National Institute for Health and Care Excellence

DIAGNOSTICS ASSESSMENT PROGRAMME

Evidence overview

Tumour profiling tests to guide adjuvant chemotherapy decisions in people with breast cancer

(update of DG10)

This overview summarises the key issues for the diagnostics advisory committee's consideration. This document is intended to be read in conjunction with the final scope issued by NICE for the assessment and the diagnostics assessment report. A glossary of terms can be found in Appendix B. Academic in confidence information is marked [redacted]. Commercial-in-confidence information is marked [redacted].

1 Background

1.1 Introduction

The purpose of this assessment is to update NICE diagnostics guidance 10 on gene expression profiling and expanded immunohistochemistry tests for guiding adjuvant chemotherapy decisions in early breast cancer management. The clinical and cost effectiveness of 5 tumour profiling tests used to guide adjuvant chemotherapy decisions in people with breast cancer have been evaluated. Provisional recommendations on the use of these technologies will be formulated by the diagnostics advisory committee at the committee meeting on 30 November 2017.

This guidance will update the existing guidance, which included the following recommendations:

- Oncotype DX is recommended as an option for guiding adjuvant chemotherapy decisions for people with oestrogen receptor positive (ER+), lymph node negative (LN−) and human epidermal growth factor receptor 2 negative (HER2−) early breast cancer if:
  - the person is assessed as being at intermediate risk and
  - information on the biological features of the cancer provided by Oncotype DX is likely to help in predicting the course of the disease and would therefore help when making the decision about prescribing chemotherapy and
  - the manufacturer provides Oncotype DX to NHS organisations according to the confidential arrangement agreed with NICE.
- NICE encourages further data collection on the use of Oncotype DX in the NHS.
- MammaPrint, IHC4 and Mammostrat are only recommended for use in research in people with ER+, LN− and HER2− early breast cancer, to collect evidence about potentially important clinical outcomes and to determine the ability of the tests to predict the benefit of chemotherapy (see section 7). The tests are not recommended for general use in these people because of uncertainty about their overall clinical benefit and consequently their cost effectiveness.

This update has been done according to the standard update process.

Tumour profiling tests are designed to provide information on the activity of genes within tumour samples from people with early breast cancer. The results of the tests provide a risk profile of an individual's breast cancer which can be combined with other clinical risk factors that are routinely assessed, such as nodal status and tumour size. It is claimed that the risk profile can be used to better predict the risk of disease recurrence in the future. Some tests also claim to predict the benefit a patient may receive from chemotherapy.

This information is intended to help treatment decision-making with regard to adjuvant chemotherapy use.

Tumour profiling tests may improve the identification of people with early breast cancer who may not benefit from having adjuvant chemotherapy because they have a low risk of disease recurrence. These people could avoid unnecessary treatment, and would therefore not be exposed to the comorbidities and negative impacts on quality of life that are associated with chemotherapy. Additionally, people with early breast cancer who have been identified as at low risk of disease recurrence could be reclassified as being at high risk of recurrence, and therefore may benefit from chemotherapy. People with breast cancer and clinicians may also benefit from improved confidence in the appropriateness of the treatment they are having or recommending.

1.2 Scope of the assessment

Table 1 Scope of the assessment

Decision question	Do tumour profiling tests used for guiding adjuvant chemotherapy decisions for people with early stage breast cancer (described in section 1.6 of NICE CG80) represent a clinically and cost-effective use of NHS resources?
Populations	People with oestrogen receptor positive (ER+) (and/or progesterone receptor positive [PR+]), human epidermal growth factor receptor 2 negative (HER2−), early stage breast cancer (stages I or II) with 0 to 3 positive lymph nodes. Where data permits, the following subgroups may be considered: • people with lymph node negative cancer; people with micrometastases in the lymph nodes; and people with 1 to 3 positive lymph nodes • premenopausal women and postmenopausal women • people predicted to be at low, intermediate or high risk using a risk assessment tool, or using clinical and pathological features • men and women • people of different ethnicities.
Interventions	• EndoPredict

	• MammaPrint • Oncotype DX Breast Recurrence Score • Prosigna • IHC4 In combination with current decision-making.
Comparators	• Current decision-making, which may include any tool, or clinical and pathological features, used to assess risk.
Healthcare setting	Secondary and tertiary care.
Outcomes	Intermediate measures for consideration may include: • time to test results • analytical validity • prognostic ability • ability to predict benefit from chemotherapy • impact of test results on decision-making.
	Clinical outcomes for consideration may include: • disease-free survival • overall survival • distant recurrence • disease-related morbidity and mortality • chemotherapy-related morbidity and mortality.
	Patient-reported outcomes for consideration may include: • health-related quality of life • anxiety.
	Costs will be considered from an NHS and Personal Social Services perspective. Costs for consideration may include: • costs of treating breast cancer, including: drug cost, administration cost, outpatient appointments, and treatment of adverse events • costs of the tests, including equipment costs and reagents when relevant • costs of staff and associated training.
	The cost effectiveness of interventions should be expressed in terms of incremental cost per quality-adjusted life year.
Time horizon	The time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.

Further details including descriptions of the interventions, comparators, care pathway and outcomes can be found in the final scope.

2 The evidence

This section summarises data from the diagnostics assessment report compiled by the External Assessment Group (EAG).

2.1 Clinical effectiveness

The EAG did a systematic review of the evidence on the effectiveness of 5 tumour profiling tests: EndoPredict (EPClin score unless EP score is specifically mentioned), IHC4/IHC4+C, MammaPrint, Oncotype DX Breast Recurrence Score (hereafter referred to as Oncotype DX) and Prosigna. A summary of the test characteristics is presented in table 2.

Table 2 Characteristics of tests included in the scope

Test	EndoPredict	MammaPrint	Oncotype DX	Prosigna	IHC4/IHC4+C
Purpose	Predict recurrence risk	Predict recurrence risk and chemotherapy benefit	Predict recurrence risk and chemotherapy benefit	Predict recurrence risk and intrinsic subtype[a]	Predict recurrence risk
Description	EP score = 12 gene assay (RT-qPCR) EPClin score = EP score + clinical factors	Microarray 70 gene array	RT-qPCR 21 gene assay	Direct mRNA counting + clinical factors 50 gene assay	IHC4 = 4 IHC tests IHC4+C = 4 IHC tests + clinical factors
Testing location	Local laboratory or test service (Germany)	Test service (the Netherlands)	Test service (US)	Local laboratory or test service (UK)	Local laboratory
Menopausal status	Pre- and postmenopausal	Pre- and postmenopausal	Pre- and postmenopausal	Postmenopausal	Postmenopausal
Test result	Low risk, high risk	Low risk, high risk	Low risk, intermediate risk, high risk	Low risk, intermediate risk, high risk	Low risk, intermediate risk, high risk

[a] Evidence relating to intrinsic subtype was not reviewed in this assessment.

Abbreviations: EP, EndoPredict; IHC, immunohistochemistry; RT-qPCR, reverse transcriptase quantitative polymerase chain reaction.

The methods of the systematic review can be found starting on page 45 of the diagnostics assessment report. Evidence on the following outcomes was of interest in the clinical-effectiveness review:

- Prognostic ability – the degree to which the test can accurately predict the risk of an outcome such as disease recurrence.
- Prediction of chemotherapy benefit - chemotherapy benefit relates to the ability of the test to predict which patients will respond to chemotherapy, and can be assessed by considering whether the effect of chemotherapy versus no chemotherapy on patient outcomes differs according to the test score.
- Clinical utility – the ability of the prospective use of the test to affect patient outcomes such as recurrence and survival compared with current practice.
- Decision impact – how the test influences decision-making in terms of which patients will be offered chemotherapy.

A total of 153 references were included in the review. Studies assessing prognostic ability and prediction of chemotherapy benefit were quality assessed using relevant criteria selected from the draft prediction model study risk of bias assessment tool (PROBAST). Clinical utility studies were quality assessed using the Cochrane risk of bias tool for randomised controlled trials (RCTs). Studies assessing decision impact were not quality assessed because of time constraints.

Evidence on prognostic ability

Study designs and patient characteristics

Studies with evidence on prognostic ability are summarised by test, starting on page 66 of the diagnostics assessment report. The EAG judged that studies done in East Asia may be less generalisable to England because usual clinical practice may differ between countries enough to affect prognostic outcomes. Also, it is possible that people of different ethnicities have different underlying risk profiles and natural history of disease.

Many studies treated some or all patients with chemotherapy. The EAG stated that results from these studies should be viewed with caution because this could reduce the apparent prognostic performance of a test as chemotherapy could affect event rates. As such, validation cohorts (a population studied to confirm the prognostic ability of a test) should ideally treat patients with endocrine monotherapy, but not chemotherapy.

Results are generally presented as unadjusted or adjusted analyses. Unadjusted analyses do not assess the question of whether a test has additional value over clinicopathological factors. Adjusted analyses show whether the test has prognostic value over clinicopathological variables. Studies with evidence of prognostic ability for the tumour profiling tests are summarised in table 3.

Table 3: Study designs and patient characteristics for studies providing evidence on prognostic ability

	Study designs	Study locations	Cohort characteristics	Treatments	Quality assessment
Oncotype DX (from page 66 of DAR)	11 data sets: 7 re-analyses of RCT data; 4 retrospective studies of routinely collected data/archived samples	RCTs: 5 USA, 1 UK, 1 France Retrospective studies: 1 USA, 2 China, 1 Japan	LN status: 3 studies - mixed; 4 studies - LN0; 3 studies - LN+; 1 study NR HR status: all studies 100% HR+ HER2 status: only 2 studies 100% HER2-	All ET no CT: 4 studies All ET all CT: 3 studies All ET and mixed/ unclear CT: 2 studies Mixed ET and CT: 1 study Unclear: 1 study	All studies were validation studies Only 4 studies with no CT treatment Concerns due to exclusion of tumour samples with insufficient tissue
MammaPrint (from page 128 of the DAR)	10 data sets: 1 re-analysis of RCT data; 9 retrospective studies of routinely collected data 4 studies pooled data on specific patients from studies above	RCT: Sweden Retrospective studies: 4 Netherlands; 2 USA; 2 Europe; 1 Japan	LN status: 3 cohorts – mixed; 6 cohorts – LN0; 1 cohort – LN+ HR status: 8 cohorts – mixed; 2 cohorts (and 1 subgroup) – 100% ER+ HER2 status: 6 cohorts – not reported; 4 cohorts – mixed	All ET no CT: 2 analyses No ET, no CT: 2 analyses Mixed ET, no CT: 1 cohorts Mixed ET, mixed CT: 6 cohorts	All studies were validation studies Only 5 analyses with no CT treatment Concerns due to exclusion of some patients recruited to the original trial/cohort and inclusion of patients out of scope
Prosigna (from page 182 of the DAR)	8 data sets: 6 re-analyses of RCT data; 3 retrospective analyses of 2 prospective cohorts	RCTs: 1 UK, 1 Austria, 1 Spain, 3 USA/Canada Retrospective studies: 1 Denmark, 1 Canada	LN status: Mixed – 6 studies; LN0 – 1 study; LN+ - 2 studies HR status: 6 studies 100% ER+ or HR+; 3 studies mixed HER2 status: 3 studies HER2-; 3 studies HER2 NR; 3 studies mixed	All ET no CT: 5 cohorts All ET and all CT: 1 cohort Some ET (or NR) and all CT: 3 cohorts	All studies were validation studies 5 cohorts had no CT treatment Concerns due to exclusion of some patients recruited to the original trial/cohort
EndoPredict (from page 198 of the DAR)	3 data sets: all re-analyses of RCT data	1 UK, 1 Austria, 1 Spain	LN status: Mixed – 2 studies; LN+ 1 study HR status: all 100% ER+ HER2 status: all 100% HER2-	All ET (5 years) no CT: 2 cohorts All ET (5 years) all CT: 1 cohort	All studies were validation studies 2 studies with no CT Concerns due to exclusion of some patients recruited to the original trial (or unclear)

IHC4 and IHC4+C (from page 212 of the DAR)	12 data sets: 6 re-analyses of RCT data (5 validation, 1 derivation); 6 analyses of routinely collected patient data	RCTs: 1 UK, 2 Germany, 1 Spain, 1 Europe, 1 various Retrospective studies: 2 UK, 1 USA, 1 France, 1 China, 1 Taiwan	LN status: Mixed – 9 studies; LN0 – 1 study; LN+ - 2 studies HR status: 10 studies 100% HR+ or ER+; 2 studies HR+ NR HER2 status: 7 studies 100% HER2-; 3 studies HER2 NR; 2 studies mixed	All ET, no CT: 2 studies All ET and mixed/unclear CT: 4 studies Some ET all CT: 2 studies Some ET no CT: 1 study Mixed/unclear ET and CT: 3	One study was the derivation cohort Only 2 studies with no CT Concerns due to exclusion of tumour samples with insufficient tissue
Abbreviations: CT, chemotherapy; DAR, diagnostics assessment report; ER, oestrogen receptor; ET, endocrine therapy; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; LN, lymph node; NR, not reported; RCT, randomised controlled trial;					

Distribution of patients across risk categories

Among studies of lymph node negative patients treated with endocrine monotherapy, around 70 to 80% of patients were categorised as low or low/intermediate risk across all tests (table 4). Some studies did not report the human epidermal growth factor receptor 2 (HER2) status of patients, which, if mixed, could affect the proportions of patients in each category. There was only 1 MammaPrint study which treated patients with endocrine monotherapy. In MammaPrint studies with mixed endocrine and chemotherapy use, cohorts generally had mixed hormone receptor status, and/or mixed HER2 status, so results may not be comparable with other tests (low risk 20 to 61%; 6 studies; not tabulated). Most IHC4/IHC4+C studies used quartiles or tertiles to define risk groups. These studies will have been specific to each cohort and do not provide useful information on the distribution of patients across risk categories (not tabulated).

Table 4: Risk categories for lymph node negative cohorts not treated with chemotherapy

	Low risk category	Intermediate risk category	High risk category	Number of studies	Patients
Oncotype DX	48 to [redacted]%	20 to [redacted]%	[redacted] to 33%	4 studies	ER+, HER2+/-
MammaPrint	71%	-	29%	1 study	ER+, HER2 NR
Prosigna	48 to [redacted]%	[redacted] to 32%	[redacted] to 20%	3 studies	Most ER+, HER2-
EndoPredict	[redacted] to [redacted]%	-	[redacted] to [redacted]%	2 studies	ER+, HER2-
IHC4+C	[redacted]%	[redacted]%	[redacted]%	1 study	ER+, 95% HER2-
Abbreviations: ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2; NR – not reported					

The proportion of low/intermediate risk patients was generally much lower in lymph node positive than in lymph node negative cohorts. When Oncotype DX was used, however, the proportion of patients with low/intermediate risk was only slightly lower in the lymph node negative group than in the lymph node positive group (table 5). Studies of MammaPrint in lymph node positive patients were all done in cohorts with mixed hormone receptor status and mixed or unknown HER2 status, so results may not be comparable with other tests (low risk 38 to 41%; 2 studies). Most IHC4/IHC4+C studies used quartiles or tertiles to define risk groups, and these will have been specific to each cohort (not tabulated).

Table 5: Risk categories for lymph node positive cohort not treated with chemotherapy

	Low risk category	Intermediate risk category	High risk category	Number of studies	Patients
Oncotype DX	[redacted]%	[redacted]%	[redacted]%	1 study	ER+, HER2-
MammaPrint	-	-	-	No studies limited to HR+ patients	
Prosigna	4 to 25%	27 to 34%	48 to 62%	3 studies	Most ER+, HER2-
EndoPredict	[redacted] to [redacted]%	-	[redacted] to [redacted]%	2 studies	ER+, HER2-
IHC4+C	[redacted]%	[redacted]%	[redacted]%	1 study	ER+, HER2-
Abbreviations: ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; NR – not reported					

Prognostic performance: Oncotype DX

The 10 year distant recurrence-free interval rates (table 6) suggest that:

- the lymph node negative, low-risk group is at very low risk of recurrence in the absence of chemotherapy
- the lymph node negative intermediate risk group maybe at slightly higher risk of recurrence
- The lymph node positive group was generally at higher risk of recurrence than the lymph node negative group in both low and intermediate categories.

Table 6 Percentage DRFI risk (0 to 10 years) by lymph node status and Oncotype DX test result

LN status	Oncotype DX risk score	10 year DRFI rates [a]	Notes
LN negative	Low risk	93 to 97%	4 studies; patients had endocrine monotherapy
		96%	1 study; patients had endocrine and chemotherapy
LN positive	Low risk	82%	1 study; patients had endocrine monotherapy
		81%	1 study; patients had endocrine and chemotherapy
LN negative	Intermediate risk	86 to 100%	4 studies; patients had endocrine monotherapy
		89%	1 study; patients had endocrine and chemotherapy
LN positive	Intermediate risk	[redacted]	1 study; patients had endocrine monotherapy
		65%	1 study; patients had endocrine and chemotherapy
LN negative	High risk	61 to 77%	4 studies; patients had endocrine monotherapy
		88%	1 study; patients had endocrine and chemotherapy
LN positive	High risk	[redacted]	1 study; patients had endocrine monotherapy
		59%	1 study; patients had endocrine and chemotherapy

Abbreviations: DRFI, distant recurrence-free interval; LN, lymph node

[a] Note than Sun et al. was excluded from the ranges because it appeared to be an outlier with very low DRFI rates (18 to 63%)

Unadjusted analyses indicated that Oncotype DX had prognostic power (there were statistically significant differences between low-risk and high-risk groups) across various recurrence outcomes, regardless of lymph node status. However, hazard ratios between the intermediate-risk group and the high- or low-risk groups were not always statistically significant, particularly in the lymph node positive group.

In adjusted analyses, Oncotype DX provided additional prognostic information over most commonly used clinicopathological variables (age, grade, size, nodal status) regardless of lymph node status. Analyses also showed that Oncotype DX provided additional prognostic information (statistically significant) [redacted].

Prognostic performance: MammaPrint

MammaPrint had prognostic power (there were statistically significant differences between low-risk and high-risk groups) for 10 year distant recurrence-free survival in almost all unadjusted analyses of lymph node negative and lymph node positive patients. The 10-year distant recurrence-free survival and distant recurrence-free interval rates for low-risk patients are shown in table 7.

Table 7 Percentage DRFS/DRFI risk (0 to 10 years) by lymph node status and MammaPrint risk category

LN status	MammaPrint risk category	10-year DRFS/DRFI rates	Notes
Pooled LN negative / LN positive	Low risk group	87%	1 analysis; 33% had endocrine and 25% chemotherapy
LN negative	Low risk group	93%	1 analysis; endocrine monotherapy
		83%	1 analysis; no endocrine or chemotherapy
LN negative	Low risk group	80% to 90%	3 analyses; varying rates of endocrine and chemotherapy use
LN positive	Low risk group	79% to 91%	2 analyses; varying rates of endocrine and chemotherapy use
Abbreviations: DRFS, distant recurrence-free survival; DRFI, distant recurrence-free interval; LN, lymph node			

In lymph node negative patients, 4 of 5 unadjusted analyses showed statistically significant differences in hazard ratios between risk groups for 10-year distant recurrence-free survival or distant recurrence-free interval rates. Among lymph node positive patients, 2 unadjusted analyses reported statistically significant prognostic performance of MammaPrint based on hazard ratios for 10-year distant recurrence-free survival between risk groups.

In adjusted analyses, a pooled analysis of lymph node negative and positive patients showed that MammaPrint had prognostic power (there were statistically significant differences between low-risk and high-risk groups) for 10-year distant recurrence-free survival in a multivariable analysis adjusting for clinicopathological variables. Among lymph node negative patients, MammaPrint had prognostic power for distant recurrence-free interval when adjusted for Adjuvant! Online or NPI in 3 cohorts. In lymph node positive patients, MammaPrint had prognostic power (statistically significant

differences or borderline statistically significant differences between low-risk and high-risk groups) for 10-year distant recurrence-free survival and interval in adjusted analyses.

Prognostic performance: Prosigna

Prosigna had prognostic power (there were statistically significant differences between low-risk and high-risk groups) for 10-year distant recurrence-free survival and 10-year distant recurrence-free interval in all unadjusted analyses of lymph node negative and lymph node positive patients. The 10-year distant recurrence-free survival and distant recurrence-free interval rates for low- and intermediate-risk patients are shown in table 8. In analyses adjusted for clinicopathological variables, Prosigna was found to be prognostic for 10-year distant metastasis-free survival and distant recurrence-free survival. In lymph node negative patients the results were statistically significant and in lymph node positive patients the results were statistically or borderline significant.

Table 8 Percentage DRFS/DRFI risk (0 to 10 years) by lymph node status and Prosigna category

LN status	Prosigna risk category	10 year DRFS/DRFI rates	Notes
LN negative	Low risk	95% to [redacted]	3 studies; patients had endocrine monotherapy
LN positive	Low risk	[redacted]	2 studies; patients had endocrine monotherapy
		92%	1 study; all patients had endocrine and chemotherapy
LN negative	Intermediate risk	[redacted] to 93%	2 studies; patients had endocrine monotherapy
LN positive	Intermediate risk	[redacted] to 94%	2 studies; patients had endocrine monotherapy
		74%	1 study; all patients had endocrine and chemotherapy
Abbreviations: DRFS, distant recurrence-free survival; DRFI, distant recurrence-free interval; LN, lymph node			

Prognostic performance: EndoPredict

EndoPredict had prognostic power (there were statistically significant differences between low-risk and high-risk groups) for unadjusted analyses of 10-year distant recurrence-free survival and distant recurrence-free interval in lymph node negative and lymph node positive patients. The 10-year distant recurrence-free survival and distant recurrence-free interval rates for low-risk

patients are shown in table 9. In adjusted analyses of one study, EndoPredict [redacted].

Table 9 Percentage DRFS/DRFI risk (0 to 10 years) by lymph node status and EndoPredict risk category

LN status	EndoPredict risk category	10 year DRFS/DRFI rates	Notes
LN negative	Low risk	[redacted] to [redacted]	2 studies, patients had endocrine monotherapy
LN positive	Low risk	[redacted]	2 studies, patients had endocrine monotherapy
LN positive	Low risk	100%	1 study, patients had endocrine and chemotherapy
Abbreviations: DRFS, distant recurrence-free survival; DRFI, distant recurrence-free interval; LN, lymph node			

Prognostic performance: IHC4 and IHC4+C

Studies have shown that IHC4 provides statistically significant prognostic information in unadjusted analyses in both lymph node negative, lymph node positive and mixed cohorts. However, none of these studies reported survival or recurrence outcomes by risk group, but instead presented hazard ratios for high risk groups compared with low risk groups. In addition, most studies used quartiles or tertiles to define risk groups, and these were specific to each cohort and did not use pre-defined cut-off values. Also, many used laboratory methods that differed from the derivation study methodology (the study population in which the test was established). In adjusted analyses, IHC4 was shown to have additional prognostic value over clinicopathological factors in some studies.

Data on IHC4+C came from the derivation cohort and 1 validation cohort. These studies showed that IHC4+C had prognostic value in unadjusted analyses. In adjusted analyses, IHC4+C provided statistically significantly more information than NPI in lymph node negative patients but not lymph node positive patients.

Evidence on prediction of chemotherapy benefit

Chemotherapy benefit: Oncotype DX

Five data sets reported across 11 published references and 1 confidential manuscript conducted analyses that assessed the ability of Oncotype DX to predict benefit of chemotherapy (table 10).

Table 10 Studies reporting the ability of Oncotype DX to predict benefit of chemotherapy

Data set	Reference(s)	Study design	Patient population	Study quality
SWOG-8814 study	Albain et al. 2010	Phase 3, open-label, parallel-group RCT	All postmenopausal, HR+, LN+ patients (38.1% with ≥4 positive lymph nodes) and 12% HER2+.	Some risk of bias, mainly because of patient spectrum bias
NSABP B-20 trial	Paik et al. 2006, Tang et al. 2011a and Tang et al. 2011b	RCT	ER+, LN0 patients, with an unreported percentage being HER2	Some risk of bias, mainly because of patient spectrum bias
MD Anderson Center		Retrospective observational	HR+, HER2-, LN0 patients	High risk of confounding and unclear generalisability to decision problem
Clalit Health Services		Retrospective observational	ER+, HER2- patients, [redacted] and with LN1-3	High risk of confounding and unclear generalisability to decision problem
SEER registry		Retrospective observational	HR+, HER2-, LN0 patients	High risk of confounding and unclear generalisability to decision problem

Abbreviations: ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; LN, lymph node; NSABP, national surgical adjuvant breast and bowel project; RCT, randomised controlled trial; SEER, surveillance, epidemiology and end results; SWOG, southwest oncology group

There is some evidence from 2 re-analyses of RCTs to suggest that Oncotype DX may predict benefit from chemotherapy. Based on hazard ratios for disease-free survival for patients having chemotherapy versus those having no chemotherapy, the greatest benefit appeared to be for patients in the Oncotype DX high risk recurrence score category. Unadjusted interaction tests between Oncotype DX risk group and chemotherapy benefit were mainly statistically significant, but adjusted interaction tests were not always statistically significant. Therefore the significant results could be a consequence of omitting potentially important covariates from the statistical model.

From the 3 observational studies evidence was mixed and at high risk from confounding; 1 reported a statistically significant interaction test but this was only adjusted for a limited number of factors; 2 reported hazard ratios for various outcomes at 5 years which were statistically non-significant.

The recurrence score pathology-clinical (RSPC) algorithm incorporates Oncotype DX plus age, tumour size and grade. The RSPC algorithm was derived in TransATAC and NSABP B-14, and validated in NSABP B-20. There was a non-significant interaction test result between chemotherapy benefit and RSPC risk group. This suggests that the interaction between treatment effect and recurrence score risk group may be confounded by clinicopathological variables.

Chemotherapy benefit: MammaPrint

Two studies reported the ability of MammaPrint to predict the benefit of chemotherapy (table 11). Knauer et al. (2010) reported a pooled analysis of 541 patients, from 6 consecutive patient series, and Mook et al. (2009) reported a pooled analysis of 2 of the 6 patient series from Knauer et al. with an extended follow-up (10 years).

Table 11 Studies reporting the ability of MammaPrint to predict chemotherapy benefit

Reference(s)	Study design	Patient population	Study quality
Knauer et al. 2010	Retrospective analysis of 6 patient series	90% ER+, 89% HER2-, 50% LN0, 50% 1-3 LN+ 100% had endocrine therapy, 42% had chemotherapy	High risk of confounding and included a proportion of patients who were ER- and HER2+
Mook et al. 2009	Retrospective analysis of 2 of the 6 patient series from Knauer et al.	Restricted to LN1-3 patients (including micrometastases)	High risk of confounding and included a proportion of patients who were ER- and HER2+
Abbreviations: ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2; LN, lymph node			

The evidence for the ability of MammaPrint to predict chemotherapy benefit is very limited. The effect of chemotherapy versus no chemotherapy was statistically significant in the MammaPrint high-risk group but not in the low-risk group in unadjusted analyses. In analyses adjusted for clinicopathological

variables results were not statistically significant. Further, the interaction test for chemotherapy treatment and risk group was non-significant. In the analysis restricted to LN1-3 patients, a statistically non-significant interaction between chemotherapy treatment and risk group was reported.

Evidence on clinical utility

Clinical utility: Oncotype DX

Five data sets reported across 9 published references and 1 confidential manuscript reported evidence relating to the clinical utility of Oncotype DX (table 12). Another study, based on the SEER database (National Cancer Institute database, US) and Genomic Health's clinical laboratory database, did not meet the inclusion criteria (because of insufficient follow-up length), but is discussed because it presents subgroup data according to age, lymph node status and race. Studies generally reported different outcomes, making comparisons across studies difficult.

Table 12 Studies providing evidence on the clinical utility of Oncotype DX

Study	Study design	Patients
TAILORx	Women with RS<11 were assigned to endocrine therapy alone. Women with RS 11 to 25 were randomised to either endocrine therapy plus chemotherapy or endocrine therapy alone. As of July 2017, only results for the low-risk (RS<11) group (n=1,626) were available. Data for this group are effectively prospective observational data.	HR+, HER2-, LN0 with tumours sized 1.1 to 5 cm (or 0.6 to 1.0 cm in intermediate or high-risk)
WSG Plan B trial	Patients with RS≥12 were randomised to 2 different sorts of chemotherapy. Another aim was to assess the risk of recurrence in patients with RS<12 who were not treated with adjuvant chemotherapy. Data for this group are effectively prospective observational data.	HR+, HER2-, clinically high-risk patients with 0-3 positive LN
MD Anderson Cancer Center (USA)	Retrospective analyses of routinely collected data. Treatment was given according to routine clinical practice, including the Oncotype DX RS, which resulted in differing levels of chemotherapy being prescribed per risk group and per study.	ER+, HER2-, LN0-LNmic patients who had had an Oncotype DX test
Clalit Health Services (Israel)	Retrospective analyses of routinely collected data. Treatment was given according to routine clinical practice, including the Oncotype DX RS, which resulted in differing levels of chemotherapy being prescribed per risk group and per study.	ER+, HER2-, LN0-LNmic or LNmic – LN3 patients who had had an Oncotype DX test.
Memorial Sloan Kettering Center (USA)	Retrospective analyses of routinely collected data. Treatment was given according to routine clinical practice, including the Oncotype DX RS, which resulted in differing levels of chemotherapy being prescribed per risk group and per study.	ER+, HER2-, LN0-LNmic patients who had had an Oncotype DX test.
Abbreviations: ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; LN, lymph node; LNmic, lymph node micrometastases; RS, risk score; TAILORx, the trial assigning individualised options for treatment WSG, West German study group		

Based on the evidence available, the EAG stated that it is difficult to conclude whether patient outcomes would be affected by using the test in a clinical setting. In lymph node negative patients, using the test in clinical practice appeared to result in low rates of chemotherapy use in low-risk patients (2% to 12%), with acceptable outcomes (distant recurrence-free survival, distant recurrence-free interval or invasive disease-free survival 96% to 99.6%). Rates of chemotherapy use increased with increasing risk category, and were generally higher in lymph node positive patients.

It was not possible to conclude whether patients in intermediate and high-risk categories had better outcomes as a result of using Oncotype DX to guide treatment as there were no comparator (no-Oncotype DX) groups.

Clinical utility: MammaPrint

Two studies reported evidence relating to clinical utility of MammaPrint (table 13).

Table 13 Studies providing evidence on the clinical utility of MammaPrint

Study	Study design	Patients
MINDACT	A partially randomised trial of MammaPrint versus clinical practice. Patients with discordant risk scores (high/low or low/high) via MammaPrint and modified AO were randomised to chemotherapy or no chemotherapy; patients with concordant risk were followed as prospective cohorts (high-risk patients were all recommended to receive chemotherapy and low-risk patients were all recommended no chemotherapy).	Overall, 88% were HR+; 90% HER2-; 79% were LN0 and 21% LN1-3. However, this varied by group.
RASTER	A prospective observational study. Chemotherapy use was guided by MammaPrint in combination with the Dutch Institute of Healthcare Improvement guidelines of 2004 and clinician and patient preference. As such, estimates of prognostic performance are confounded by the differing rates of chemotherapy in different risk groups. [redacted]	LN0 patients, age <61 years, 80% ER+ and 84% HER2-
Abbreviations: AO, Adjuvant! Online; ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; LN, lymph node		

In the MINDACT study, for the high-clinical, low-MammaPrint risk group, 5 year distant metastasis-free survival was 95.9% with chemotherapy and 94.4% without chemotherapy, an absolute difference of 1.5%. The authors suggested that chemotherapy could possibly be avoided in these patients. For

the low-clinical, high-MammaPrint risk group, 5-year distant metastasis-free survival was 95.8% with chemotherapy and 95.0% without chemotherapy, an absolute difference of 0.8%. The EAG suggested that this result shows that low-clinical risk patients with a high-risk MammaPrint result have little benefit from chemotherapy, implying that MammaPrint should not be used to guide treatment in low clinical risk patients as it would result in patients receiving chemotherapy but not gaining any benefit. However, the comparator was modified Adjuvant! Online, and it is unclear whether the same would be true for other clinical risk scores.

Results from the RASTER study suggest that distant recurrence-free interval rates were sufficiently low in the MammaPrint low-risk group for these patients to avoid chemotherapy. The 5-year distant recurrence free interval rate for lymph node negative patients was 97.0% for low-risk patients (15% had chemotherapy) and 91.7% for high-risk patients (81% received chemotherapy). [redacted] Further, MammaPrint provided additional prognostic information over Adjuvant! Online and NPI, but not over another NHS risk scoring tool, PREDICT Plus.

Evidence comparing tests with each other

Studies comparing more than one test

Data were reported for 6 cohorts: 4 re-analyses of RCTs and 2 observational studies. The design and results from these studies are described in more detail in the diagnostics assessment report starting on page 241. The most comprehensive analysis in terms of the number of tests assessed was from TransATAC, which assessed 4 tests: EndoPredict, Prosigna, Oncotype DX and IHC4+C. A bespoke analysis of the TransATAC data which focused on the population in the scope for this assessment was provided by the trial investigators.

As the data comparing the tests with each other are limited, only broad conclusions can be drawn. Evidence shows that generally when a test placed

more patients in a low-risk category compared with another test, the event-free survival was reduced. Also, the tests generally performed differently in lymph node negative and lymph node positive patients. In TransATAC, [redacted] In terms of additional prognostic information the tests provide over clinicopathological variables or algorithms, [redacted] provided statistically significantly more prognostic information [redacted]

Microarray studies

Thirteen studies reported data from microarray analyses on more than one of the tests. These studies had methodological limitations, but they have value because they provide comparative prognostic data. All the studies reported data on Oncotype DX and MammaPrint, and 2 also report data on EndoPredict. The results of these studies are described in the diagnostics assessment repot starting on page 263.

The microarray studies support conclusions from studies using the commercial versions of the assays in suggesting that Oncotype DX, MammaPrint and EndoPredict can discriminate between high- and low-risk patients regardless of lymph node status. In terms of additional prognostic performance of the tests over clinicopathological variables, EndoPredict appeared to have the greatest benefit, followed by Oncotype DX and then MammaPrint, though the evidence base was limited.

OPTIMA prelim

The OPTIMA prelim study analysed concordance between different tests, that is, the degree to which the tests assign the same patients to the same risk groups. It is described starting on page 282 of the diagnostics assessment

report. The study included Oncotype DX, MammaPrint, Prosigna and IHC4 plus 2 other tests not included in this assessment. OPTIMA prelim selected women who would routinely be offered chemotherapy, specifically women aged 40 years or older with ER+, HER2- early breast cancer with either 1-9 positive lymph nodes, or a tumour of 30 mm or greater if node negative.

Out of the 4 in-scope tests, MammaPrint assigned the most patients to the low-risk category (table 14), but unlike the other 3 tests it does not have an intermediate category. When low and intermediate categories were treated as 1 category for the 3 tests that have 3 risk groups, Oncotype DX assigned the most to the low/intermediate category, and MammaPrint the least. Kappa statistics indicated modest agreement between tests ranging from 0.33 to 0.53. Further, across 5 tests in the study that reported risk groups, only 39% of tumours were uniformly classified as either low/intermediate or high by all 5 tests. Of these, 31% were low/intermediate by all tests and 8% were high-risk by all tests.

The authors of the study concluded that although the tests assigned similar proportions of patients to low/intermediate and high risk categories, test results for an individual patient could differ markedly depending on which test was used.

Table 14 Proportion of patients assigned to each risk category in OPTIMA prelim

Test	% Low risk	% Intermediate risk	% High risk
Oncotype DX	54	28	18
MammaPrint	61	-	39
Prosigna	36	29	35
IHC4	24	48	28

Evidence on decision impact

The review of decision impact focused on studies done in the UK or the rest of Europe. The studies are described starting on page 284 of the diagnostics assessment report. The following studies were identified:

- Oncotype DX: 6 UK studies and 12 other European studies

- EndoPredict: 1 UK study and 3 other European studies
- IHC4+C: 1 UK study and 0 other European studies
- Prosigna: 0 UK studies and 3 other European studies
- MammaPrint: 0 UK studies and 8 other European studies

The percentage of patients with any change in treatment recommendation or decision (either to or from chemotherapy) in UK studies was 29% to 49% across 4 Oncotype DX studies, 37% in 1 EndoPredict study and 27% in 1 IHC4+C study. Ranges across European (non-UK) studies were 5% to 70% for Oncotype DX, 38% to 41% for EndoPredict, 14% to 41% for Prosigna and 13% to 51% for MammaPrint.

The net change in the percentage of patients with a chemotherapy recommendation or decision (pre-test to post-test) among UK studies was a reduction of 8% to 23% across 4 Oncotype DX studies, an increase of 1% in 1 EndoPredict study, and a reduction of between 2% and 26% in 1 IHC4+C study (depending on whether the decision was defined as 'recommend chemotherapy' or 'discuss chemotherapy'). Net changes across European (non-UK) studies were a reduction of 0% to 64% for Oncotype DX, a reduction of 13% to 26% for EndoPredict, a reduction of 2% to an increase of 9% for Prosigna, and reduction of 31% to an increase of 8% for MammaPrint.

Evidence on anxiety and health-related quality of life

The EAG identified 6 studies that reported outcomes relating to anxiety (including worry and distress) and health-related quality of life, which are discussed starting on page 298 of the diagnostics assessment report. The lack of use of a comparator in the studies made it difficult to tell whether changes in anxiety experienced with the use of tumour profiling tests would also have occurred if patients received a definitive decision based on clinical risk factors alone. Overall, evidence suggests that tumour profile testing may reduce state anxiety in some patients in some contexts, but generally there was little impact on health-related quality of life.

2.2 Costs and cost effectiveness

The EAG conducted a review of existing studies investigating the cost effectiveness of tumour profiling tests to guide treatment decisions in people with early breast cancer, and critiqued economic analyses provided by Agendia (MammaPrint), Genomic Health (Oncotype DX) and, and the chief investigator of a UK decision impact study (EndoPredict). The EAG also constructed a de novo economic model to assess the cost effectiveness of Oncotype DX, MammaPrint, Prosigna, IHC4+C, and EndoPredict compared with current practice.

Systematic review of cost-effectiveness evidence

A total of 26 studies were identified that had been published since the original assessment for diagnostics guidance 10. The models reported in the studies were developed to assess the cost effectiveness of tumour profiling tests across a variety of different countries including the UK, the US, Canada, Mexico, Japan, Austria, Germany, France and the Netherlands. Most studies compared Oncotype DX (18 studies) or MammaPrint (8 studies) with comparators such as Adjuvant! Online, the St Gallen guidelines, standard practice or other conventional diagnostic tools. Only 1 study compared EndoPredict against a comparator, which was comprised of a combination of 3 different guidelines. There was variation between the analyses in the patient populations evaluated, their disease type and other patient characteristics.

There was a high level of consistency in terms of the general modelling approach and structure, and several studies were based on a previously published model. The majority of the models used a Markov or hybrid decision tree–Markov approach, 2 studies used a partitioned survival approach, and 1 study used a discrete event simulation approach. The time horizons ranged from 10 years to the patient's remaining lifetime, with cycle lengths ranging from 1 month to 1 year when reported. Most of the models that evaluated Oncotype DX assumed that the test could predict the benefit of chemotherapy.

None of the identified models included all of the intervention tests included in the scope of the assessment. Further details of the published models can be found starting on page 305 of the diagnostics assessment report.

Review and critique of economic analyses provided by companies

Economic analyses were provided by the manufacturers of Oncotype DX (Genomic Health) and MammaPrint (Agendia) and the chief investigator of the EndoPredict (Myriad) decision impact study. The EAG review and critique of these models is presented in the diagnostics assessment report starting on page 317.

Economic analysis

The EAG developed a de novo economic model designed to assess the cost effectiveness of Oncotype DX, MammaPrint, Prosigna, IHC4+C and EndoPredict compared with current practice. It is described in the diagnostics assessment report starting on page 346. The model assessed the health outcomes and costs associated with each test over a lifetime horizon (42 years) from the perspective of the UK NHS and Personal Social Services. All costs and health outcomes were discounted at a rate of 3.5% per annum. Unit costs were valued at 2015/16 prices. The principal source of evidence used to inform the analyses of Oncotype DX, Prosigna, IHC4+C and EndoPredict was the TransATAC study. ATAC was an international trial with a translational research continuation (TransATAC) that investigated the prognosis of breast cancer. Only UK data were included in the bespoke analysis provided to the EAG, which was also restricted to hormone receptor positive, HER2 negative patients with 0 to 3 positive lymph nodes. A comparison of the TransATAC data with other study data is provided in the diagnostics assessment report on page 303. As this study excluded MammaPrint, the MINDACT study was used as the basis for evaluating the cost effectiveness of MammaPrint. PREDICT scores were not available in either dataset, and so this tool could not be considered as a comparator or used to determine different risk subgroups. Therefore, the comparator for the analyses of Oncotype DX, Prosigna, IHC4+C and EndoPredict was current practice (various tools and algorithms),

and the comparator for the analysis of MammaPrint was a modified version of Adjuvant! Online.

Model structure

The de novo model was a hybrid decision tree–Markov model, and was based on the model previously developed by Ward et al. to inform diagnostics guidance 10. The decision tree component of the model classified patients in the current practice (no test) group and the tumour profiling test group into high, intermediate and low risk categories. For EndoPredict and MammaPrint, the intermediate risk category was excluded as the test provides results in terms of high and low risk only. Within both the test group and the current practice group, the decision tree determined the probability that a patient would be in 1 of 6 groups: low-risk, chemotherapy; low risk, no chemotherapy; intermediate risk, chemotherapy; intermediate risk, no chemotherapy; high risk, chemotherapy, and high risk, no chemotherapy (for the analyses of EndoPredict and MammaPrint, 4 branches were used due to the absence of an intermediate risk category). Each of the branches was then linked to a Markov model which predicted lifetime quality-adjusted life-years (QALYs) and costs according to the patient's risk of distant recurrence and whether or not they received chemotherapy. The decision tree structure is shown in Figure 1.

Figure 1: Decision tree

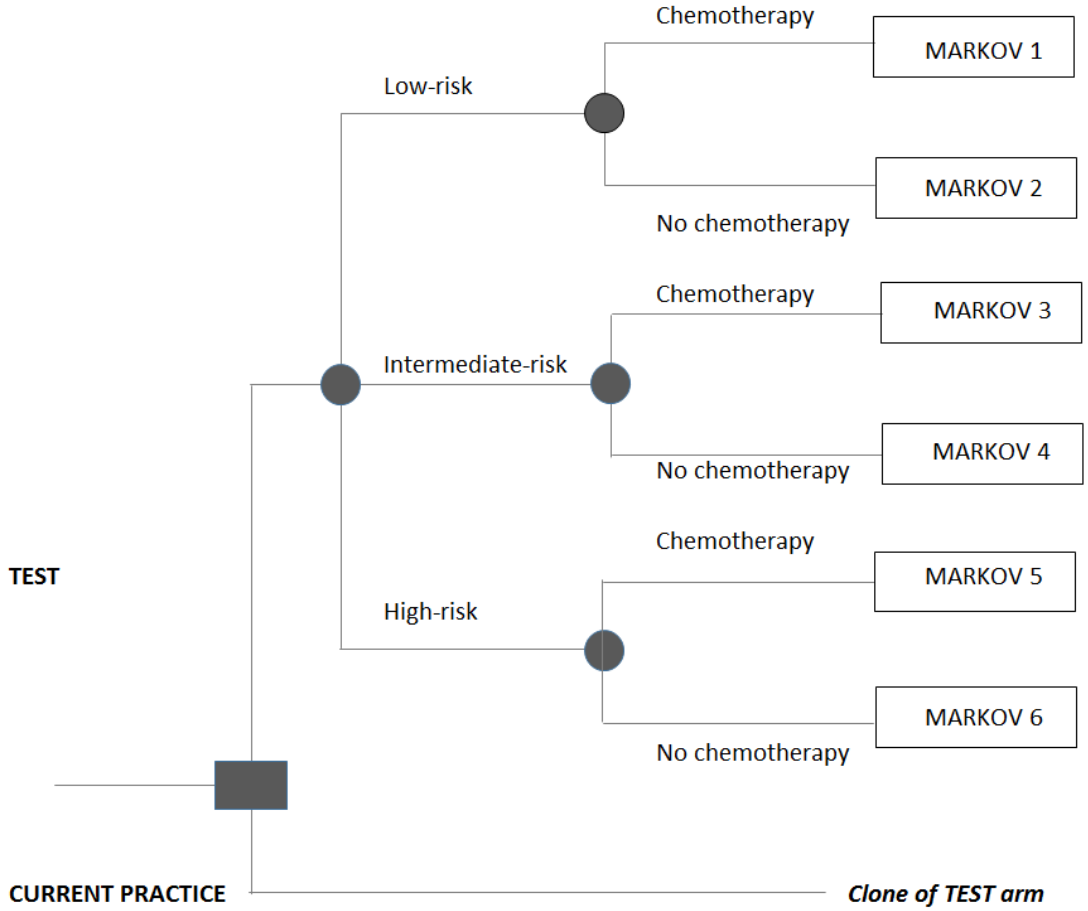
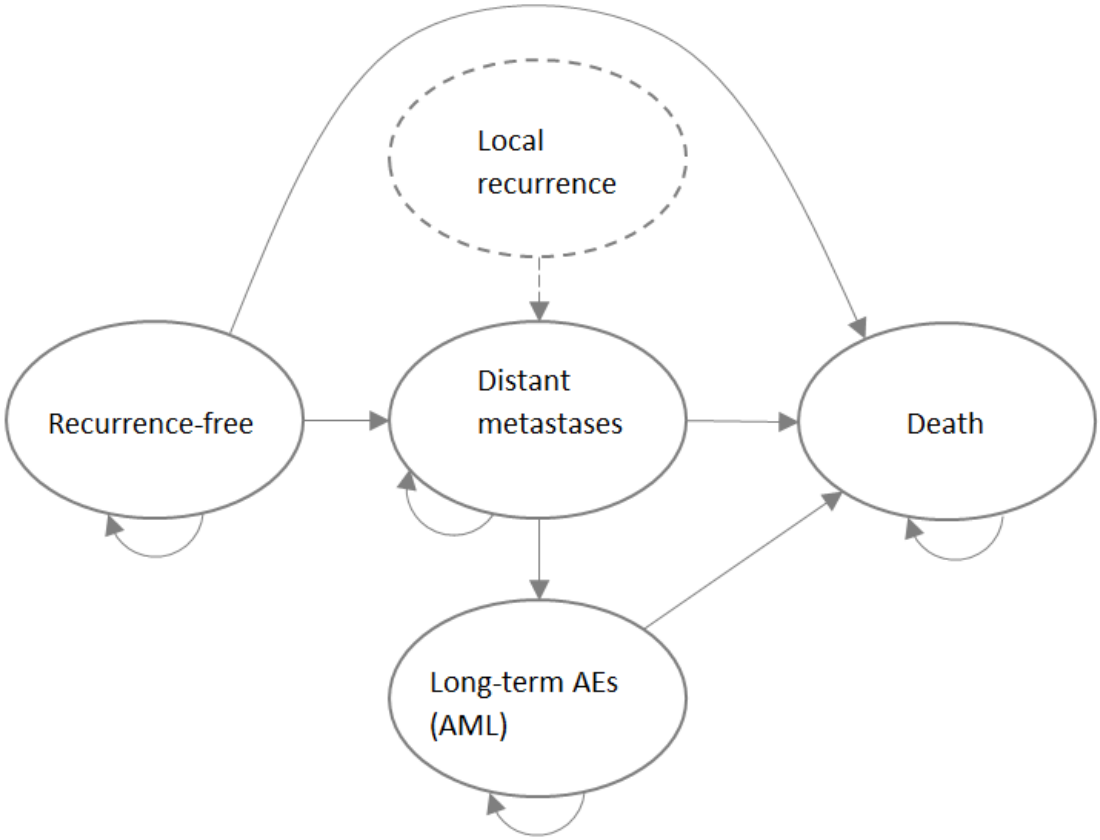


Figure 2 shows the Markov nodes of the model. Each Markov node was evaluated over 84 cycles of 6 months (42 years). Each node included 4 health states: recurrence-free; distant recurrence; long-term adverse events (acute myeloid leukaemia); and dead. Patients entered the model in the recurrence-free health state. A health-related quality of life decrement was applied during the first model cycle to account for health losses associated with short-term adverse events for patients receiving adjuvant chemotherapy. The benefit of adjuvant chemotherapy was modelled using a relative risk reduction for distant recurrence within each risk classification group. The benefit of the test was therefore captured in the model by changing the probability that patients with each test risk classification received adjuvant chemotherapy. A full

description of the model structure can be found starting on page 348 of the diagnostics assessment report.

Figure 2: Markov nodes



Abbreviations: AEs, adverse events; AML, Acute myeloid leukaemia

Model inputs

The model was populated with data from the clinical-effectiveness review, the NHS England access scheme dataset (provided as commercial-in-confidence by Genomic Health), the National Cancer Registration and Analysis Service (NCRAS), published literature and expert opinion. Full details of the model inputs can be found starting on page 352 of the diagnostics assessment report.

Risk classification probabilities

Risk classification probabilities for Oncotype DX, Prosigna, IHC4+C and EndoPredict were obtained from a bespoke data analysis of the TransATAC

trial (table 15). The analysis provided data on hormone receptor positive, HER2 negative patients for the 3 modelled subgroups (node-negative NPI≤3.4, node-negative NPI>3.4, and 1 to 3 positive nodes). Risk classification probabilities for MammaPrint were obtained from the MINDACT trial (table 16).

Table 15 Risk classification probabilities from TransATAC

Test (number samples)	Proportion of patients with risk classification		
	Low-risk	Intermediate-risk	High-risk
Node-negative, NPI≤3.4			
Oncotype DX (541)	0.72	0.24	0.04
Prosigna (410)	0.72	0.24	0.03
IHC4+C (510)	0.88	0.11	0.01
EndoPredict (423)	0.90	-	0.10
Node-negative, NPI>3.4			
Oncotype DX (284)	0.50	0.31	0.19
Prosigna (253)	0.27	0.38	0.35
IHC4+C (279)	0.36	0.38	0.25
EndoPredict (254)	0.47	-	0.53
Node-positive (1-3 nodes)			
Oncotype DX (219)	0.57	0.32	0.11
Prosigna (192)	0.08	0.32	0.60
IHC4+C (213)	0.28	0.34	0.38
EndoPredict (198)	0.24	-	0.76
Abbreviations: NPI, Nottingham prognostic index			

Table 16 Risk classification probabilities from MINDACT

Population	Proportion of patients with risk classification	
	MammaPrint low-risk	MammaPrint high-risk
MINDACT overall population (n=6,693)	0.64	0.36
MINDACT Adjuvant! Online clinical high-risk subgroup (n=3,370)	0.46	0.54
MINDACT Adjuvant! Online clinical high-risk subgroup (n=3,324)	0.82	0.18

Probability of developing distant metastases

For Oncotype DX, Prosigna, IHC4+C and EndoPredict the probability of developing distant metastases in each cohort and risk category was based on 10-year recurrence-free interval data taken from the bespoke data analysis of the TransATAC trial (table 17). For MammaPrint the probability of developing distant metastases was based on an adjusted analysis of 5-year distant metastases-free survival data from Cardoso et al. 2016 (table 18). The model

assumed that the risk of distant metastases between 10 and 15 years was halved, and after 15 years was zero. This assumption was made because there is uncertainty about the sustained benefit of adjuvant chemotherapy.

Table 17 Ten year distant recurrence rates by risk classification

Population	Proportion distant recurrence-free at 10 years (95% CI)			
	Oncotype DX	Prosigna	IHC4+C	EndoPredict
LN negative, NPI≤3.4, low risk	0.983 (0.963-0.992)	0.986 (0.962-0.995)	0.975 (0.954-0.987)	0.971 (0.947-0.984)
LN negative, NPI≤3.4, intermediate risk	0.931 (0.867-0.965)	0.933 (0.857-0.969)	0.878 (0.747-0.943)	n/a
LN negative, NPI≤3.4, high risk	0.838 (0.577-0.945)	0.636 (0.297-0.845)	0.800 (0.204-0.969)	0.870 (0.714-0.944)
LN negative, NPI>3.4, low risk	0.854 (0.776-0.907)	0.923 (0.825-0.967)	0.873 (0.787-0.926)	0.848 (0.761-0.905)
LN negative, NPI>3.4, intermediate risk	0.798 (0.694-0.869)	0.796 (0.687-0.870)	0.788 (0.688-0.859)	n/a
LN negative, NPI>3.4, high risk	0.749 (0.598-0.851)	0.699 (0.584-0.788)	0.769 (0.645-0.855)	0.774 (0.688-0.838)
LN positive, low risk	0.818 (0.727-0.880)	1 (n/a)	0.961 (0.851-0.990)	0.95 (0.811-0.988)
LN positive, intermediate risk	0.754 (0.630-0.842)	0.807 (0.679-0.889)	0.758 (0.635-0.845)	n/a
LN positive, high risk	0.686 (0.447-0.839)	0.707 (0.604-0.788)	0.672 (0.546-0.771)	0.716 (0.629-0.785)
Abbreviations: LN, lymph node; NPI, Nottingham prognostic index				

Table 18 Calculated 10-year distant metastases-free survival by group for MammaPrint

Risk group	Treatment	10-year distant metastases-free survival probability for group[a]
Modified Adjuvant! Online low, MammaPrint low	Chemotherapy	0.953
	No chemotherapy	0.953
Modified Adjuvant! Online high, MammaPrint low	Chemotherapy	0.920
	No chemotherapy	0.891
Modified Adjuvant! Online low, MammaPrint high	Chemotherapy	0.918
	No chemotherapy	0.903
Modified Adjuvant! Online high, MammaPrint high	Chemotherapy	0.821
	No chemotherapy	0.766[b]
[a] Extrapolated from 5-year data assuming a constant event rate [b] Estimated by adjusting to remove the effect of chemotherapy		

Probability of having chemotherapy

The probability of having chemotherapy in the current practice group was taken either from data resulting from a bespoke request placed with NCRAS to obtain aggregate data on the use of adjuvant chemotherapy in women with early breast cancer in England (described on page 359 of the diagnostics assessment report), or from the NHS England Access dataset (relating only to node-negative disease and NPI>3.4; described on page 360 of the diagnostics assessment report). The NCRAS dataset reflects unselected patients who were not necessarily eligible for tumour profile testing; therefore the proportion of women who are eligible for testing who go on to receive chemotherapy may be greater than the estimates generated using this dataset.

In the groups for which tumour profiling tests were used, the probability of having chemotherapy was taken from either:

- The NHS England access scheme dataset: [redacted]
- Bloomfield et al. (2017): A UK study on the impact of EndoPredict results on adjuvant treatment decisions (149 patients). This is the only decision impact study on a 2-level tumour profiling test. It is unlikely to accurately represent the use of chemotherapy in node-positive disease.
- Loncaster et al. (2017): A prospective study to evaluate the clinical value of Oncotype DX testing in 201 patients who had been recommended chemotherapy. It provides separate estimates for node-negative and node-positive disease.
- Holt et al. (2011): A UK study on the impact of Oncotype DX on adjuvant treatment decisions with results available for 74 patients.
- UK breast cancer group (UKBCG) survey: a bespoke survey designed by the EAG, with 11 usable responses from oncologists. The results indicate considerable variation in practice.
- Expert opinion.

The probability of receiving adjuvant chemotherapy in the current practice group by test risk classification is presented in table 19. Where appropriate, the source not selected for inclusion in the base case was tested in the sensitivity analyses.

Table 19 Probability of receiving chemotherapy in the base case

Population	Source	Proportion of patients receiving chemotherapy		
		Low risk	Intermediate risk	High risk
Current practice group				
Node-negative, NPI≤3.4	NCRAS dataset	0.07		
Node-negative, NPI>3.4	NHS England access scheme dataset	0.43		
Node-positive (1-3 nodes)	NCRAS dataset	0.63		
Overall population (MammaPrint)	Expert opinion	0.46		
3-level tests (Oncotype DX, Prosigna and IHC4+C)				
Node negative, NPI≤3.4	UKBCG survey data	0.00	0.17	0.74
Node negative, NPI>3.4	NHS England access scheme dataset	0.01	0.33	0.89
Node-positive (1-3 nodes)	Loncaster et al. node-positive estimates	0.08	0.63	0.83
2-level tests (EndoPredict and MammaPrint)				
EndoPredict: All 3 subgroups	Bloomfield et al. (2017) study	0.07	-	0.77
MammaPrint: all subgroups	Bloomfield et al. (2017) study	0.07	-	0.77
Abbreviations: NCRAS, national cancer registration and analysis service; NPI, Nottingham prognostic index; UKBCG, UK breast cancer group				

Adjuvant chemotherapy treatment effect on distant recurrence

In the base-case analysis, the benefit of chemotherapy was assumed to be the same across all test risk groups, that is, all tests were assumed to be associated with prognostic benefit only. For Oncotype DX, Prosigna, IHC4+C and EndoPredict the relative risk of recurrence for chemotherapy versus no chemotherapy was based on a meta-analysis reported by the early breast cancer trialists' collaborative group (EBCTCG; 2012). A 10-year relative risk of distant recurrence of 0.76 was estimated, and was assumed to apply to the lymph node negative and positive groups. For MammaPrint the relative risk of distant recurrence for chemotherapy versus no chemotherapy was based on

data from the MINDACT trial on the discordant clinical and genomic risk groups. A 10-year relative risk of distant recurrence for the discordant populations was estimated to be 0.77. Sensitivity analyses explored the relative risks of distant recurrence in the modified Adjuvant! Online low- and high-risk subgroups, which were estimated to be 0.84 and 0.74, respectively.

In sensitivity analyses, the impact of assuming that Oncotype DX could predict the benefit of chemotherapy was explored, based on the studies reported by Paik et al. (2006) and Albain et al. (2010). In the node-negative group, the 10 year relative risks of relapse with chemotherapy versus no chemotherapy were 1.31, 0.61 and 0.26 for the low-, intermediate- and high-risk categories respectively. For the node-positive group, the 10-year relative risks of relapse with chemotherapy versus no chemotherapy were 1.02, 0.72 and 0.59 respectively.

Survival following onset of distant metastases

The survival prognosis of patients with distant metastases was based on an analysis of 77 women randomly selected from 232 women who had relapsed breast cancer between 2000 and 2005 (Thomas et al. 2009). Median survival was 40.1 months following distant recurrence. From this, the 6-month probability of death following distant recurrence was estimated to be 0.098, assuming a constant rate. The rate of death due to distant metastases was assumed to be the same across the different subgroups and across each test risk group.

The probability of local recurrence, developing acute myeloid leukaemia and survival thereafter

The model assumes that 10.5% of patients entering the distant recurrence health state had previously had a local recurrence, based on de Bock et al. (2009). The 6-month probability of developing acute myeloid leukaemia (AML) was estimated to be 0.00025, based on Wolff et al. (2015). Survival following the onset of AML was estimated to be approximately 8 months; assuming a constant event rate, this gave a 6-month probability of death following AML of 0.53. This was taken from the NICE technology appraisal guidance on

azacitidine for the treatment of myelodysplastic syndromes, chronic myelomonocytic leukaemia and acute myeloid leukaemia.

Costs

Costs and resource use inputs are described starting on page 371 of the diagnostics assessment report.

Test costs

The costs of the tumour profiling tests were based on prices submitted by companies, as shown in table 20.

Table 20 Test costs used in the model

Test	Cost	Comments
Oncotype DX	£2,580	Tests carried out in Genomic Health laboratory in US. Cost includes sample handling and customer service. A commercial-in-confidence discounted test cost was used in the model.
Prosigna	£1,970	Based on conducting the test in an NHS laboratory, which includes the laboratory costs (£240), the Prosigna kit (£1,650) and the nCounter System (£194,600)
EndoPredict	£1,500	Tests carried out in Myriad's laboratory in Munich
IHC4	£203	The cost was submitted using 2014 prices. The total cost of the test (£198) was uplifted using the HCHS indices to current prices.
MammaPrint	£2,326	Converted from Euros to UK Pounds Sterling assuming exchange rate of 1 British Pound to 1.15 Euro.
Abbreviations: HCHS, hospital and community health services		

Costs of adjuvant chemotherapy acquisition and administration

The costs associated with adjuvant chemotherapy were obtained from a previous costing analysis of the OPTIMA Prelim trial (Hall et al. 2017). The EAG model assumed that women with ER+, HER2-, early breast cancer with 0 to 3 nodes typically receive 1 of 4 adjuvant chemotherapy regimens:

- FEC100-T (fluorouracil, epirubicin, cyclophosphamide and docetaxel; 3+3 cycles; assumed to be given to 25% of patients)
- TC (docetaxel and cyclophosphamide; 4 cycles; assumed to be given to 20% of patients)
- FEC75 (fluorouracil, epirubicin and cyclophosphamide; 6 cycles; assumed to be given to 45% of patients)

- FEC100-Pw (fluorouracil, epirubicin, cyclophosphamide and weekly paclitaxel; 3+3 cycles; assumed to be given to 10% of patients)

The weighted mean cost of adjuvant chemotherapy acquisition, delivery and toxicity was estimated to be £3,145 per course.

Costs of endocrine therapy

The model assumed that all surviving patients received endocrine therapy for a period of between 5 and 8 years and may have received 1 of 4 endocrine therapy regimens:

- tamoxifen for 5 years (40% of patients, annual cost £35.06)
- anastrozole for 5 years (20% of patients, annual cost £14.09)
- letrozole for 5 years (20% of patients, 10% of patients were assumed to receive extended letrozole for 3 further years, annual cost £32.87)
- tamoxifen for 2 years then exemestane for 3 years (20% of patients, annual cost of exemestane £69.52).

Costs of additional treatments

The model assumed that 30% of women with early breast cancer would receive 4 milligrams of bisphosphonates (zoledronic acid) every 6 months by intravenous infusion for up to 3 years (cost per 36-month course = £58.50).

Follow-up costs

The model assumed that all patients received 2 routine follow-up visits during the first year after surgery, with annual visits thereafter for 5 years. Patients were also assumed to have a routine annual mammogram for up to 5 years. The cost of a routine follow-up visit was estimated to be £162.84, and the cost of a mammogram was estimated to be £46.37.

Costs of treatments for local and distant recurrence

Costs associated with treating local recurrence were taken from a UK-based costing analysis (Karnon et al. 2007) and uplifted to current prices using the HCHS (Hospital and Community Health Service) index (£13,913). This was applied as a once-only cost upon the incidence of distant recurrence.

Costs associated with treating distant metastases were derived from Thomas et al. (2009), and included visits, drugs, pharmacy, hospital admission and intervention, imaging, radiotherapy, pathology and transport. Cost components specifically associated with terminal care were excluded. The 6-monthly cost of treating metastatic breast cancer was estimated to be £4,541.

Health-related quality of life and QALY decrements

Health utilities were taken from various published studies as shown in table 21. The studies are described in detail starting on page 367 of the diagnostics assessment report.

Table 21 Health utilities applied in the model

Health state / event	Duration applied in model	Mean	Standard error	Source
Recurrence-free	Indefinite	0.824	0.002	Lidgren et al.
Distant metastases	Indefinite	0.685	0.004	
Disutility distant metastases	Indefinite	-0.14	0.11	Calculated using difference method
Local recurrence	Once-only QALY loss applied on transition to distant recurrence state	-0.108	0.04 (assumed)	Campbell et al.
Chemotherapy AEs	6 months	-0.038	0.004	Campbell et al.
AML	Indefinite	0.26	0.04 (assumed)	Younis et al.
Abbreviations: AEs, adverse events; AML, acute myeloid leukaemia; QALY, quality-adjusted life year				

Base-case results

For the purposes of decision-making, the incremental cost-effectiveness ratios (ICERs) per QALY gained or lost are considered. The following assumptions were applied in the base-case analysis:

- The proportion of patients who received chemotherapy under current practice (no test) was assumed to be the same for each test risk classification (low, intermediate, and high risk). This proportion was however assumed to differ between subgroups defined according to clinical risk (LN0 NPI≤3.4, LN0 NPI>3.4, LN1-3, MINDACT overall population, MINDACT modified Adjuvant! Online low risk, and MINDACT modified Adjuvant! Online high risk).

- Clinicians interpret each of the 3-level tests in the same way (for example, an Oncotype DX high-risk score would lead to the same chemotherapy decision as a Prosigna high-risk score).
- Clinicians interpret each of the 2-level tests in the same way (for example, a MammaPrint high-risk score would lead to the same chemotherapy decision as an EndoPredict high-risk score).
- The benefit of adjuvant chemotherapy was the same across all risk score categories for all tests.
- A proportion of patients (10.5%) who developed distant metastases had previously developed local recurrence (QALY losses and costs associated with local recurrence were applied once only).
- The prognosis of patients with AML and the costs and QALYs accrued within the AML state were independent of whether the patient had previously developed distant metastases.
- A disutility associated with adjuvant chemotherapy was applied once during the first model cycle only (while the patient is receiving the regimen).
- Costs associated with endocrine therapy, bisphosphonates, follow-up appointments and mammograms were assumed to differ according to time since model entry.
- Across all 3 analysis subgroups, patients entered the model aged 58 years, based on the mean age of patients in the NHS England Access dataset.
- The model included both pre- and postmenopausal women; however, the TransATAC study related only to postmenopausal women.

The results of the model are presented in the diagnostics assessment report starting on page 379, and are summarised below. In addition, the modelled chemotherapy use with and without the tumour profiling tests is presented in appendix 7 in the diagnostics assessment report (page 502). All estimates presented here are based on the probabilistic version of the model.

In the node-negative population, in the subgroup with an NPI of 3.4 or less, for tumour profiling tests compared with current practice the model gave ICERs of £147,419 per QALY gained (EndoPredict), £122,725 per QALY gained

(Oncotype DX), £91,028 per QALY gained (Prosigna) and £2,654 per QALY gained (IHC4+C).

In the node-negative population, in the subgroup with an NPI greater than 3.4, for tumour profiling tests compared with current practice the model gave ICERs of £46,788 per QALY gained (EndoPredict) and £26,058 per QALY gained (Prosigna). Oncotype DX was dominated by current practice (that is, Oncotype DX was more expensive and less effective) and ICH4+C was dominant over current practice (that is, ICH4+C was less expensive and more effective).

In the node-positive population, the tumour profiling tests compared with current practice (NPI) had ICERs of £28,731 per QALY gained (Prosigna) and £21,458 per QALY gained (EndoPredict). Oncotype DX was dominated by current practice and ICH4+C was dominant over current practice.

In the overall MINDACT population, MammaPrint compared with current practice (modified Adjuvant! Online) had an ICER of £131,482 per QALY gained. In the modified Adjuvant! Online high-risk subgroup, MammaPrint was dominated by current practice, and in the modified Adjuvant! Online low-risk subgroup, MammaPrint compared with current practice had an ICER of £414,202 per QALY gained.

QALYs, costs and ICERs for each test compared with current practice are presented in tables 22 to 26.

Table 22 Probabilistic ICERs for Oncotype DX compared with current practice

Option	QALYs	Costs	Inc. QALYs	Inc. costs	ICER (per QALY gained)
Node-negative NPI≤3.4					
Oncotype DX	13.89	£5,474	0.01	£1,313	£122,725
No test	13.88	£4,161	-	-	-
Node-negative NPI>3.4					
Oncotype DX	12.73	£11,806	-0.01	£881	Dominated
No test	12.74	£10,925	-	-	-
Node-positive (1-3 nodes)					
Oncotype DX	12.48	£13,212	-0.07	£687	Dominated

No test	12.55	£12,525	-	-	-
Abbreviations: ICER, incremental cost-effectiveness ratio; NPI, Nottingham prognostic index; QALY, quality adjusted life year					

Table 23 Probabilistic ICERs for IHC4+C compared with current practice

Option	QALYs	Costs	Inc. QALYs	Inc. costs	ICER (per QALY gained)
Node-negative NPI≤3.4					
IHC4+C	13.86	£4,291	0.01	£22	£2,654
No test	13.86	£4,269	-	-	-
Node-negative NPI>3.4					
IHC4+C	12.73	£10,941	0.01	-£90	Dominating
No test	12.72	£11,031	-	-	-
Node-positive (1-3 nodes)					
IHC4+C	12.59	£12,268	0.05	-£287	Dominating
No test	12.54	£12,554	-	-	-
Abbreviations: ICER, incremental cost-effectiveness ratio; NPI, Nottingham prognostic index; QALY, quality adjusted life year					

Table 24 Probabilistic ICERs for Prosigna compared with current practice

Option	QALYs	Costs	Inc. QALYs	Inc. costs	ICER (per QALY gained)
Node-negative NPI≤3.4					
Prosigna	13.87	£6,201	0.02	£1,884	£91,028
No test	13.84	£4,318	-	-	-
Node-negative NPI>3.4					
Prosigna	12.65	£13,330	0.06	£1,686	£26,058
No test	12.59	£11,644	-	-	-
Node positive (1-3 nodes)					
Prosigna	12.47	£15,172	0.07	£1,936	£28,731
No test	12.40	£13,236	-	-	-
Abbreviations: ICER, incremental cost-effectiveness ratio; NPI, Nottingham prognostic index; QALY, quality adjusted life year					

Table 25 Probabilistic ICERs for EndoPredict compared with current practice

Option	QALYs	Costs	Inc. QALYs	Inc. costs	ICER (per QALY gained)
Node-negative NPI≤3.4					
EndoPredict	13.85	£6,034	0.01	£1,679	£147,419
No test	13.84	£4,355	-	-	-
Node-negative NPI>3.4					
EndoPredict	12.71	£12,612	0.03	£1,388	£46,788
No test	12.68	£11,224	-	-	-
Node-positive (1-3 nodes)					
EndoPredict	12.52	£14,080	0.05	£1,164	£21,458

No test	12.46	£12,916	-	-	-
Abbreviations: ICER, incremental cost-effectiveness ratio; NPI, Nottingham prognostic index; QALY, quality adjusted life year					

Table 26 Probabilistic ICERs for MammaPrint compared with current practice

Option	QALYs	Costs	Inc. QALYs	Inc. costs	ICER (per QALY gained)
Overall MINDACT population					
MammaPrint	13.51	£9,151	0.01	£1,760	£131,482
No test	13.49	£7,391	-	-	-
MINDACT modified Adjuvant! Online high-risk group					
MammaPrint	12.86	£12,727	-0.04	£1,413	Dominated
No test	12.90	£11,313	-	-	-
MINDACT modified Adjuvant! Online low-risk group					
MammaPrint	13.70	£7,777	0.01	£2,410	£414,202
No test	13.69	£5,366	-	-	-
Abbreviations: ICER, incremental cost-effectiveness ratio; NPI, Nottingham prognostic index; QALY, quality adjusted life year					

Probabilistic sensitivity analysis

Probabilistic sensitivity analyses of pairwise (test compared with current practice) results (table 27) indicated that:

- In the lymph node negative, NPI of 3.4 or less subgroup, the only test with a non-zero probability of producing more net benefit compared with current practice at maximum acceptable ICERs of £20,000 and £30,000 per QALY gained was IHC4+C.
- In the lymph node negative, NPI of greater than 3.4 subgroup, at a maximum acceptable ICER of £20,000 per QALY gained, IHC4+C had a probability of 0.69 of being the most cost-effective option. All other tests had less than 0.24 probability of bring more cost effective than current practice. In the same subgroup, at a maximum acceptable ICER of £30,000 per QALY gained, IHC4+C had a probability of 0.67 of being the most cost-effective option and Prosigna had a probability of 0.60 of being the most cost-effective option. Oncotype DX had a probability of 0.04 and EndoPredict had a probability of 0.26 of producing more net benefit compared with current practice.

- In the lymph node positive subgroup, IHC4+C had probabilities of 0.95 and 0.94 of producing more net benefit compared with current practice at maximum acceptable ICERs of £20,000 and £30,000 respectively. In the same subgroup at the same maximum acceptable ICERs, the probability of EndoPredict producing more net benefit than current practice ranged from 0.44 to 0.73, and for Prosigna the range was 0.24 to 0.55. In this subgroup Oncotype DX had very low probabilities of producing more net benefit than current practice at the same maximum acceptable ICERs (0.01 or lower).
- In the overall MINDACT population and in the subgroups, the probability that MammaPrint would produce more net benefit than current practice at maximum acceptable ICERs of £20,000 and £30,000 per QALY gained was approximately zero.

Table 27: Probabilities of tests being cost effective

Test	Subgroup	Probability of being cost effective compared with current practice	
		At maximum acceptable ICER of £20,000	At maximum acceptable ICER of £30,000
Oncotype DX	LN0 NPI≤3.4	0.00	0.00
	LN0 NPI>3.4	0.01	0.04
	LN+ (1-3 nodes)	0.00	0.01
IHC4+C	LN0 NPI≤3.4	0.95	0.97
	LN0 NPI>3.4	0.69	0.67
	LN+ (1-3 nodes)	0.95	0.94
Prosigna	LN0 NPI≤3.4	0.00	0.00
	LN0 NPI>3.4	0.24	0.60
	LN+ (1-3 nodes)	0.24	0.55
EndoPredict	LN0 NPI≤3.4	0.00	0.00
	LN0 NPI>3.4	0.09	0.26
	LN+ (1-3 nodes)	0.44	0.73
MammaPrint	MINACT overall population	0.00	0.00
	Modified AO high risk	0.00	0.00
	Modified AO low risk	0.00	0.00

Abbreviations: AO, Adjuvant! Online; ICER, incremental cost-effectiveness ratio; LN, lymph node; NPI, Nottingham prognostic index

Deterministic sensitivity analyses

The EAG did several deterministic sensitivity analyses for each test. Full details are on page 377 of the diagnostics assessment report. Results for Oncotype DX were:

- Node-negative with NPI≤3.4: ICERs for Oncotype DX compared with current practice remained over £34,000 per QALY gained across all analyses.
- Node-negative with NPI>3.4: Oncotype DX was either dominated or had an ICER of more than £35,000 per QALY gained across almost all analyses. The only exception was when Oncotype DX was assumed to predict chemotherapy benefit. Within this analysis, Oncotype DX dominated current practice.
- Node-positive (1 to 3 nodes): Oncotype DX remained dominated across the majority of analyses. The exceptions were: when Oncotype DX was assumed to predict chemotherapy benefit (was dominant), and when the cost of chemotherapy was doubled (£3,700 saved per QALY lost).

Deterministic sensitivity analysis results for IHC4+C:

- Node-negative with NPI≤3.4: ICERs for IHC4+C compared with current practice remained below £16,000 per QALY gained across all analyses, except when post-test chemotherapy probabilities were derived from Holt et al. 2011 (£36,259 per QALY gained); in addition, IHC4+C dominated current practice when the cost of chemotherapy was doubled.
- Node-negative with NPI>3.4: IHC4+C dominated current practice or had an ICER below £6,000 per QALY gained across all scenarios.
- Node-positive (1 to 3 nodes): IHC4+C dominated current practice across all but 1 scenario. When the probability of receiving chemotherapy was based on the UKBCG survey, the ICER was estimated to be £1,929 per QALY gained.

Deterministic sensitivity analysis results for Prosigna:

- Node-negative with NPI≤3.4: ICERs for Prosigna compared with current practice were greater than £71,000 per QALY gained across all analyses.
- Node-negative with NPI>3.4: ICERs for Prosigna compared with current practice were below £34,000 per QALY gained across all analyses.
- Node-positive (1 to 3 nodes): ICERs for Prosigna compared with current practice were below £38,000 per QALY gained across all analyses.

Deterministic sensitivity analysis results for EndoPredict:

- Node-negative with NPI≤3.4: ICERs for EndoPredict compared with current practice remained greater than £91,000 per QALY gained across all analyses.
- Node-negative with NPI>3.4: ICERs for EndoPredict compared with current practice remained greater than £30,000 per QALY gained across all but 2 of the analyses. Exceptions were: when the UKBCG survey was used to inform the probability of receiving chemotherapy (£25,250 per QALY gained), and when Cusumano et al. (2014) was used to inform the probability of receiving chemotherapy conditional on the EndoPredict test result (£26,689 per QALY gained).
- Node-positive (1 to 3 nodes): ICERs for EndoPredict compared with current practice remained below £30,000 per QALY gained across all scenarios.

Deterministic sensitivity analysis results for MammaPrint:

- In the overall MINDACT population, the ICER for MammaPrint compared with current practice was estimated to be greater than £76,000 per QALY gained across all scenarios.
- In the modified Adjuvant! Online high-risk subgroup, MammaPrint was dominated by current practice across almost all scenarios.
- In the modified Adjuvant! Online low-risk subgroup, the ICER for MammaPrint compared with current practice was greater than £161,000 per QALY gained across all analyses.

Comparison between the new EAG model and the EAG model used for diagnostics guidance 10

The differences between the models are described in the diagnostics assessment report starting on page 398. The new EAG model suggested that in the lymph node negative with NPI>3.4 subgroup, Oncotype DX was dominated by current practice. In the same subgroup, the previous EAG model produced a base-case ICER for Oncotype DX compared with current practice of £22,600 per QALY gained. This ICER was also based on Oncotype DX offered at a confidential price through a patient access scheme.

The models had a similar general modelling approach. In both models, data on risk reclassification and risk of distant recurrence in the absence of chemotherapy were taken from analyses of the ATAC trial, although different datasets were used. The proportions of women who were assumed to receive chemotherapy conditional on the Oncotype DX risk score were taken from the NHS England access scheme dataset in the current EAG model, but the previous EAG model used unpublished data (Holt et al. 2013) to estimate this. In addition, the proportion of patients receiving chemotherapy in the standard care arm was taken from the NHS England access scheme dataset in the current EAG model, but was taken from English cancer registry datasets in the previous model.

When both models used pre- and post-test chemotherapy probabilities from the NHS England access scheme dataset and no predictive benefit is assumed, both models produce the same economic conclusion: Oncotype DX is dominated by current practice.

3 Summary

Clinical effectiveness

Among studies of lymph node negative patients receiving endocrine monotherapy, percentages of patients categorised as high risk ranged from 9 to 33% across all 5 tests. In studies of patients receiving endocrine monotherapy, 3 tests (Prosigna, EndoPredict and IHC4+C) categorised more

lymph node positive patients as high risk than lymph node negative patients. Oncotype DX, however, categorised similar numbers of patients as high risk in lymph node negative and positive groups. Oncotype DX also categorised more lymph node positive patients as low risk than other tests, but led to worse 10-year distant recurrence free survival/interval outcomes in this group compared with other tests.

All tests had statistically significant prognostic power in unadjusted analyses in lymph node negative and lymph node positive populations. All tests provided additional prognostic information over most commonly used clinicopathological factors and over clinical treatment score and Nottingham Prognostic Index (NPI) in lymph node negative patients. Results were more varied in lymph node positive patients.

There was some evidence of differential chemotherapy benefit between risk groups assessed by Oncotype DX, shown by significant interaction tests between risk group and chemotherapy treatment in unadjusted analyses. However, the interaction test results sometimes became non-significant when clinicopathological factors were adjusted for. Evidence on the ability of MammaPrint to predict benefit from chemotherapy was extremely limited, but suggested no statistically significant difference in effect of chemotherapy between risk groups. Evidence of differential chemotherapy benefit was not available for the other 3 tests.

For Oncotype DX and MammaPrint, evidence from observational, non-comparative studies assessing the impact of the test used prospectively in clinical practice suggested that recurrence/survival outcomes in low-risk groups were acceptable even with low rates of chemotherapy. There was no similar evidence relating to the other tests.

Decision impact studies reported that the percentage of patients with any change in chemotherapy recommendation or decision pre-/post-test ranged from 27% to 49% across UK studies (Oncotype DX, EndoPredict and IHC4+C) and from 5% to 70% across European studies (all tests except

IHC4). Across all tests, the net change in the percentage of patients with a chemotherapy recommendation or decision pre-/post-test ranged from an increase of 1% to a decrease of 23% among UK studies, and a decrease of 0% to 64% in European studies

Cost effectiveness

The EAG developed a de novo health economic model to assess the cost effectiveness of Oncotype DX, MammaPrint, Prosigna, EndoPredict and IHC4+C, each versus current practice. The base-case model suggested the following results:

Oncotype DX: In the lymph node negative NPI≤3.4 group, the incremental cost-effectiveness ratio (ICER) for Oncotype DX compared with current practice was estimated to be £122,725 per quality-adjusted life year (QALY) gained (£34,245 per QALY gained assuming prediction of chemotherapy benefit). In the lymph node negative NPI>3.4 and lymph node positive groups, Oncotype DX was dominated by current practice (but Oncotype DX dominated current practice if prediction of chemotherapy benefit was assumed). The results were primarily driven by the modelled reduction in the use of adjuvant chemotherapy using the Oncotype DX test because it categorised more lymph node positive patients as low risk than other tests, but this led to worse 10-year distant recurrence free survival/interval outcomes.

IHC4+C: In the lymph node negative NPI≤3.4 group, the ICER for IHC4+C compared with current practice was estimated at £2,654 per QALY gained. In the lymph node negative NPI>3.4 and lymph node positive groups, IHC4+C dominated current practice.

Prosigna: In the lymph node negative NPI≤3.4 group, the ICER for Prosigna compared with current practice was estimated to be £91,028 per QALY gained. In the lymph node negative NPI>3.4 and lymph node positive groups, the ICERs for Prosigna compared with current practice were estimated to be £26,058 and £28,731 per QALY gained, respectively.

EndoPredict: In the lymph node negative NPI≤3.4 group, the ICER for EndoPredict compared with current practice was estimated to be £147,419 per QALY gained. In the lymph node negative NPI>3.4 and lymph node positive groups, the ICERs for EndoPredict compared with current practice were estimated to be £46,788 and £21,458 per QALY gained, respectively.

MammaPrint: In the overall MINDACT population, the ICER for MammaPrint compared with current practice was estimated to be £131,482 per QALY gained. In the modified Adjuvant! Online high-risk group, MammaPrint was expected to be dominated by current practice. In the modified Adjuvant! Online low-risk group, the ICER for MammaPrint compared with current practice was estimated to be £414,202 per QALY gained.

4 Issues for consideration

Clinical effectiveness

Many of the included studies were retrospective analyses of randomised controlled trials (RCTs) or observational data sets which used stored tumour samples. Nearly all of these studies excluded patients who did not have a large enough tissue sample for testing, which leaves the evidence base at potential risk of spectrum bias, as patients with smaller tumours (who may be systematically different to those with large tumours) are likely to be under-represented. However, this issue is unavoidable in retrospective analyses.

The IHC4/IHC4+C evidence base was limited in that most of the data related to the IHC4 score alone, without the clinical score, and most studies used tertiles and quartiles to define low-, intermediate- and high-risk patients, which may not be useful in a clinical setting where fixed cut-offs are likely to be more practicable. In addition, there are known problems with conducting the analyses required for IHC4, in particular the reliability and reproducibility of the Ki-67 marker measurement.

Prosigna, EndoPredict and IHC4+C categorise more lymph node positive patients than lymph node negative patients as high risk. However, Oncotype

DX categorised similar percentages of lymph node positive and lymph node negative patients as high risk.

In terms of the prognostic ability of the tumour profiling tests, much of the evidence base was results from unadjusted analyses, which did not assess whether a test had additional value over clinicopathological factors. In adjusted analyses, the clinicopathological variables included were not consistent. Further, the retrospective observational studies reporting evidence on prognostic ability were at risk of confounding and spectrum bias, which can affect estimates of prognostic performance. This is because chemotherapy rates may differ by risk group, and if patients who received chemotherapy were excluded, these patients would be likely to be systematically different to those who did not. These problems were particularly relevant to the MammaPrint evidence base, as most studies were observational in nature rather than re-analyses of RCTs.

There were relatively limited data relating to the ability of Oncotype DX and MammaPrint to predict benefit from chemotherapy and on the ability of the tests to affect patient outcomes. These types of evidence were not available for the other 3 tests.

Concordance between tests was not fully reviewed, but 1 UK study (OPTIMA prelim) which compared Oncotype DX, MammaPrint, Prosigna and IHC4 concluded that although the tests assigned similar proportions of patients to low-, intermediate- and high-risk categories, test results for an individual patient could differ markedly depending on which test was used.

No data were available for men, who do account for a proportion of breast cancer cases seen in practice. It is not certain whether the prognostic and clinical-effectiveness data are applicable to men.

Cost effectiveness

The EAG model is subject to a number of uncertainties and limitations.

With the exception of Oncotype DX in the lymph node negative NPI>3.4 group, the evidence surrounding the pre- and post-test chemotherapy probabilities is subject to considerable uncertainty. The model results are sensitive to the assumptions made about pre- and post-test chemotherapy use. The inclusion of data collected through the NHS England access scheme dataset has a significant impact on the model results for Oncotype DX compared with the model results from the original assessment for diagnostics guidance 10 (in the lymph node negative NPI>3.4 subgroup, Oncotype DX was dominated by current practice in the current model, but had an incremental cost-effectiveness ration (ICER) of £22,600 per QALY gained in the previous EAG model). Further, there is only 1 UK-based decision impact study relating to a 2-level tumour profiling test (Bloomfield et al. 2017). Sensitivity analyses showed that alternative estimates of post-test chemotherapy use generally led to more favourable cost-effectiveness estimates for EndoPredict and MammaPrint.

The comparator in the model is defined as a modified version of Adjuvant! Online for the MammaPrint analyses, and as current practice for the other 4 tests. In clinical practice in England other tools may be used to define risk, such as the PREDICT algorithm. It was not possible to do a comparison with PREDICT, or to define clinical risk groups by PREDICT because data were not available from the TransATAC trial, the NCRAS data set or the MINDACT trial. The cost effectiveness of the tumour tests compared with current NHS practice is therefore highly uncertain.

There is uncertainty about whether Oncotype DX can predict chemotherapy benefit. The inclusion of this potential test characteristic in the model has a substantial impact on the results. When a predictive benefit was included, Oncotype DX dominated current practice in both the lymph node negative NPI>3.4 and lymph node positive (1 to 3 nodes) groups, and had an ICER of £34,245 per QALY gained in the lymph node negative NPI≤3.4 group.

The analysis of MammaPrint was based on a different data source from the other 4 tests. In addition, the MINDACT trial, which was used to inform the analysis of MammaPrint, had a follow-up period limited to 5 years.

The test cost for Prosigna was based on an efficient level of throughput. This may not hold if centres do not undertake the anticipated number of tests.

There is the potential for the prognostic performance of IHC4+C to have been overestimated. This is because the TransATAC trial was the derivation study for IHC4 and it is not certain how generalisable the prognostic model fitted from this dataset is.

The test risk classification probabilities and distant metastases-free survival probabilities for Oncotype DX, Prosigna, IHC4+C and EndoPredict were based on a postmenopausal population only (TransATAC). It is expected that the tumour profiling tests will also be used in premenopausal women.

5 Equality considerations

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others.

Breast cancer can occur in men, and it is often underdiagnosed and undertreated in this group. No data were identified for male only cohorts.

Women of African family origin are more likely to develop breast cancer at an earlier age and to have a more aggressive form of the disease compared with other women. Data relating to people of different ethnicities were difficult to interpret because of differences in treatment practices in different countries.

6 Implementation

NanoString does not offer a centralised testing service for Prosigna, so a local testing service would need to be established.

Standardisation and quality assurance programmes would be required before IHC4 could be used routinely in the NHS.

7 Authors

Frances Nixon

Topic Lead

Rebecca Albrow

Technical Adviser

November 2017

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

A. The diagnostics assessment report for this assessment was prepared by the School of Health and Related Research (ScHARR), The University of Sheffield:

Tumour profiling tests to guide adjuvant chemotherapy decisions in people with breast cancer (update of DG10)

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.

Manufacturers of technologies included in the final scope:

- Agendia NV
- Genomic Health UK
- Myriad Genetics
- Nanostring Technologies
- Royal Marsden Hospital Trust

Other commercial organisations:

- Decision Resources Group, Abacus
- Oncomark
- Roche Diagnostics

Professional groups and patient/carer groups:

- Association of Breast Surgery
- Breast Cancer Now
- The Royal College of Physicians
- The Royal College of Radiologists

Research groups:

- Cancer Research UK

Associated guideline groups:

None

Others:

- Colchester Hospital NHS Foundation Trust
- Department of Health
- Greater Manchester Cancer / NHS Trafford clinical commissioning group
- Healthcare Improvement Scotland
- Medicines and Healthcare products Regulatory Agency
- NHS England
- Peony Breast Care Unit
- The London Breast Clinic
- Welsh Government

Appendix B: Glossary of terms

Adjuvant therapy
Additional cancer treatment given after primary treatment to lower the risk that cancer will come back. Adjuvant therapy may include chemotherapy, radiotherapy, hormone therapy or biological therapy.

Distant recurrence
Cancer that comes back in a different area to the original cancer after initial treatment.

Hormone (endocrine) therapy
Hormones such as oestrogen and progesterone can fuel the growth of breast cancer. Hormone therapies, such as tamoxifen and aromatase inhibitors, aim to block the availability of hormones such as oestrogen and progesterone and prevent the cancer growing.

Local recurrence
Cancer that comes back in the same place as the original cancer after initial treatment.