Evidence overview

Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in breast cancer management: MammaPrint, Oncotype DX, IHC4 and Mammostrat

This overview summarises the key issues for the Diagnostics Advisory Committee’s consideration. It includes a description of the topic, a description of the analytical structure and model, a discussion of the analytical difficulties, and a brief summary of the main results. It is not a complete summary of the diagnostics assessment report, and it is assumed that the reader is familiar with that document. This overview contains sections from the original scope and the diagnostics assessment report, as well as referring to specific sections of these documents.

1 Background

1.1 Introduction

The Randox Breast Cancer Assay (Randox BCA) was selected by the Medical Technologies Advisory Committee for recommendations on its use for guiding adjuvant chemotherapy decisions in breast cancer management. A wide variety of technologies may be used to guide decisions about adjuvant chemotherapy in breast cancer and were potentially available for inclusion in this evaluation. Following discussions with specialist Committee members and the External Assessment Group (EAG), eight further technologies were identified during the scoping stage and included in the assessment as alternative tests. During the assessment conducted by the EAG it became apparent that 5 out of the 9 technologies did not meet the criteria for inclusion in the EAG’s economic analysis. These five technologies (Randox BCA,
BluePrint, Breast Cancer Index, PAM50 and NPI+) have been omitted from the evaluation as information for evaluating the cost-effectiveness of these technologies in the NHS in England is not available. Therefore, four technologies are considered in this evaluation. Two are based on gene expression profiling: MammaPrint (Agendia) and Oncotype DX (Genomic Health). Two are based on immunohistochemistry (also referred to as protein expression profiling in the diagnostics assessment report): IHC4 (academic sponsor) and Mammostrat (Clariant). The four technologies considered in this evaluation are eligible for DAP guidance.

The purpose of the meeting of the Diagnostics Advisory Committee on January 4th 2012 is to formulate provisional recommendations on the clinical and cost-effectiveness of MammaPrint, Oncotype DX, IHC4 and Mammostrat in guiding the use of adjuvant chemotherapy in women with early breast cancer in England.

1.2 The condition(s)

Epidemiology and incidence
Breast cancer is the most commonly diagnosed cancer in women in England and Wales. In 2009 there were 42,305 new cases diagnosed. Breast cancer is the second largest cause of cancer death in women after lung cancer, with an age-standardised mortality rate of 26 per 100,000 women. In 2008 this constituted 10,716 deaths for women in England and Wales.

Incidence varies most with gender. Women are far more likely to get breast cancer than men. For both genders, incidence varies with age (see table 3 of the diagnostics assessment report). Just over 80% of cases occur in women aged 50 years and over. In England and Wales, 2006 data demonstrate highest rates for women in the 60- to 70-year age range.

Incidence also varies with ethnicity. Asian, Chinese and black ethnic groups and those with mixed heritage have a lower incidence than the white ethnic group in England. Incidences are 0.65, 0.75, 0.49 and 0.58 that of the white ethnic group respectively.
Prognosis
Overall, 5-year, age-standardised, survival rates for breast cancer are around 80%. Survival varies with age, stage of disease, ethnicity, socioeconomic status and tumour characteristics.

Clinicians estimate prognosis using tools such as the Nottingham Prognostic Index, which takes into account grade as well as size and spread of tumour, or Adjuvant! Online, which uses age of the patient, tumour size, nodal involvement, hormonal receptor status and histological grade to predict disease course and treatment options. Good prognosis is associated with small tumour size, node-negative (LN–) status, younger age, oestrogen receptor positive (ER+) and progesterone receptor positive (PgR+) status. HER2 over-expression (also known as HER2 positive (HER2+)) is associated with poor prognosis.

The economic analysis performed by the EAG modelled patients with ER+, LN- and HER2- early breast cancer. This was considered to be the population in which the tests were most likely to be used in the first instance in England.

1.3 Diagnostic and care pathways
Patients diagnosed with early breast cancer currently follow the diagnosis/treatment pathway described in Figure 1.
Chemotherapy is defined as the use of cytotoxic drugs with the intention of preventing cancer recurrence. Chemotherapy regimens containing anthracycline are used routinely after cancer surgery (in the adjuvant setting). It should be noted that, for the
purposes of this assessment, chemotherapy does not include other forms of systemic therapy such as endocrine treatments or targeted biological therapy (trastuzumab – used to treat certain HER2+ tumours).

Meta-analyses of randomised clinical trials by the Early Breast Cancer Trialists’ Collaborative Group (EBCTCG) have indicated that the use of adjuvant chemotherapy (chemotherapy following surgery) is associated with a reduced risk of relapse and death in women with early stage breast cancer. However chemotherapy is associated with considerable adverse effects. Short-term and long-term adverse effects will affect a proportion of patients receiving chemotherapy, imposing costs and reducing the quality and possibly length of life. Although chemotherapy may prevent relapse in some, not all women with early stage breast cancer will benefit and many women remain recurrence free at 10 years without chemotherapy.

Further information on diagnosing and staging of breast cancer can be found in section 3.1.4 of the diagnostics assessment report.

**Current guidelines**

NICE cancer service guidance ‘Improving outcomes in breast cancer’ (2002), recommends that women at intermediate or high risk of recurrence who have not had neoadjuvant chemotherapy should normally be offered multi-agent chemotherapy which includes anythracyclines.

Some patients with a ‘good’ prognosis may still have recurrence after curative surgery and adjuvant therapy. This presents a challenge to clinicians to estimate prognosis and make the most appropriate decisions about the use adjuvant chemotherapy in women with early stage breast cancer.

Historically factors such as patient age, tumour size, nodal involvement, histological grade, oestrogen receptor (ER) expression, human epidermal growth factor receptor type 2 (HER2) overexpression and comorbidities have been assessed and considered alongside patient preference when assessing risk and recommending therapies. In the UK, local guidance based on the Nottingham Prognostic Index (NPI) and Adjuvant! Online has been developed to help clinicians decide the benefits of
adjuvant chemotherapy for a particular patient. The guidance provides information about prognosis which is largely based on pathological parameters (for example, tumour size, grade and lymph node status) for NPI with the addition of ER receptor status, age and comorbidity for Adjuvant! Online. However these are imperfect because different local guidance can give different results and it has been suggested that a proportion of women with early stage breast cancer are over- or under-treated. This may result in unnecessary use of expensive chemotherapy with its associated adverse effects for women who derive little or no benefit. In addition, there may be avoidable deaths in women who had chemotherapy withheld.

‘Early and locally advanced breast cancer: diagnosis and treatment’ (NICE clinical guideline 80 [2009]) recommends that adjuvant therapy should be considered for all patients with early invasive breast cancer after surgery, based on assessment of the prognostic and predictive factors, the potential benefits and side effects of the treatment. These guidelines do not refer to the use of gene expression profiling or expanded immunohistochemistry tests to aid decision-making. NICE clinical guideline 80 recommends that decisions should be made following discussion of these factors with the patient and that Adjuvant! Online should be considered to support estimations of individual prognosis and the absolute benefit of adjuvant treatment. The Nottingham Prognostic Index (NPI) is also commonly used locally to aid decisions about chemotherapy for patients with early stage breast cancer.

Role of new tests
Gene expression profiling and expanded immunohistochemistry (or protein expression) aim to improve the targeting of chemotherapy in breast cancer by improving the stratification and identification of patients who will gain most benefit. The rationale is based on the knowledge that certain biological features of cancers may indicate an increased likelihood of rapid growth and metastatic potential. The aim of this assessment is to evaluate whether, by guiding the selection of patients to receive adjuvant chemotherapy, gene expression profiling and expanded immunohistochemistry tests improve health outcomes and quality of life in patients with early stage breast cancer compared with currently used decision-making protocols.
The technologies

2.1 Tests under evaluation

Gene expression profiling and expanded immunohistochemistry tests typically report two types of information – breast cancer sub-types and/or risk of recurrence. Tests developed to provide information on sub-types can be used either before surgery to inform decisions on neoadjuvant therapy or after primary surgery to inform decisions on adjuvant chemotherapy. Tests predicting the risk of recurrence in a specific population are likely to be used after surgery, in conjunction with other information such as tumour size and grade to guide the use of adjuvant chemotherapy. These tests are typically indicated for women with ER+ and LN– (and sometime LN+ if the number of nodes is small).

Test results are likely to be used, in conjunction with other information available about tumour size, grade etc, to guide the decision on which patients should be offered adjuvant chemotherapy. Some tests involving sending samples away for central review, following surgery, which may introduce a delay (of 2 or 3 weeks) before the decision of whether or not to offer chemotherapy.

Four technologies are considered in this evaluation: two are based on gene expression profiling and two on immunohistochemistry.

Gene-expression profiling

Gene expression profiling identifies and quantifies mRNA transcripts in a specific tissue sample. Because only a fraction of the genes encoded in the genome of a cell are transcribed into mRNA, gene expression profiling provides information about the activity of genes that give rise to these mRNA transcripts.

Various profiling tests are used in breast cancer to investigate the expression of specific panels of genes (also known as a gene profile or gene signature). They work by using different techniques to measure mRNA levels in breast cancer specimens, including real-time reverse transcription polymerase chain reaction (RT-PCR) and DNA microarrays. Many of these tests have been designed to measure the risk of cancer recurrence. Other uses include breast cancer sub-typing (using molecular
classification systems), predicting the likely benefit from certain types of therapy (for example, chemotherapy) and diagnosis.

Different tests use different methods for preparing the RNA and different protocols for preparing the specimens (for example, formalin fixation, paraffin embedding, snap freezing and fresh samples). Furthermore, there are different algorithms for combining the raw data to obtain a summary measure. All of these factors can affect the reproducibility and reliability of gene expression profiling tests.

The two gene expression profiling tests considered in this evaluation are:

- **MammaPrint (Agendia)** is based on microarray technology which uses a 70-gene expression profile. MammaPrint is intended as a prognostic test for women of all ages, with LN- and LN+ (up to 3 nodes positive) breast cancer with a tumour size of 5.0 cm or less. MammaPrint is used to estimate the risk of distant recurrence of early breast cancer. Patients are stratified into two distinct groups — low risk (good prognosis) or high risk (poor prognosis) of distant recurrence. MammaPrint has been cleared by the FDA as an In Vitro Diagnostic Multivariate Index Assay. The test uses fresh samples.

- **Oncotype DX** (Genomic Health) quantifies gene expression for 21 genes in breast cancer tissue by RT-PCR. It predicts the likelihood of recurrence in women of all ages with newly diagnosed stage I or II, ER+ LN− or LN+ (up to 3 nodes) breast cancer treated with tamoxifen. The test assigns the breast cancer a recurrence score (RS) and a risk category – low (RS < 18), intermediate (19 < RS < 31) or high (RS > 31). The test also reports ER progesterone receptor (PgR) and HER2 status. The test uses formalin-fixed paraffin-embedded samples.

Key details of individual gene expression profiling tests are included in table 6 of the diagnostics assessment report.

**Expanded immunohistochemistry (protein expression profiling) tests**

Immunohistochemistry markers are being developed to provide similar information to that given by the gene expression profiling tests. Instead of measuring RNA levels,
protein expression profiling measures protein levels in the tumour sample. Some of these tests offer the advantage of using existing immunohistochemical markers (such as ER, and HER2) which are routinely tested in all UK pathology departments.

The two expanded immunohistochemistry tests included in this evaluation are:

- The IHC4 test assesses levels of four key proteins in a breast cancer sample, ER, PgR, HER2 and Ki-67. The IHC4 score is calculated from the percentage of cells positive for Ki67 and PgR (0–100%), the Histoscore (a measure of the percentage of cells positive multiplied by intensity, range 0–300) for ER status, and the tumour HER2 status (expressed as positive or negative). The final algorithm for IHC4 calculates a risk score for distant recurrence based on ER, PgR, HER2 and Ki67 in addition to classical clinical and pathological variables (composite risk score IHC4 plus clinical is referred to as IHC4). An online calculator is expected to be available at the beginning of 2012 (personal communication). The test uses formalin-fixed paraffin-embedded samples.

- The Mammostrat test uses five immunohistochemical markers (SLC7A5, HTF9C, P53, NDRG1, and CEACAM5) to stratify patients into risk groups. These markers are independent of one another and do not directly reflect either proliferation or hormone receptor status. The current version of the test provides categorical classification of breast cancer sub-type, and quantitative values for ESR1/ER, PR/PgR, ERBB2/HER2, Proliferation, and Luminal score (ER-pathway). The test uses formalin-fixed paraffin-embedded samples.

Key details of individual expanded immunohistochemistry tests are included in table 7 of the diagnostics assessment report.

2.2 Comparator

The comparator is standard UK practice. This varies between trusts and encompasses the use of Adjuvant! Online (recommended in NICE clinical guideline 80) and/or guidance based on NPI (discussed in NICE clinical guideline 80) to guide decisions on which patients with early breast cancer should be offered adjuvant chemotherapy.
Adjuvant! Online
The Adjuvant! Online computer programme is designed to provide estimates of the benefits of adjuvant endocrine therapy and chemotherapy. The current version of Adjuvant! Online does not include HER2 status and the potential benefit of trastuzumab. Patient and tumour characteristics are entered into the programme and provide an estimate of the baseline risk of mortality or relapse for patients without adjuvant therapy. Information about the efficacy of different therapy options are derived from EBCTCG meta-analyses and provide estimates of reduction in risk at 10 years of breast cancer-related death or relapse for selected treatments. These estimates are then provided on printed sheets in simple graphical and text formats to be used in consultations.

Nottingham Prognostic Index (NPI)
The NPI is a composite prognostic parameter involving both time-dependent factors and aspects of tumour aggressiveness. The NPI score is based on a mix of grade, lymph node involvement and tumour size. The score is calculated by adding numerical grade (1, 2, or 3), lymph node score (negative = 1, 1 to 3 nodes = 2, > 3 nodes = 3) and 0.2 times tumour size in centimetres. Patients can be divided into three prognostic groups on the basis of the NPI: a good prognostic group (NPI < 3.4), a moderate prognostic group (3.4 < NPI < 5.4), and a poor prognostic group (NPI > 5.4).

3 The evidence

3.1 Systematic review of clinical effectiveness
A systematic review of the evidence on the clinical effectiveness of four gene expression profiling and expanded immunohistochemistry tests was undertaken by the EAG. In addition, supplementary evidence provided by the manufacturers of the tests is presented within the diagnostics assessment report.

For two of the four tests (Oncotype DX and MammaPrint), the current review updates an existing systematic review of gene expression profiling tests for breast cancer. Two previous systematic reviews (one an update of the other) reviewed the literature relating to both Oncotype DX and MammaPrint. Marchionni et al.2008 conducted an
exhaustive literature review of various electronic databases (covering biomedical literature) between 1990 and 2006. In 2010, Smartt updated this systematic review and included all relevant evidence between 2007 and 2009.

In the present review, new search strategies were developed for all of the four tests based on scoping searches (and strategies reported in the two existing systematic reviews for the Oncotype DX and MammaPrint).

Full details of the systematic review can be found in chapter 4 of the diagnostics assessment report. The outcome measures included in the review of clinical effectiveness were as follows:

- **Analytic validity**— the ability of the test to accurately and reliably measure the expression of mRNA or proteins by breast cancer tumour cells (that is, repeatability and reproducibility).
- **Clinical validity** – the degree to which the test can accurately predict the risk of an outcome (typically distant recurrence) and discriminate patients with different outcomes.
- **Clinical utility** – the tests’ ability to discriminate between those who will have more or less benefit from a therapeutic intervention.

Clinical utility relates to improvements in clinical outcomes such as overall survival, disease-free survival, adverse effects of chemotherapy, or quality of life. Based on the conclusion of previous and current systematic reviews, prospective studies reporting long-term outcomes such as overall survival were not available. In the absence of such studies the following outcomes were included:

- Reclassification of risk compared with existing prognostic variables (correlations between test score and score on existing measures such as NPI, Adjuvant! OnLine). That is, how does the test change the classification of risk for patients?
- Impact of the test results on clinical decision-making – how do the tests results translate into changes in decision-making (for example, changes in the proportion of patients receiving adjuvant chemotherapy)?
3.2 Results

In summary, the vast majority of the clinical evidence was related to the Oncotype DX and MammaPrint tests, indicating that these tests were significantly further along the validation pathway than the other tests examined in the review. The highest quality evidence was reported for Oncotype DX, although limitations or gaps in the clinical data were identified for all tests. Most studies, for all tests, were retrospective in design, analysing archived tumour samples from a cohort of patients with documented information regarding patient characteristics and outcomes. This type of design is associated with increased biases in comparison with prospective randomised controlled trials (RCTs).

The current review identified evidence on analytical validity, clinical validity and clinical utility. The study populations were generally heterogeneous although most of evidence on Oncotype DX came from ER+, LN− populations. Most studies included a small number of participants, but seven studies included more than 1000 samples (three for Oncotype DX, three for Mammostrat and one for IHC4). Follow-up was short or not reported for a large number of studies. Five studies were specific to a UK population, including three for Oncotype DX, one for IHC4, and one for Mammostrat.

A summary of the clinical effectiveness of the four tests can be found in chapter 2 (executive summary) and detailed results in section 4.2 of the diagnostics assessment report. A summary is provided below (section 3.2.1) of the clinical effectiveness results of the four tests included in this evaluation. The results of the economic analyses are highly sensitive to the assumptions regarding the benefit of chemotherapy, therefore, further details are provided below in section 3.2.2.
3.2.1 Clinical effectiveness results summary

Oncotype DX

Oncotype DX was reported to be furthest along the validation pathway by the previous systematic reviews. In terms of clinical validity these reviews reported evidence that the Oncotype DX recurrence score was significantly correlated with disease-free survival and overall survival. Furthermore, only recurrence score was shown to be a better predictor of distant recurrence at 10 years than traditional clinicopathological predictors. The evidence on clinical utility was limited. One study demonstrated a significant benefit from the use of chemotherapy in the Oncotype DX high-risk group, although the review highlighted that the study may have potential flaws. Key gaps were identified in the evidence base related to the extent to which the test added to the management of breast cancer and the proportion of patients who would benefit from the test along with the stability of risk categories in populations other than patients with ER+, LN– early breast cancer. The previous systematic reviews indicated that prospective confirmation of the clinical utility of Oncotype DX was needed.

The EAG’s review identified 12 additional studies of Oncotype DX. Further larger studies now support the prognostic capability of Oncotype DX. One large scale UK study in postmenopausal women with ER+, LN– early breast cancer found that an increase in risk score was significantly associated with an increased risk of distant recurrence. Furthermore, the evidence base has been extended to include the LN+ population. The EAG’s did not identify any prospective studies of the impact of Oncotype DX on long-term outcomes such overall survival. Four studies were identified which presented further evidence on the impact of Oncotype DX on clinical decision-making. These indicated that the use of Oncotype DX leads to changes in treatment decisions for between 31.5% and 38% of patients. However, only one of these was in the UK and limitations in relation to study design were identified. Specifically, these data were based on a small sample (n = 106) and lack of a description of how treatment decisions were made limit the generalisability of the study. Three studies provided evidence supporting the case that Oncotype DX predicts benefit from chemotherapy. The first evidence relating to improvements in quality of life and reductions in patient anxiety as a result of using Oncotype DX have
been reported, but generalisations should be made with caution because of small sample sizes.

**MammaPrint**

There are a range of studies that provided evidence on the prognostic ability of MammaPrint in heterogeneous populations. However, the previous systematic reviews indicated that evidence relating to the clinical validity of MammaPrint was not always conclusive nor supported the prognostic value of the test. Four studies suggested that the test could predict prognosis, one study failed to verify the prognostic utility of the test and in another the methods and results were at variance with other studies. In terms of clinical utility, the previous reviews identified one study demonstrating that MammaPrint had an impact on clinical decision-making. The addition of MammaPrint to the standard Dutch clinical assessment of risk (modified by patient preference) in a cohort of 427 increased the number of patients receiving adjuvant systemic therapy by 20 (5%). However, follow-up was not long enough to provide evidence of its effect on clinical endpoints such as distant metastasis-free survival or its utility in predicting treatment benefit. It was recommended that further evidence from randomised controlled trials was needed in addition to robust evidence on the prediction of chemotherapy benefit.

The EAG review identified seven (non-UK-based) additional studies on the MammaPrint test. Four additional studies on the clinical validity of MammaPrint demonstrated that the MammaPrint score is a strong independent prognostic factor, and may provide additional value to standard clinicopathological measures. However the test may only be reliable at predicting outcome at 5 years, rather than 10 years, although the population in all these studies was relatively small. One of the studies was of a Japanese population, and follow-up was limited to only 5 years in two of the studies. Six studies with data on the clinical utility of MammaPrint were identified by the current review. Five studies reported use of MammaPrint to reclassify patients into high- and low-risk groups and compared this with the risk assigned according to current guidance. They reported a high level of discordance between MammaPrint and current classification, although these studies did not demonstrate how this would impact on treatment decisions. One study reported that the use of MammaPrint 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, rather than by providing evidence of actual changes in practice.

A study of the benefit of chemotherapy according to MammaPrint risk group was identified but was omitted from the systematic review because it was based on a pooled analysis of six primary studies rather than a meta-analysis of the separate studies. The pooled analysis approach is not considered methodologically appropriate.

Robust evidence of clinical utility is not yet available. It is not yet clear whether use of the MammaPrint will change the management of breast cancer. In summary, most studies of MammaPrint were retrospective in design, used small sample sizes and were heterogeneous with respect to patient populations. Moreover, no studies were conducted in the UK setting.

Mammostrat
The EAG identified three studies that provided data to support the use of Mammostrat as an independent prognostic tool for women with ER+, tamoxifen-treated breast cancer. These studies included a large sample size, appeared to be of reasonable quality, and provided data from a UK setting (one study). In addition, clinical utility data of Mammostrat (from one study) suggest that low- and high-risk groups benefit from chemotherapy, with high-risk groups benefiting more than low-risk patients. However, no benefit was observed in moderate-risk groups. This anomaly makes assessing clinical utility questionable. There was no published evidence on reclassification of risk groups compared with current means of classification, and no evidence of the impact of the test on decision-making.

IHC4
No studies on analytical validity of IHC4 were identified. However three of the four individual tests which make up IHC4 (ER, PgR and HER2) are currently used in the NHS. There are however outstanding issues around the reproducibility of Ki-67 detection. The current review identified one study on the clinical validity of IHC4,
which claims that the IHC4 score is a highly significant predictor of distant recurrence. This initial study included a large sample size and detailed the development of the test in one cohort and the external validation of the test in an independent cohort. The study was rated as high quality. The review did not identify any published evidence on the clinical utility of IHC4 in terms of its ability to change treatment decisions or its ability to predict chemotherapy benefit. In summary, the evidence base for IHC4 is currently limited to clinical validity (prognostic ability), although this evidence is considered to be relatively robust. Further evidence is required on its analytic validity and clinical utility.

3.2.2 Chemotherapy benefit

NICE clinical guideline 80 states ‘Meta-analyses of randomised clinical trials by the EBCTCG have indicated that the use of adjuvant chemotherapy is associated with a reduction in the risk of relapse and death in women with early stage breast cancer (EBCTCG, 2005). The reduction in risk of relapse and death attributable to adjuvant chemotherapy is dependent on age at diagnosis but is independent of prognosis. The absolute benefit of chemotherapy therefore varies according to both patient age and underlying prognosis. Estimates of the benefits of adjuvant chemotherapy are therefore made on the basis of patient age and prognosis derived from pathological features.’

Results of the EBCTCG overview (published in 2005) of the effects of chemotherapy for early breast cancer show that the proportional reductions in recurrence attributable to chemotherapy were dependent on age but independent of prognosis (as stated in NICE clinical guideline 80). The EBCTCG recently published, in December 2011, the updated results of the overview comparing different polychemotherapy regimens for early breast cancer. The results showed that in all meta-analyses involving taxane-based regimens or anthracycline-based regimens, the proportional reductions in recurrence (and breast cancer mortality) appeared largely independent of age, nodal status, tumour diameter, tumour differentiation, ER status or tamoxifen use.

GEP and expanded IHC tests have not been evaluated in previous NICE guidance and, therefore, conclusions regarding the benefit of chemotherapy in individual risk
groups assigned by GEP or immunohistochemistry-based tests are not available. Data on the ability of each of the four tests to predict the benefit of chemotherapy based on the outcome of the test are summarised below.

**Oncotype DX:** In total, 3 studies were identified that reported that the chemotherapy benefit by Oncotype DX risk groups. The previous systematic reviews identified Paik et al. (2006) as part of their review of clinical utility. Paik et al., (2006) used assessable specimens and data from an existing trial (NSABP B-20) to address the potential value of the RS in predicting chemotherapy benefit in a population of ER+, LN- patients (n = 651). This study compared a group treated with tamoxifen and chemotherapy with a group of patients who were randomized to tamoxifen only. The RS was found to be correlated with chemotherapy benefit, defined in terms of 10-year distant recurrence free survival, with a significant benefit from the use of chemotherapy in the high RS group (p = 0.001). However, in a multivariate analysis the benefit from chemotherapy was unclear due to large confidence intervals in the low and intermediate RS risk groups. The review highlighted that the study may have been subject to bias as some patients in the validation dataset (tamoxifen treated patients of the NSABP B-20 trial) were also in the training dataset. Furthermore, there were specific limitations in that the chemotherapy and tamoxifen arms were unbalanced (227 randomly assigned to tamoxifen and 424 randomly assigned to tamoxifen plus chemotherapy), the study compared the use of CMF/MF in addition or not to tamoxifen. Clinical experts indicated that more effective chemotherapy regimes are currently used in the UK. Paik et al. also explored the interaction of clinical variables with chemotherapy benefit. When assessing age, it was found that more than 44% of patients were aged below 50 years old and the benefit of chemotherapy (reduction in distant recurrence) was greater in this population compared to women aged over 50 years old. The hazard ratio for the benefit of chemotherapy (reduction in distant recurrence) in women aged above 50 years old compared to younger women was 2.02 (CI: 0.75 – 5.47, p = 0.162). The corresponding hazard ratio for the entire NSABP B-20 cohort (n = 2,299) was 1.78 (CI: 1.06 – 2.97, p = 0.029). Paik et al. (2006) was used to inform the estimates of chemotherapy benefit in the Oncotype DX and IHC4 base case analysis.
Two new studies (with 3 related citations, Tang et al., 2010 [abstract only]; and Tang et al., 2011 used the same trial data as Paik et al. 2006; and Albain et al. 2010) provided evidence supporting the case that OncotypeDX predicts chemotherapy benefit. The Albain et al. study employed a sample of 367 postmenopausal ER+ and LN+ US and Canadian patients from the SWOG-9914 trial. RS was found to be prognostic for tamoxifen-treated patients with positive nodes and predicts significant benefit of CAF in tumours with a high recurrence score; a low score identifies women who might not benefit from anthracycline-based chemotherapy, despite positive nodes. The Tang 2010 study employed a sample of 625, ER+, LN- US patients from the NSABP B-20 trial. They examined the value of the RSPC (integration of RS and clinco-pathological factors) in prediction of chemotherapy benefit in reducing risk of recurrence and concluded that RS used alone remains the best predictor of chemotherapy benefit in ER+, LN- breast cancer. Tang 2011 study employed a sample of 1319, ER+, N- patients from two large US trials (NSABP-B14 and B20). Showed that in the B20 cohort RS was significantly predictive of chemotherapy benefit (for distant recurrence free interval, for overall survival, and disease-free survival.

**MammaPrint**: A study on the benefit of chemotherapy by MammaPrint risk group (Knauer et al. 2010) was identified but omitted from the systematic review because it was based on a pooled analyses of six primary studies (which were included in the review in their own right). This study reports findings on chemotherapy benefit for high and low MammaPrint risk groups but the findings are not considered to be robust as the authors do not reanalyse the tumour samples and it is unclear how individual patient data were combined. As the Knauer et al. 2010 study was used to inform the estimates of chemotherapy benefit in the MammaPrint base case analysis further details of the study are provided below.

Knauer et al. 2010 created a pooled database of patients from six prior studies. They included 541 women with unilateral stage T1-3, LN-1, M0 invasive breast cancer diagnosed between 1984 and 2006. Each tumour had been classified as having a high or low risk signature using the MammaPrint test: 252 (47%) as low risk and 289 (53%) as high risk. Median follow-up was 7.1 years, but all analyses were censored at
five years. Women in both risk categories appeared to benefit from chemotherapy, although the estimates were not statistically significant in the low risk group. For breast cancer disease specific survival the unadjusted hazard ratio for chemotherapy was 0.58 (0.07-5.0) in the low risk group and 0.21 (0.07-0.59) in the high risk group. The p-value for interaction between use of chemotherapy and the risk signature was not statistically significant (p=0.45). For distant metastases free survival the unadjusted hazard ratio for chemotherapy was 0.26 (0.03-2.0) in the low risk group and 0.35 (0.17-0.71) in the high risk group. The p-value for the interaction was not reported, but in this case the trend was towards greater relative benefit from chemotherapy in the low risk group. This study, however, has some major statistical flaws. For instance, data were truncated arbitrarily at 5 years, despite that fact that the median follow-up was 7.1 years. Censoring the follow-up at five years biased the results in favour of the utility of prognostic signature because the association between the 70-gene signature and recurrent disease falls quickly after five years of follow-up. As the majority of distant recurrences and deaths from breast cancer occur more than five years after diagnosis, this is a significant limitation.

**Mammostrat:** One study was identified that reported the chemotherapy benefit by Mammostrat risk groups. Ross et al. 2008 presented evidence on the tests ability to identify patients who have greater absolute benefit from adjuvant chemotherapy compared with unstratified patient populations. These analyses were based on the tamoxifen and cytotoxic chemotherapy treated (n=269) and B20 tamoxifen only treated (n=161) patients from the trial data. In terms of recurrence-free interval patients in the low risk improved by 5% from 86% to 91%, HR 0.4 (95%CI: 0.2 – 0.8), and patients in the high risk group 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. The Ross et al. 2008 study was used to inform the estimates of chemotherapy benefit in the Mammostrat base case analysis.

**IHC4:** There is currently no published evidence on the clinical utility of IHC4 including its ability to predict chemotherapy benefit. The benefit of chemotherapy was applied from the RS risk group (using Oncotype DX – Paik et al. 2006 study) as no specific data for IHC4 exist.
4 Patient outcomes and cost effectiveness modelling

4.1 Systematic review of cost-effectiveness evidence

Four studies (corresponding to five references) were identified as meeting the inclusion criteria of the systematic review of economic evaluations. This included an economic analysis developed as part of an evaluation in the Ontario Health Technology Assessment Series.

Of the four identified economic studies, two compared MammaPrint against Adjuvant! Online and two compared OncotypeDX against Adjuvant! Online. None were conducted in a UK setting (see section 5.1 of the diagnostics assessment report).

A summary of the main results from all four studies can be found in table 1.

Table 1 Summary of the main results from the cost-effectiveness studies identified in the systematic review

<table>
<thead>
<tr>
<th>Reference</th>
<th>Test evaluated</th>
<th>Costs</th>
<th>QALYs</th>
<th>ICER (when compared to Adjuvant! Online)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retel et al. (2010)</td>
<td>MammaPrint</td>
<td>€28,045</td>
<td>12.44</td>
<td>€4,614</td>
</tr>
<tr>
<td></td>
<td>Adjuvant! Online (comparator)</td>
<td>€26,915</td>
<td>12.20</td>
<td></td>
</tr>
<tr>
<td>Chen et al. (2010)</td>
<td>MammaPrint</td>
<td>$163,580</td>
<td>21.218</td>
<td>$9,428</td>
</tr>
<tr>
<td></td>
<td>Adjuvant! Online (comparator)</td>
<td>$162,140</td>
<td>21.065</td>
<td></td>
</tr>
<tr>
<td>Tsoi et al. (2010)</td>
<td>Oncotype DX</td>
<td>19,747 ($CAD)</td>
<td>13.638</td>
<td>63,064 ($CAD) (≈ £39,917*)</td>
</tr>
<tr>
<td></td>
<td>Adjuvant! Online (comparator)</td>
<td>15,645 ($CAD)</td>
<td>13.573</td>
<td></td>
</tr>
</tbody>
</table>
4.2 Economic models submitted by manufacturers

Genomic Health (Oncotype DX) and Clarient (Mammostrat) submitted economic evaluations considering the cost effectiveness of their technologies in the UK. A summary of the main results from these analyses can be found in table 2.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Test evaluated</th>
<th>Costs (£)</th>
<th>QALYs</th>
<th>ICER (when compared to the comparator) (£)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genomic Health</td>
<td>Oncotype DX</td>
<td>12,735</td>
<td>11.54</td>
<td>6,232</td>
</tr>
<tr>
<td>'Usual care’</td>
<td></td>
<td>11,847</td>
<td>11.39</td>
<td></td>
</tr>
<tr>
<td>Clarient</td>
<td>Mammostrat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These evaluations were potentially more relevant for decision-making in England. However, the following issues in the evaluations need further consideration:
• assumption about the baseline level of chemotherapy in clinical practice in England
• assumption about the risk of distant recurrence in England
• assumption about the proportion that would be offered chemotherapy after reclassification with the new test in England
• assumption about who would be offered the test in England
• assumption about the cost of chemotherapy and therapy used to prevent or treat associated adverse event in England

Full details and critiques of these economic models can be found in sections 5.2 and 5.3 of the diagnostics assessment report.

4.3 De novo economic analysis performed by the EAG
The EAG constructed a de novo economic model to address the limitations described above and to estimate the cost effectiveness of a wider range of gene expression profiling and expanded immunohistochemistry tests. Notably, the EAG’s economic assessment uses UK data and addresses the following:

• proportion of patients who currently receive chemotherapy
• the risk of distant recurrence
• offering the test to the subgroup of patients most likely to benefit
• a more accurate estimation of the cost of chemotherapy

4.4 Population assessed
Women with ER+, LN–, HER2– early breast cancer
The economic assessment focused on women with ER+, LN–, HER2– early breast cancer. This subgroup was selected after review of the evidence (see section 4.2 of the diagnostics assessment report) and the indications of the tests (see tables 6 and 7 of the diagnostics assessment report), discussion with clinical experts and the perceived likelihood of the use of the test resulting in a change in current clinical practice. Patients with HER2+ or LN+ early breast cancer were not considered in this assessment because of time and resource constraints and lack of evidence, but should be the subject of future research.
4.5 Subgroups for whom the new tests are the most likely to be used

**Women with an NPI score above 3.4**

Previous economic evaluations have typically assumed that the new tests may be offered to all women with ER+, LN−, HER2− early breast cancer. Clinical experts suggested that the new tests may be targeted at a subgroup of this population – those at intermediate risk (and typically those aged less than 75 years) for whom the decision about whether or not to give chemotherapy is most uncertain.

This ‘intermediate group’ is not well defined, but clinical experts suggested that patients with a NPI of 3.4 or less are unlikely to receive chemotherapy (except for a few very young women with aggressive early breast cancer).

Consequently two analyses are included:

- the new test is given to all women with ER+, LN−, HER2− early breast cancer
- the new test is given only to women with a NPI above 3.4, based on the assumption that most women with an NPI below 3.4 would not be considered for chemotherapy.

This last group is considered to include the intermediate risk group that might benefit the most from the test, but also patients at the top end of the NPI distribution for whom the decision of chemotherapy is more certain. This subgroup was used because it was not possible to create an ‘intermediate only’ group by separating out the high NPI risk group. However, because the population is ER+, LN−, HER2−, it is rare to have a patient with a NPI over 5.4.

4.6 Interventions

**MammaPrint, Oncotype DX, IHC4 and Mammostrat**

The EAG defined four minimum criteria that a test had to fulfil to be included in the economic evaluation:

- The test has been validated in an external cohort (clinical validity).
• There is direct or indirect evidence about risk reclassification against one of the comparators defined within the scope (that is, NPI, Adjuvant! Online or clinical practice in the UK).
• The test provides an estimate of risk of recurrence in the form of a risk score or risk category. Tests that only provide information about sub-typing were excluded, following discussion with clinical experts, because it is not yet clear how sub-typing will impact on treatment decisions.
• The outputs of the test which will be used to inform the decision about whether or not to offer chemotherapy are well defined.

Overall, four tests (of the nine tests originally considered in this evaluation) met the criteria for economic evaluation defined by the EAG: Oncotype DX, MammaPrint, IHC4 and Mammostrat.

4.7 Comparator used in the economic model

Cancer registry data were used in the base-case analysis to reflect current levels of chemotherapy in England and Wales for women with ER+, LN−, HER2− early breast cancer. These data were used to represent current clinical practice in the UK which includes the use of NPI and Adjuvant! Online.

Clinical opinion indicated that although NPI and Adjuvant! Online are used to aid the decision-making process, the decision whether or not to offer adjuvant chemotherapy to a specific patient is complex and includes other demographic and pathological parameters. Consequently, the EAG economic assessment used cancer registry data to reflect the current clinical practice in England and Wales in terms of the proportion of women who currently receive chemotherapy. Summary data from the Eastern Cancer Registration and Information Centre (ECRIC) and the West Midland Cancer Intelligence Unit (WMCIU) were obtained to populate the economic model (personal communication). Characteristics of each database can be found in section 5.5.3.2 of the diagnostics assessment report.

Cancer registry data from the ECRIC and WMCIU were combined by the EAG and used in the base-case analysis to reflect the current levels of chemotherapy in England and Wales for women with ER+, LN−, HER2− early breast cancer. Data were
obtained for patients with a NPI ≤ 3.4 and patients with a NPI > 3.4 to perform a subgroup analysis and to account for the prognostic value of current treatment decision based on clinicopathological parameters. Registry data reflect how both NPI and Adjuvant! Online are used in current practice. However, it is not known which particular tools/local guidance (NPI or Adjuvant! Online, both, others) were used within these cancer registries. For the purposes of the economic model it is assumed that data from these two areas are representative of all trusts in England and Wales. The terminology 'clinical practice' will be used from now on to define the comparator selected for this assessment (that is, levels of adjuvant chemotherapy based on cancer registry data from the ECRIC and WMCIU).

4.8 Organisation of the economic analysis

The systematic review of evidence indicated that Oncotype DX is the furthest along the validation pathway compared with other similar tests and the evidence base, in particular in relation to prognostic ability, was reasonably sound. There is less evidence for IHC4, but there is evidence relating to the performance of IHC4 compared with Oncotype DX and so only a few additional assumptions were necessary to evaluate the cost effectiveness of IHC4. The final algorithm for IHC4 calculates a risk score for distant recurrence based on ER, PgR, HER2 and Ki67 in addition to classical clinical and pathological variables (composite risk score IHC4 plus clinical score). This version of the algorithm was considered in the economic analysis (the term IHC4 will be used for simplicity but refers to the composite risk score IHC4 plus clinical score). An online calculator is expected to be available at the beginning of 2012 (personal communication). The primary analysis therefore evaluated the cost effectiveness of adjuvant chemotherapy guided using Oncotype DX and IHC4 because this had the most robust evidence to populate the economic model.

Analyses were also performed for MammaPrint and Mammostrat and there were significant gaps and/or limitations in the evidence base available for both models (see section 5.5.5.4 of the diagnostics assessment report).

Three separate analyses were performed:
• Oncotype and IHC4 (used with current clinical practice)
• MammaPrint (used with current clinical practice)
• Mammostrat (used with current clinical practice)

All were modelled against current clinical practice and used the best direct sources of data available for each test. The results of these three models should not be directly compared. The basis of each model is different because the data came from different studies with different patient characteristics and methodologies. The EAG considered combining evidence from different studies into a single model, but concluded that limited conclusions could be drawn from such an analysis.

4.9 Modelling approach
The economic model aims to determine the cost effectiveness of the new tests relative to current practice by considering the difference in the proportion and risk levels of patients receiving chemotherapy (and the resultant costs and benefits of chemotherapy on short term and long term outcomes) between current practice (comparator arm) and practice with the addition of the new tests (intervention arm).

The economic model was constructed using three steps that were common to all three analyses, these steps are described below. In addition, details of the primary economic analysis (assessment of Oncotype DX and IHC4) are also provided as an example to help illustrate the steps. Full details of the economic model are provided in section 5.5.5 of the diagnostics assessment report.

• Step 1 - patients were assigned to risk categories according to the assigned risk score/group using the new test
• Step 2 - women who would receive chemotherapy as well as endocrine therapy were identified using the additional information from the new test (risk categories defined by the new test)
• Step 3 - breast cancer outcomes for patients treated with endocrine therapy alone or with the addition of chemotherapy was then simulated using a state transition model
4.10 Step 1 - Patients were assigned to risk categories according to the assigned risk score/group using the new test

*Common to all analyses*

All women in the model were assumed to be treated with endocrine therapy.

In this initial element, patients were categorised into different groups. NPI score and the output (or assumed output) of the new test(s) were used to categorise patients. Each patient is allocated to one of two NPI groups (NPI ≤ 3.4 or NPI > 3.4) and a risk group defined by the new test. For OncotypeDX and Mammostrat these subcategories are low-, intermediate- and high-risk. For MammaPrint the subcategories are poor and good prognosis. The IHC4 test provides a risk score only; however patients have been allocated into risk groups, similar to the OncotypeDX risk groups for the purposes of this assessment (discussed further below).

Each group has a different prognosis, defined in the model in terms of the probability of being free from distant recurrence at 10 years. The prognosis for each group, for patients on endocrine therapy alone, was calculated from the best available data for each test.

Notes:

(1) Grouping patients by NPI score allowed the EAG to explore the cost effectiveness of the new tests in different subgroups. In particular, the group of women with a NPI > 3.4 is of interest because this group is being used as a proxy for those at intermediate risk for whom the decision about whether or not to give chemotherapy is most uncertain and, therefore, who might benefit from the new test the most (further details in section 4.5 above). Another advantage of this initial categorisation is that it allows the EAG to take in to account how clinicopathological information, represented by the NPI distribution in the model, impacts the decision to offer chemotherapy (such that the proportion of women likely to receive chemotherapy in the NPI > 3.4 group is likely to be higher than the NPI ≤ 3.4 group).

(2) Information from the new test was used to categorise patients for the modelling exercise only and has no bearing on prognosis at this stage. Once categorised, the
prognosis (probability of being free from distant recurrence at 10 years) for each group was calculated. Categorisation of patients and calculation of prognosis was based on retrospective analysis of previous studies. Setting up the model in this way allows the baseline characteristics of those patients in the comparator arm (current clinical practice) and the intervention(s) arm (new tests) to remain the same.

**Primary economic analysis - Oncotype Dx and IHC4**

Women were first divided into two groups: those with a NPI ≤ 3.4 and women with a NPI > 3.4. Each of the two NPI groups were further subdivided using the Oncotype DX result (low-, intermediate- or high-risk) to give six groups of patients for the Oncotype DX evaluation. The proportion of the total population in each of these categories was computed as was the expected prognosis and effect of chemotherapy. Although IHC4 provides a continuous risk score, similar risk categories (low, intermediate and high) were developed for that test. Six similar groups were developed for IHC4 in conjunction with the two NPI groups. However, unlike Oncotype DX, there is no direct evidence of prognosis and chemotherapy effectiveness for these six groups. In order to compute the prognosis and chemotherapy effectiveness for these IHC4/NPI groups, data from the ATAC trial was used to link these groups to the relevant Oncotype DX/NPI groups. Data were available to show what proportion of those with low IHC4 scores would have low Oncotype DX scores, what proportion would have intermediate Oncotype scores and what proportion would have high Oncotype scores. Similar breakdowns for intermediate and high IHC4 scores were done. Outcome estimates for the six IHC4/NPI groups were based the outcomes that would arise from those expected for the Oncotype/NPI groups and the proportions that would exist in the IHC4/NPI group. Tables 51 and 52 of the diagnostics assessment report detail the group breakdowns. Information on the risk of distant recurrence at 10 years is available in tables 53 and 54 of the diagnostics assessment report.

Notes:

(3) Because Oncotype DX and IHC4 provide a continuous risk of recurrence score there are some limitations of assuming these tests classify patients into risk categories (that is, low, moderate and high risk):
• In the economic model, patients were classified according to the original cut-offs defined by the manufacturer of the technology; low (recurrence score <18), intermediate (recurrence score between 18 and 31) and high (recurrence score > 31).

• The IHC4 test does not present risk groups. The test is intended to be used as a continuous risk score and the interpretation is at the discretion of the clinician. For the purpose of the economic assessment, investigators provided risk classification evidence of IHC4 based on low, intermediate and high risk of distant recurrence. Cut-offs were defined using a similar approach to the classification with Oncotype DX (< 10%, 10 – 20%, and > 20% predicted risk of distant recurrence). The cut-offs used for IHC4 are therefore exploratory and were defined only to populate the economic model.

4.11 Step 2 - Women who would receive chemotherapy as well as endocrine therapy were identified using the additional information from the new test (risk categories defined by the new test)

Common to all analyses

In the comparator (current practice) arm the proportion of women receiving chemotherapy was based on cancer registry data, split for women with a NPI ≤ 3.4 and a NPI > 3.4 separately. When assessing the comparator, the risk categories assigned by the new tests would not be known so the probability of receiving chemotherapy is not influenced by the new test.

In the intervention arm the proportion of patients who would receive chemotherapy was based on new test results and the expected interpretation of the test (for example, women categorised at high risk of recurrence are far more likely to receive chemotherapy compared with women categorised at low risk).

Although the risk groups identified by the new test are expected to strongly influence the targeting of chemotherapy, other factors will also influence the decision to offer chemotherapy, including clinical and pathological factors and patient choice. An adjustment for such factors was therefore used in the model (that is, not all high risk patients receive chemotherapy).
Notes:

(4) At this stage, information from the new test is used to calculate the number of patients who would receive chemotherapy and, therefore, any impact on the patient’s prognosis (that is, the impact on the probability of being free from distant recurrence at 10 years). The proportion of patients receiving chemotherapy is the key difference between the comparator and intervention arms.

**Primary economic analysis - Oncotype Dx and IHC4**

To determine how the test results will translate into decisions about whether or not to give chemotherapy data from the only identified UK study were used for Oncotype DX in the base case. The Holt study reported the Oncotype DX recurrence score and the chemotherapy decision based on the recurrence score and traditional clinical and pathological parameters. This indicated that 89.5% of high risk score patients received chemotherapy, compared with 1.5% of low risk patients. Full results are presented in table 55 of the diagnostics assessment report.

No evidence on the proportion of women that would be given chemotherapy according to the results of the IHC4 test is available. It was assumed that the same proportions as for correspondingly named Oncotype DX would apply. This assumption was tested in a sensitivity analysis.

4.12 Step 3 - Breast cancer outcomes for patients treated with endocrine therapy alone or with the addition of chemotherapy was then simulated using a state transition model

**Common to all analyses**

Since the focus of the analysis is on the diagnostic test, costs and QALYs were summed to include the entire cohort of patients receiving the test. This cohort includes women who receive endocrine therapy only and those women who receive endocrine therapy plus chemotherapy. Long term outcomes of patients are influenced by the proportion of patients receiving each therapy and their risk levels.
Chemotherapy reduces the risk of breast cancer recurrence, but has cost implications (cost of treatment and cost of preventing or treating AEs) and a detrimental impact on quality of life.

The expected benefit of chemotherapy depends on the risk level of the patient as determined by NPI and other clinical factors. In addition there is a developing evidence base suggesting that risk groups defined by the new tests are predictive of chemotherapy benefit. That is, patients in the high risk group benefit more in relative terms than other patients. The benefit of chemotherapy (reduction in the risk of distant recurrence) was derived from studies evaluating the effect of chemotherapy by the risk category defined by each of the interventions (except IHC4). The benefit of chemotherapy was applied to both arms (comparator and intervention) of each analysis. For example, in the MammaPrint model the benefit of chemotherapy was derived for those with a poor prognosis and those with a good prognosis with MammaPrint. These were then applied to those patients receiving chemotherapy in the MammaPrint arm and also in the current practice arm (using the proportion of patients who were high risk and low risk by the MammaPrint test in the two NPI groups derived in element 1).

Outcomes associated with breast cancer were simulated using five health states. As shown in figure 2 (figure 11 from the diagnostics assessment report), patients enter the model in the recurrence free survival health state (A) and remain in that health state until developing a distant recurrence (B), having an adverse event after chemotherapy (D) or dying from general causes (E). After a distant recurrence (B), patients remain in this health state until dying from either breast cancer or general causes (E) or developing an adverse event for women treated with chemotherapy (D). Patients developing an adverse event after chemotherapy can remain in that health state, die from their adverse event or die from general causes (E).
The estimation of long-term adverse events is simplistic. No distinction was made for patients developing long-term adverse events after a recurrence (B) or in the recurrence free health state (A). Furthermore, patients with a long-term adverse event were assumed to remain in that health state (D) until death (E) and were not allowed to move to other health states, in particular, it seems that moving from a long-term adverse event to distant recurrence isn’t possible.

Local/regional recurrences have been modelled by considering the cost and quality of life decrements (disutility) assuming that a proportion of patients entering the distant recurrence state (B) have previously experienced a local recurrence (C).

*Primary economic analysis* - Oncotyple DX and IHC4

The base-case analysis used data from the Paik et al. study (2006) to estimate the benefit of chemotherapy by Oncotyple DX risk categories. The relevance of this study to the model population is not certain. In the overall population, the addition of chemotherapy compared with tamoxifen alone was estimated to reduce the risk of distant recurrence by 44% (HR 0.56, 95% CI 0.34 to 0.91). No chemotherapy benefit was found for women classified as low risk of distant recurrence with Oncotyple DX (HR 1.31, 95% CI 0.46 to 3.78, p = 0.61). A reduction of 39% (HR 0.61, 95% CI 0.24 to 0.91) was estimated for women with high Oncotyple DX risk.

**Figure 2: Schematic of the Markov model structure**

![Markov model diagram](image)
to 1.59, p = 0.39) and 74% (HR 0.26, 95% CI 0.13 to 0.53, p < 0.001) was found for the risk of distant recurrence for women classified as intermediate and high risk of distant recurrence with Oncotype DX who received chemotherapy in addition to tamoxifen compared with tamoxifen alone.

These levels of chemotherapy benefit were applied, indirectly, to the comparator arm and the IHC4 arm (data on chemotherapy benefit were not available for IHC4).

4.13 Assumptions used in the economic analysis

General assumptions used in the base-case analysis are detailed in section 5.5.5.4 and specific assumptions used to model each test are detailed in table 64 of the diagnostics assessment report. Specific assumptions used in the primary economic analysis (Oncotype DX and IHC4) are detailed below.

**OncotypeDX**

- Oncotype DX was modeled using three discrete risk groups (low, intermediate, high recurrence score) instead of a continuous recurrence score.
- Original recurrence score groups are used – low (RS < 18), intermediate (RS between 18 and 31), high (RS > 31).
- Reclassification evidence from the ATAC trial (UK population) by NPI, by Oncotype DX and by IHC4.
- Risk of recurrence from the ATAC trial (UK population) by NPI, by Oncotype DX and by IHC4.
- Holt study (UK population) to inform the proportion of patients receiving chemotherapy classified as low, intermediate or high risk of recurrence.
- Benefit of chemotherapy in terms of reduction of distant recurrence taken from Paik et al. (2006) in a US cohort treated with cyclophosphamide, methotrexate and 5FU or cyclophosphamide and methotrexate. The study included a large proportion of pre-menopausal women and some women with HER2+ breast cancer. No benefit was assumed for women with a low Oncotype DX recurrence score.
IHC4

- Composite score of IHC4 in addition to clinicopathological parameters – IHC4 plus clinical score.
- The test was assumed to be reproducible (notably the Ki 67 element).
- IHC4 was modeled using three discrete risk groups (low, intermediate, high risk) instead of continuous risk of distant recurrence.
- Patients with a low, intermediate and high risk of distant recurrence were defined as having a predicted risk of distant recurrence < 10%, 10–20%, and > 20% (similar approach to Oncotype DX).
- Reclassification evidence from the ATAC trial (UK population) by NPI, by Oncotype DX and by IHC4.
- Risk of recurrence from the ATAC trial (UK population) by NPI, by Oncotype DX and by IHC4.
- Clinicians interpret Oncotype DX and IHC4 the same way. The Holt study (UK population) was used to inform the proportion of patients receiving chemotherapy classified as low, intermediate or high risk.
- The benefit of chemotherapy in terms of reduction of distant recurrence was applied from the RS risk group (using Oncotype DX – Paik et al. study) because there were no specific data for IHC4. This assumes that the reclassification with IHC4 does not provide any additional benefit (possibly a conservative assumption).
- An additional cost of £100–200 to perform the IHC4 test over and above current practice.

4.14 Results of the de novo economic analysis

The primary analysis compared current clinical practice with treatment guided by Oncotype DX and IHC4. All analyses assumed that the new tests were used in addition to current prognostic tools.

In addition to the primary economic analysis, additional economic analyses were undertaken for Mammostrat and MammaPrint. There are significant limitations in the evidence base, concerns about the assumptions made and the robustness of the evidence used in these additional models.
Full details can be found in section 5.6 of the diagnostics assessment report. A brief summary of the key results (including ICERs (incremental cost-effectiveness ratios) of the base-case analysis and sensitivity analyses is presented below.

Two analyses are presented:

- the tests were used for all women with ER+, LN–, HER2– early breast cancer
- the tests were used only for women with ER+, LN–, HER2– early breast cancer with a NPI score above 3.4 (on the assumption that most women with a NPI below 3.4 would not be considered for chemotherapy).

The base case results were computed two different ways. The first way was deterministic with the model parameters set at their central values. The second way was probabilistic. The model was run multiple times with some parameters having a different value randomly selected based on its probability distribution on each run. The expected value (mean) of the results of the multiple runs are then presented.

4.15 Cost-effectiveness results – Primary economic analysis - Oncotype DX and IHC4

All women with ER+, LN–, HER2– early breast cancer

Deterministic results

For a cohort of 1000 women with ER+, LN–, HER2– early breast cancer, Oncotype DX was predicted to lead to an increase in the proportion of patients receiving chemotherapy under the base case assumptions (19.11% vs. 14.42%) compared with current clinical practice. Treatment guided using IHC4 was predicted to lead to a reduction in the proportion of patients receiving chemotherapy (9.57% vs. 14.42%) under the base case assumptions. More chemotherapy was given for Oncotype DX because more women were classified as high or intermediate risk with Oncotype DX compared with IHC4, and therefore were more likely to receive chemotherapy. It was predicted that 76 distant recurrences would occur under current clinical practice. Guiding treatment using Oncotype DX and IHC4 was predicted to reduce the number of distant recurrence to 64 and 71 respectively, with the assumptions used for the base-case analysis. The results of the deterministic analysis are shown in table 3.
Table 3 Deterministic ICER for the primary analysis comparing Oncotype DX and IHC4 with current clinical practice assuming the test to be used for all women with ER+, LN–, HER2– early breast cancer in England and Wales

<table>
<thead>
<tr>
<th></th>
<th>Mean Cost (£)</th>
<th>Mean QALYs</th>
<th>ICER</th>
<th>Incremental analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oncotype DX</td>
<td>£9094</td>
<td>13.54</td>
<td>£26,940*</td>
<td>£55,406*</td>
</tr>
<tr>
<td>IHC4</td>
<td>£6340</td>
<td>13.49</td>
<td>Dominant</td>
<td></td>
</tr>
<tr>
<td>Current clinical practice</td>
<td>£6519</td>
<td>13.44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. *rounding error contributes to the difference from expected

Probabilistic results

Results for the probabilistic analysis are shown in table 4.

Table 4 Probabilistic ICER (2500 iterations) for the primary analysis comparing Oncotype DX and IHC4 with current clinical practice assuming the test to be used for all women with ER+, LN–, HER2– early breast cancer in England and Wales

<table>
<thead>
<tr>
<th></th>
<th>Mean Cost (£)</th>
<th>Mean QALYs</th>
<th>ICER</th>
<th>Incremental analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oncotype DX</td>
<td>£9100</td>
<td>13.52</td>
<td>£29,503*</td>
<td>£64,111*</td>
</tr>
<tr>
<td>IHC4</td>
<td>£6332</td>
<td>13.48</td>
<td>Dominant</td>
<td></td>
</tr>
<tr>
<td>Current clinical practice</td>
<td>£6507</td>
<td>13.44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b. *rounding error contributes to the difference from expected

At a threshold of £20,000 per QALY gained the probability of IHC4 being cost effective compared with Oncotype DX and current clinical practice was 99.48%. The probability of treatment guided using Oncotype DX being cost effective at a £20,000 threshold was 0.4% when compared with IHC4 and current clinical practice.

The probability for treatment guided using Oncotype DX to be cost effective at a £20,000 threshold when compared with current clinical practice alone was 12.44%.
Women with ER+, LN–, HER2– early breast cancer with a NPI above 3.4

Deterministic results

Assuming the tests were offered only to women with ER+, LN–, HER2– early breast cancer with a NPI above 3.4, a greater proportion of patients were predicted to receive chemotherapy when using Oncotype DX (under the base-case assumptions) and a lower proportion using IHC4 compared with current clinical practice (34.72%, 26.31% and 33.60% respectively). The results of the deterministic analysis are shown in table 5.

Table 5 Deterministic ICER for the primary analysis comparing Oncotype DX and IHC4 with current clinical practice assuming the test to be used for women with a NPI > 3.4 only

<table>
<thead>
<tr>
<th></th>
<th>Mean Cost (£)</th>
<th>Mean QALYs</th>
<th>ICER</th>
<th>Incremental analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oncotype DX</td>
<td>£10,911</td>
<td>13.06</td>
<td>£9,007*</td>
<td>£26,859*</td>
</tr>
<tr>
<td>IHC4</td>
<td>£8318</td>
<td>12.97</td>
<td>Dominant</td>
<td></td>
</tr>
<tr>
<td>Current clinical practice</td>
<td>£8816</td>
<td>12.83</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*rounding error contributes to the difference from expected

Probabilistic results

Results for the probabilistic analysis are shown in table 6.
Table 6 Probabilistic ICER for the primary analysis comparing Oncotype DX and IHC4 with current clinical practice assuming the test to be used for women with a NPI > 3.4 only

<table>
<thead>
<tr>
<th></th>
<th>Mean Cost (£)</th>
<th>Mean QALYs</th>
<th>ICER</th>
<th>Incremental analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oncotype DX</td>
<td>£10,924</td>
<td>13.05</td>
<td>£9,774*</td>
<td>£31,125*</td>
</tr>
<tr>
<td>IHC4</td>
<td>£8305</td>
<td>12.96</td>
<td>Dominant</td>
<td></td>
</tr>
<tr>
<td>Current clinical practice</td>
<td>£8797</td>
<td>12.83</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

d. *rounding error contributes to the difference from expected

At a threshold of £20,000 per QALY gained the likelihood of IHC4 being cost effective compared with Oncotype DX and current clinical practice was 81.24%. The probability of treatment guided using OncotypeDX being cost effective at a £20,000 threshold was 18.60% when compared with IHC4 and current clinical practice.

The probability for treatment guided using Oncotype DX to be cost effective at a £20,000 threshold when compared to current clinical practice alone was 91.56%.

Univariate sensitivity analyses – parameters
A range of univariate sensitivity analyses were carried out to explore the impact of varying the main model parameters. Results of the univariate sensitivity analysis assuming the tests were offered to all women with ER+, LN–, HER2– are presented in Table 70 of the diagnostics assessment report. Results for the univariate sensitivity analysis assuming the test were offered to women with a NPI above 3.4 only are presented in Table 71 of the diagnostics assessment report.

Main model parameters were varied within reasonable ranges. The ICER for Oncotype DX compared with current clinical practice was mainly sensitive (defined as changes in the ICER by 10% or more) to the assumptions about the time horizon, the starting age of the cohort, the risk of recurrence, the proportion of patients receiving chemotherapy after reclassification with the new test, the benefit of chemotherapy and the NPI distribution.
The ICER increased (less favourable to Oncotype DX) as the time horizon decreased or the age increased, because fewer benefits can be accrued over time. Similarly, a reduction in the risk of distant recurrence increased the ICER (less favourable to Oncotype DX) whereas an increase in the risk of distant recurrence reduced the ICER. Given that more recurrences can be avoided if treatment is offered to patients with an increased risk of distant recurrence, additional benefits (QALYs) accrue.

Furthermore, the ICER was very sensitive to the assumption about the probability of chemotherapy in patients classified as intermediate with Oncotype DX. The probability of patients at intermediate risk receiving chemotherapy was set equal to that of high risk patients and low risk patients in turn; the ICER was £22,818 and £35,629 respectively if the test was given to all women and £8,371 and £10,022 respectively if the test was given to women with a NPI above 3.4 only.

The ICER also declined (more favourable to Oncotype DX) when using the NPI distribution from the Holt study because more patients were classified as having an NPI > 3.4. This group of patients was shown to derive a greater benefit from the new tests.

Finally, the ICER increased when a lower benefit of chemotherapy was assumed (benefit reduction of 20% to 40%). The ICER also increased when assuming patients classified as low, intermediate or high risk according to the Oncotype DX recurrence score classification had the same relative reduction (in the risk of distant recurrence) as a result of chemotherapy.

The ICER for IHC4 was sensitive to more assumptions (such as the time spent in the distant recurrence health state, the proportion of patients receiving chemotherapy under current clinical practice and the cost of chemotherapy), but IHC4 remained dominant compared with current clinical practice (that is, it provided more QALYs at a lower cost) except when the cost of IHC4 was set at £400 (when the ICER was £1557 per QALY gained against current clinical practice).
Univariate sensitivity analyses – structural assumptions
In addition to input parameter values, the EAG also examined the impact of two structural assumptions; the exclusion of IHC4 and the impact of modelling patients as a single group (instead of two groups with NPI ≤ 3.4 or NPI >3.4).

Assuming no further reclassification using IHC4 (exclusion of IHC4)
In this scenario analysis, the EAG used data for Oncotype DX only (table 49 of the diagnostics assessment report presents data for the risk classification) from the ATAC trial, assuming no further reclassification with IHC4.

In this scenario analysis, patients were split into six possible risk categories (by NPI and Oncotype DX recurrence score) compared with 18 risk categories in the base-case model (by NPI, Oncotype DX recurrence score and IHC4).

The impact on the ICER was minimal with a reduction from £26,940 (base case) to £25,574 per QALY gained assuming the test to be used for all women with ER+, LN–, HER2– early breast cancer or an increase from £9007 (base case) to £10218 per QALY gained assuming the test to be used for women with a NPI > 3.4 only.

Results of this scenario analysis suggested that the base-case ICER for Oncotype DX was minimally affected by the choice of model structure to accommodate the evaluation of IHC4.

Assuming no further reclassification using IHC4 (exclusion of IHC4) and modelling the entire cohort as a single group (no split by NPI)
A second scenario analysis investigated the extent to which not separating patients by NPI, and therefore assuming that patients with a NPI ≤ 3.4 and patients with a NPI > 3.4 have the same prognosis affected the ICER.

Again, in this scenario analysis data for Oncotype DX only (table 49 of the diagnostics assessment report presents data for the risk classification) from the ATAC trial were used, assuming no further reclassification with IHC4. In this scenario analysis, the model separated patients into three possible risk categories (by Oncotype DX recurrence score only) compared with 18 risk categories in the base-case model.
As expected, this assumption had a positive impact (more favourable to Oncotype DX) on the ICER with a reduction from £26,940 (base case) to £18,859 per QALY gained assuming the test is given to all women with ER+, LN–, HER2– early breast cancer.

By modelling the entire cohort as a single group, the prognostic value of current decision-making using clinicopathological parameters is ignored (that is, patients with a low NPI have a lower risk of recurrence but are also less likely to receive chemotherapy than patients with a NPI > 3.4) and therefore the analysis is more favourable to Oncotype DX. This scenario assumed that patients within the defined Oncotype DX recurrence score group are homogenous; however, it is feasible that patients with a low Oncotype DX recurrence score and low NPI would have a better prognosis than patients with a low Oncotype DX recurrence score but high NPI.

4.16 Cost-effectiveness results - Additional analyses - Mammostrat

Compared with Oncotype DX, the evidence base for Mammostrat was less developed and some gaps were identified by the systematic review of the literature. Furthermore, the EAG’s economic model considered distant recurrence as an outcome, whereas most of evidence for Mammostrat was drawn from analyses of disease-free survival and therefore included all recurrences.

An analysis was carried out to assess the cost effectiveness of Mammostrat compared with current clinical practice in England and Wales. The EAG’s economic model was repopulated using the evidence for Mammostrat for reclassification (unpublished) and benefit of chemotherapy, but assumptions were necessary because of limitations of the evidence, especially about the interpretation of the test and whether the reclassification data can be applied to the UK population. Data from a subset of the Ring study (2006) were used in the economic model; however, based on these data Mammostrat appeared to have little value in patients with a NPI > 3.4. Results of the deterministic analysis are shown in table 5.

Because of the gap in the evidence base, any conclusions drawn from this analysis are subject to uncertainty Notably there is a major uncertainty regarding the robustness of the data from the subset of the Ring study (2006).
Table 5 Deterministic ICER for the analysis comparing Mammostrat with current clinical practice in women with ER+, LN−, HER2− early breast cancer in England and Wales

<table>
<thead>
<tr>
<th></th>
<th>Mean Cost (£)</th>
<th>Mean QALY</th>
<th>ICER</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mammostrat</td>
<td>£9040</td>
<td>12.91</td>
<td>£26,598*</td>
</tr>
<tr>
<td>Current clinical practice</td>
<td>£7699</td>
<td>12.86</td>
<td></td>
</tr>
<tr>
<td>NPI &gt; 3.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mammostrat</td>
<td>£10,985</td>
<td>12.29</td>
<td>Dominated</td>
</tr>
<tr>
<td>Current clinical practice</td>
<td>£9717</td>
<td>12.34</td>
<td></td>
</tr>
</tbody>
</table>

*rounding error contributes to the difference from expected

4.17 Cost-effectiveness results - Additional analyses - MammaPrint

An analysis was also carried out to assess the cost effectiveness of MammaPrint. Although MammaPrint had more evidence than Mammostrat, there were significant gaps and data that were used to populate the economic model were considered to be less robust. Therefore, any conclusions from this analysis are subject to considerable uncertainty.

The EAG expressed particular concerns about the existing evidence on the benefit of chemotherapy because this is likely to have a significant influence the ICER. There were also other issues, notably the lack of UK data and the fact that the data available were derived mainly from pre-menopausal women. Because of these issues, the EAG was not able to estimate a single ICER, but presented a range of ICERs based on the confidence interval for the benefit of chemotherapy, which was believed to be the main uncertainty in the model. Although a range of ICERs is presented, there were also significant concerns about the study design, which were not addressed by presenting ICERs as a range.

More assumptions have been made within this analysis than with the other analyses, and results are more uncertain. Only a limited number of univariate sensitivity analyses were carried out. A sensitivity analysis was conducted assuming no additional costs for the NHS for the use of fresh frozen tissue. A second sensitivity
analysis was conducted assuming that 5% of patients classified as having a good prognosis and 95% of patients classified as having a poor prognosis received chemotherapy.

In addition to the univariate sensitivity analyses, a multivariate sensitivity analysis was performed to examine different values for the benefit of chemotherapy.

No probabilistic sensitivity analysis was conducted because the EAG considered there were too many uncertainties in the studies used. These mostly related to the quality of the study design and the fact that populations in the studies were younger and at higher risk than the population in the economic model, uncertainties which could not be adequately captured by the parameter uncertainty within probabilistic sensitivity analysis.

**Deterministic results**

Compared with current clinical practice, the incremental cost for treatment guided using MammaPrint was estimated to range between £12,240 and £53,058 per QALY gained within the reported confidence interval for the chemotherapy reduction from the Knauer (2010) study assuming the test was used for all women with ER+, LN–, HER2– early breast cancer. If MammaPrint was only used for women with a NPI above 3.4, the ICER ranged between £6053 and £29,569 per QALY gained.

The proportion of patients receiving chemotherapy increased significantly with the use of MammaPrint under our base-case assumptions: from 14.42% to 44.18% for all women and from 33.60% to 90.31% for women with a NPI > 3.4. Results of the deterministic analysis are shown in table 6 (results include an assumed additional cost of £250 per patient for handling frozen tissue samples).
Overview – Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in breast cancer management

Table 6 Deterministic ICER for the analysis comparing MammaPrint with current clinical practice in women with ER+, LN–, HER2– early breast cancer in England and Wales

<table>
<thead>
<tr>
<th></th>
<th>Mean QALYs</th>
<th>Mean cost (£)</th>
<th>ICER (£/QALYs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current practice</td>
<td>13.49 - 13.39</td>
<td>£6408 to £6629</td>
<td></td>
</tr>
<tr>
<td>MammaPrint</td>
<td>13.78 - 13.47</td>
<td>£10,017 - £10,748</td>
<td>£12,240 to £53,058*</td>
</tr>
<tr>
<td>NPI &gt; 3.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current practice</td>
<td>13.07 - 12.81</td>
<td>£8,281 to £8,872</td>
<td></td>
</tr>
<tr>
<td>MammaPrint</td>
<td>13.73 - 12.99</td>
<td>£12,278 to £14,014</td>
<td>£6053 to £29,569</td>
</tr>
</tbody>
</table>

f. contributes to the difference from expected

*rounding error

5 Issues for consideration by the Committee

In order to draft recommendations on these diagnostic technologies, the Committee will need to consider whether, by guiding the selection of patients to receive adjuvant chemotherapy, gene expression profiling and expanded immunohistochemistry tests improve health outcomes and quality of life in patients with early stage breast cancer compared with currently used decision-making protocols.

Each of the four tests included in the evaluation has varying levels of supporting evidence regarding their analytic validity, clinical validity and clinical utility. The Committee needs to consider these data carefully as the outcomes of such consideration are likely to impact the Committee’s recommendations for each of the four tests. For example, the base-case analyses assume a differential benefit of chemotherapy dependent on the risk group assigned by the test in question. The quantity and quality of evidence on the differential benefit of chemotherapy varies significantly between the four tests from little or no evidence to the availability of multiple studies.

Another factor relevant to the Committee’s decision is the processing options available for the tests, which range from local processing in the hospital (IHC4) to

National Institute for Health and Clinical Excellence

Overview – Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in breast cancer management

Issue date: December 2011
The manufacturer of Mammostrat is in the process of establishing a UK-based central processing laboratory. They aim to have the laboratory set-up by February 2012.

The MammaPrint test can currently be used on fresh tissue samples only, this may make implementation more complex (see section 7). From January 2012 the manufacturers of MammaPrint will accept Formalin Fixed Paraffin Embedded samples for the MammaPrint test.

Reproducibility of IHC4 will need to be taken into account when considering use of the test in local UK laboratories. Variability in IHC results can occur as a result of variability in several factors including fixation, antigen retrieval, reagents and interpretation. Guidance has recently been published to help standardise the measurement of Ki-67.

Most of the tests require samples to be sent to central processing laboratories and therefore time delays may be imposed on management.

Economic analysis:

Three separate economic models were created: Oncotype DX and IHC4, MammaPrint, Mammostrat. These models cannot be directly compared with each other because the methodologies employed and patient characteristics of the studies used to populate each model vary.
- The primary economic analysis was of Oncotype DX and IHC4 (when added to current clinical practice) against current clinical practice.

- The analyses of MammaPrint and Mammostrat were considered exploratory by the EAG because of the amount or quality of evidence identified in the systematic review and the results of these analyses should be interpreted with consideration.

- In each of the three separate analyses a subgroup analysis of those patients most likely to benefit from gene expression profiling and expanded immunohistochemistry tests was performed (NPI > 3.4).

- IHC4 was shown to dominate current clinical practice in the primary economic model. However, these results are based on low levels of evidence and assumptions on how the test may be used and interpreted in the clinic. It is not clear if the assumptions regarding the use of the test accurately reflect how the test will be used in the NHS.

Some gene expression profiling and expanded immunohistochemistry tests classify a proportion of patients into an intermediate risk category. Evidence for the benefit of chemotherapy (reduction in the risk of recurrence) in those patients is less clear. Whether clinicians would recommend chemotherapy in addition to endocrine therapy for patients classified as intermediate risk is unknown.

MammaPrint and Mammostrat classify patients into risk groups only (categorical) and do not calculate a continuous risk score. This is likely to be less informative than a continuous risk score because all patients are either classified as low, intermediate or high risk. This does not differentiate between patients who are at the lower end or upper end of the distribution and those who are borderline (which may impact the decision to give or withhold chemotherapy), although it makes the test easier to use and interpret in practice.
Immunohistochemistry tests such as Mammostrat offer the advantage that biomarker expression is interpreted in situ which allows the pathologist to ensure that the test is not confounded by expression of biomarker in non-tumour tissue. Gene expression assays that require homogenisation of the tissue and measure biomarkers that may be expressed in stroma, run a greater risk of confounding the interpretation of biomarker expression levels which may lead to higher error rates.

The Diagnostics Assessment Programme methods state that DAC does not use a precise Incremental Cost Effectiveness Ratio threshold below or above which a technology would automatically be defined as cost effective or not. NICE considers instead that it is most appropriate to use a threshold range of £20,000-£30,000. The influence of other factors on the decision to recommend a technology is greater if the ICER is closer to the top of the range. Below an ICER of £20,000 per QALY gained, the decision to recommend the use of a technology is normally based on the cost-effectiveness estimate and the acceptability of a technology as an effective use of NHS resources. If the estimated ICERs presented are less than £20,000 per QALY gained and the Committee considers that the tests should not be provided by the NHS, the recommendations made by the Committee should make specific reference to its view on the plausibility of the inputs to the economic modelling and/or the certainty around the estimated ICER.

6 Equality considerations

The incidence of breast cancer varies with ethnicity. Asian, Chinese and black ethnic groups and those with mixed heritage have a lower incidence than the white ethnic group in England. Incidences are 0.65, 0.75, 0.49 and 0.58 that of the white ethnic group respectively. People with cancer are protected under the Equality Act 2010, from the point of diagnosis.

7 Implementation issues

Use of fresh tissue

Collection of fresh tissue, which is currently required for MammaPrint (the manufacturer will accept Formalin Fixed Paraffin Embedded samples from January
is not routine in the NHS. There would be additional costs and these could be considerable at hospitals in which dissection facilities are already filled to capacity (which is likely to be a significant proportion of hospitals) and there are no designated staff for collection of fresh tissue. Discussion with local clinicians indicated that capital costs could be at least £75,000 per hospital if new dissection tables are needed. If fresh tissue sampling is not routine (only a few research centres currently have this working arrangement), then additional staff costs for biomedical scientists and histopathologists could be incurred. If a full fresh tissue service was needed to cover all theatre time then additional staff costs could be £20,000 to £50,000 per year.

8 Summary

Four tests that aim to assist clinicians in identifying breast cancer patients for adjuvant chemotherapy have been included in this assessment. These tests are at different stages of development. There is a variable amount of evidence. The economic analysis performed by the EAG assessed the cost-effectiveness of the four tests for women with ER+, LN–, HER2– early breast cancer. The most robust evidence was available for Oncotype DX and was considered the primary economic analysis. IHC4 was also included in this analysis as evidence relating to the performance of IHC4 compared with Oncotype DX was available. Separate economic analyses were performed for Mammostrat and MammaPrint.

The results of the primary economic analysis showed that IHC4 was dominant when compared to current clinical practice when either offered to all women or restricted to those with an NPI > 3.4. However, a range of assumptions (in particular, regarding the interpretation of the test in the clinic), were required. Results of the primary economic analysis showed that Oncotype Dx resulted in a deterministic ICER of £26,940 per QALY gained when compared to current clinical practice and when offered to all women. The ICER drops to £9,007 per QALY gained when Oncotype Dx is only offered to women with an NPI > 3.4.

The analysis of Mammostrat was based on less robust evidence and subject to an anomaly in that chemotherapy was beneficial for high and low risk patients, but not...
moderate risk ones. This meant that Mammostrat was dominated for the NPI > 3.4 cohort and an ICER of £26,598 for the entire cohort.

The data for MammaPrint were also less robust than the data for Oncotype DX. Ranges of ICERS were computed that were broad. For the cohort with NPI > 3.4 the range ran from £6053 to £29,569. For the cohort including all patients, the ICER was higher running from £12,240 to £53,058.

This overview was prepared by:

Gurleen Jhuti
Topic Lead

Hanan S. Bell, PhD
Technical Adviser

December 2011
Appendix A: Sources of evidence considered in the preparation of the overview

A The diagnostics assessment report for this evaluation was prepared by the Aberdeen Health Technology Assessment.


B The following organisations accepted the invitation to participate in this assessment as stakeholders. They were invited to attend the scoping workshop and to comment on the diagnostics assessment report.

I. Manufacturers/sponsors:

- Royal Marsden Hospital and Queen Mary University London - academic sponsor (IHC4)
- Agendia Bvd (MAMMAPRINT)
- Clarient (MAMMOSTRAT)
- Genomic Health (ONCOTYPE)

II. Professional/specialist and patient/carer groups:

- Breakthrough Breast cancer
- Bupa
- Roche Diagnostics
- Royal College of Physicians
- Southport & Ormskirk Hospital NHS Trust