinical–pathological factors and patient input.	RS led to changes in treatment decisions for 31/96 (32%) patients, including 17/62 (27%) LN- patients and 14/34 (41%) LN+ patients. The proportion of patients with a CHT recommendation decreased from 48% pre- to 34% post-assay.
Biroscak, et al. (2013) (28)	Retrospective study	ER+, LN- Mean age: 60.2 (range 39 to 78) N=50	Treatment was recommended based on histologic assessment.	RS led to changes in treatment decisions in 36 and 18% of cases by breast surgeons and medical oncologists, respectively. Breast surgeons increased recommendations for CHT in 15 (30%) of cases and decreased to no CHT in 3 (6%) of cases; and the medical oncologist increased to CHT in 4 (8%) of cases and decreased to no CHT in 5 (10%) of cases.
Davidson, et	Prospective	ER+, HER2-, LN- Mean age: 53	Treatment recommendation based on clinician	RS led to changes in CHT recommendations in 45/150 cases (30%) either to add (10%) or omit (20%) CHT.

Author (year)	Design	Population	Prior treatment recommendation	Results
al. (2013) (29)	study	(range 23-78 years) N=150	and patient pre-assay questionnaires.	
De Boer, et al. (2013) (30)	Prospective study	ER+, HER2-, LN-/LN1-3 Mean age: 56.2 N=151	Treatment recommendation base on routine pathology.	RS led to treatment recommendation changes for 24/101 patients with LN- tumours (24%) and for 13/50 patients with LN+ tumours (26%). For patients with LN- tumours there was a change from CHT to HT for 23%, and for 25% of patients with LN+ tumours.
Eiermann, et al. (2013) (31)	Prospective study.	ER+, HER2-, LN-/LN1 Mean age: 56 N=366	Treatment was recommended based on available clinical and histopathological data.	Treatment recommendations changed in 33% of patients. In 25% of patients CHT was changed to no CHT.
Holt, et al. (2013) (32)	Prospective cohort	ER+, LN-/LN1 N=106	Treatment recommendation base on NPI.	35 patients (33%) had their initial recommendation changed as a result of RS whilst for 71 patients (67%) there was no change. 25 (23.5%) changed from CHT to no CHT.
Cheung, et al. (2014) (33)	Retrospective study	ER+, HER2-, LN-/LN1 Mean age: 48 (range 24-67 years) N=154	Treatment recommendation based on clinical factors and AoL.	RS led to a change in treatment recommendations for 20 (31%) patients. 16 (10%) of these were changes to lower-intensity regimens (either equipose or HT).
Fried, et al. (2014) (34)	Retrospective study.	ER+, LN-/± (All intermediate RS) N=111	Clinician treatment recommendation before and after GEP testing.	RS led to a change in treatment recommendations for 24 patients (21.6%). Of 78 patients recommended HT alone, 11 changed to CHT (14.1%); of 33 recommended CHT, 13 received HT alone (39.4%).
Jaafar, et al. (2014) (35)	Retrospective study	ER+, LN- N=47	Treatment recommendation based on clinical factors.	RS led to a treatment change for 13 patients (27.7%), and CHT use decreased overall, from 48.9 to 25.5%.
Yamauchi, et al. (2014) (36)	Prospective study.	ER+, LN-/LN1-3, HER2- N=124	Treatment recommendations based on local and international guidelines.	RS led to a change in treatment recommendations in 33% of node-negative (N0) and 65% of node-positive (ND) patients. In 27 of 48 (56%) of N0 and 13 of 15 (87%) of N+ patients an initial recommendation for CHT was revised to HT after RS, and in 7 of 56 (13%) of N0 and 0 of 5 N+ patients from HT to CHT.
Gligorov, et al. (2015) (37)	Prospective study	ER+, LN-/LN1, HER2- N=95	Treatment recommendation based on clinical factors.	RS led to changes in treatment recommendations in 37% of patients, predominantly from CHT to HT alone. Patients recommended CHT decreased from 52% to 25% post RS.
Lee, et al.	Retrospective chart	ER+, LN-/LN1	Treatment was recommended using institutions'	RS led to a change in treatment decisions in 115 of 212 patients (54.2%), in 109 (51.4%) from CHT to HT, and in 6

Author (year)	Design	Population	Prior treatment recommendation	Results
(2015) (38)	review.	N=212	guidelines, based on clinicopathologic characteristics.	(2.8%) from HT to CHT.
Zhang, et al. (2015) (39)	Prospective study.	ER+, LN- N=134	Treatment recommendations based standard clinicopathologic criteria according to on St. Gallen and AoL.	RS led to a change in treatment decisions for 29% of patients, with 6% (8/134) changing to receive CHT and 23% (31/134) changing to reject CHT.
Kuchel, et al. (2016) (40)	Prospective study	ER+, HER2-, LN-/LN1-3 N=135	Clinician and patient recommendation before and after GEP testing.	RS led to changes in clinician treatment recommendations in 40.7% of patients. Of 69 patients with a pre-testing CHT recommendation, 43 (62.3%) had a recommendation change to HT only. Of the 66 patients with a pre-testing HT recommendation, 12 (18.2%) had a recommendation change to CHT. These changes led to a net reduction in the oncologists' CHT recommendation rate from 50.4 to 27.7%.  RS also led to 41 patients (31.3%) changing their treatment choice. Of the 52 patients with an initial CHT choice, 28 patients (53.8%) changed their choice to HT only. Of the 79 patients with an initial HT choice, 13 (16.5%) changed their choice to CHT. These changes led to a net reduction in CHT use from 39.7 to 28.2%.
Levine, et al. (2016) (41)	Prospective study.	ER+, LN-, HER2- N=979	Treatment recommendations based on AoL.	RS led to a change from unsure or CHT to no CHT in 365 (38%), and changed from unsure or no CHT to CHT in 143 (15%). CHT was recommended for 236 patients, 81% of whom received CHT.
Ozmen, et al. (2016) (42)	Prospective study.	ER+, LN-/LN1, HER2- N=165	Treatment was recommended based on histologic assessment.	RS led to a change in treatment decision for 33% of patients. Pre RS CHT was recommended to 92 (56%) of all patients, which decreased to 61 (37%) patients post-RS.
<b>MammaPrint</b>				
Bueno-de-Mesquita, et al. (2007) (43); Drukker, et al. (2013) (44)	Prospective multicentre study.	ER+/-, LN-/+ Mean age: 48 N=427	Treatment recommendations based on the Dutch Institute for Healthcare Improvement (CBO) guidelines.	Guidelines in addition to the prognosis signature and patient preferences led to an actual change in treatment for 19% of patients. 2% more CHT, 5% more HT, and 6% more CHT+HT.  At follow-up 124 patients were categorized as "low-risk" by the 70-gene signature, but high-risk by other measures, such as age, tumor size, nodal status, and other clinicopathological factors. Of these, 76% did not receive chemotherapy, and 98% survived 5 years with no recurrence of disease.
Gevensleben	Consecutive	ER+/-, LN-/+ N=136	Not reported.	GEP testing showed 40% of patients with either over (45%) - or undertreated (32%).



Author (year)	Design	Population	Prior treatment recommendation	Results
, et al. (2010) (45)	cohort			
Hartmann, et al. (2012) (46)	Prospective study.	ER+, LN-/LN1, HER2- Mean age: 70.3. N=60	Treatment according to national guidelines.	The prognosis signature used in combination with the clinico-pathological factors, would have led to changes in 18% of patients. Recommendations for CHT for 6 additional patients (10%) and withheld in 5 patients (8%).
Cusumano, et al. (2014) (47)	Prospective study.	ER+/-, LN-/LN1-3, HER2+/- N=194	Treatment recommendation based on clinicopathological factors.	MammaPrint led to changes in treatment advice for 37% of patients by the Dutch (14% decrease in CHT), 24% by the Belgian (0% decrease in CHT), 28% by the Italian (13% increase in CHT) and 35% by the Spanish teams (2% decrease in CHT).  MammaPrint increased the inter-institutional agreement in treatment advice (CHT or no CHT) from 51% to 75%.
Drukker, et al. (2014) (48)	Retrospective case review	N=37 (other factors not reported)	Treatment recommendation based on clinicopathological factors.	MammaPrint led to changes in treatment advice in 24% of cases. Pre MammaPrint recommended treatments were CHT in 48%, and HT in 46% of the cases. After adding MammaPrint recommended treatments were CHT 37%, and HT in 57% of cases.  Adding MammaPrint resulted in 14.3% of the cases in a change from CHT to HT or no treatment. In 2.1% of the cases the advice of no treatment or HT was changed to CHT. This resulted in a reduction in CHT use of 12.2%.
Exner, et al. (2014) (49)	Prospective study.	ER+, LN-/LN1-3, HER- N=75	Treatment according to clinicopathological factors and St Gallen guidelines.	MammaPrint led to changes in treatment advice in 18.6% of cases. In 10 patients (13.3%), there was a decision change towards HT and in 4 patients (5.33%) towards CHT.
<b>IHC4</b>				
Yeo, et al. (2015) (50)	Prospective study	ER+, LN-/LN1, HER2- Median age: 59 N=124	Treatment recommendation based on clinicopathological factors.	IHC4 led to changes in treatment advice for 42 patients (34%). Prior to the IHC4 score becoming available clinicians recommended CHT (or at least its discussion) to 74 (59%) patients, this fell to 32 (34%) after the IHC4 score.

RS = recurrence score, AoL = Adjuvant! Online, NPI = Nottingham Prognostic Index, ER+ = oestrogen receptor positive, ER- = oestrogen receptor negative, LN- = lymph node negative, LN 1-3 = 1-3 lymph nodes involved, HT = hormone therapy, CHT = chemotherapy.

**Figure 1: PRISMA Diagram**

