Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer Addendum: EAG responses to key themes within the Comments on the Diagnostics Consultation Document

As part of the Diagnostic Assessment Programme topic "Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer", following the 1st Diagnostics Appraisal Committee meeting on 30 November 2017, NICE produced a Diagnostics Consultation Document (DCD, dated 10 January 2018).[1] Commentators provided comments on the DCD, and the EAG has responded to these comments in a separate document. This addendum provides responses to key themes within the comments document.

1. Use of TransATAC data in the economic model

1.1. Rationale for using TransATAC data in the EAG health economic model

All studies reporting prognostic ability or prediction of chemotherapy benefit and meeting the inclusion criteria were included in the clinical review. The rationale for using the TransATAC data in the EAG model was that it could be restricted to the population in the NICE scope (ER+ HER2- 0-3 positive nodes) and it was possible to split the node-negative patients into clinically low-risk and clinically intermediate-risk (according to NPI score above or below 3.4).

1.2. The TransATAC analysis is unreported and has not been subjected to scientific peer review

Several analyses of TransATAC focussing on different tumour profiling tests have been published in peer-reviewed journals. On behalf of the EAG, the TransATAC authors produced a bespoke analysis[2] which covered four of the five tests included in the DAR (Oncotype DX, EndoPredict, Prosigna and IHC4+C) and which was restricted to the relevant population as above.

Subsequent to the publication of the EAG report, the TransATAC authors have published a pre-planned analysis of these data in a peer-reviewed journal (Sestak *et al.,* 2018[3]).

Table **1** presents some key data from the bespoke analysis for the EAG[2] alongside the data from Sestak *et al*., 2018.[3] Whilst there are some small differences, these data are largely consistent. It is not possible to use the newly-published data[3] in our model since LN0 patients are not stratified into clinically low-risk and clinically intermediate-risk, and hazard ratios (HRs) are reported for a 1 standard deviation (1SD) change rather than between risk groups.

Table 1: A comparison of key analyses reported in the data request analysis[2] and in Sestak 2018[3]

Test	LN0 HR (95% CI) for 1SD 10 year		LN1-3 HR (95% CI) for 1SD 10 year		ΔLR-χ2 to CTS 10 year		
	Data request[2]	**Sestak 2018[3]**	**Data request[2]**	**Sestak 2018[3]**	**Data request[2] LN0**	**Data request LN1-3**	**Sestak 2018[3] LN0-3**
Oncotype DX	1.67 (1.39-2.01)	1.69 (1.40-2.03)	1.42 (1.05-1.91)	1.39 (1.05-1.85)	22.78 *p*<0.0001	4.75 *p*=0.023	15.2
IHC4+C	2.56 (1.98-3.33)	NR	1.83 (1.31-2.56)	NR	48.55 *p*<0.0001	12.60 *p*<0.001	NR
IHC4	NR	1.95 (1.55-2.45)	NR	1.33 (0.99-1.78)	NR	NR	20.1
Prosigna	2.58 (1.97-3.38)	2.56 (1.96-3.35)	1.59 (1.16-2.17)	1.58 (1.16-2.15)	50.77 *p*<0.0001	8.51 *p*=0.004	26.3
EPClin	2.34 (1.82-3.02)	2.14 (1.71-2.68)	1.84 (1.34-2.53)	1.69 (1.29-2.22)	40.60 *p*<0.0001	12.91 *p*<0.001	24.4

1.3 Patient numbers per subgroup are small

The number of patients per subgroup were: at least 410 for LN0 NPI<3.4 (more for some tests), at least 253 for LN0 NPI>3.4, and at least 192 for LN1-3. The EAG do not consider the subgroups to be unreliably small.

1.4. Overlapping confidence intervals for recurrence rates between risk groups and between tests

The EAG agrees that there is some overlap between confidence intervals. However, this does not prevent the data from being useable. The point estimates for recurrence per test risk group (for LN0 and LN+ patients) are consistent with estimates from other studies (see point 2 of this addendum, distant recurrence rates by risk classification). The EAG's probabilistic sensitivity analysis fully characterises the uncertainty surrounding these estimates.

1.5. Bias in the patient spectrum due to exclusion of small tumours with insufficient tissue

The EAG report noted this limitation. This is a limitation of most analyses using stored tumour samples and is not limited to TransATAC. A comparison of some basic population-level statistics between the MINDACT trial and the TransATAC data population was provided for the previous round of comments on the DAR, and no major differences were observed.

1.6. TransATAC includes postmenopausal women who were not suitable for chemotherapy

TransATAC selected patients who had not received chemotherapy in order to assess prognostic ability of tumour profiling tests, which required calculation of distant recurrence rates in the absence of

chemotherapy. The EAG report noted this limitation. Many other prognostic studies included in the systematic review also included patients receiving no chemotherapy, to allow a consistent assessment of prognostic ability. TransATAC does appear to include some patients who would be currently indicated for chemotherapy in the UK (e.g. LN>3).

2. Distant recurrence rates by risk classification

2.4. Consistency of Oncotype 10-yr outcomes across re-analyses of RCTs included in the review

Table 2 shows distant recurrence-free rates at 10 years across re-analyses of RCTs with endocrine monotherapy. Distant recurrence-free rates at 10 years in LN0 Oncotype DX low-risk patients (not subgrouped by clinical risk) are consistent across TransATAC publications (94.9% in the bespoke analysis;[2] 94.1% in the Sestak 2016 SABCS presentation;[4] 96% in Dowsett et al. 2010,[5] the latter being measured at 9 years rather than 10 years). These rates are also consistent with those from other studies: B14[6] (93.2%) and B20[7] (96.8%), for patients in the no-chemotherapy arms. Outcomes for other risk groups were also consistent across studies (Table 2).

Table 2: 10-year distant recurrence for Oncotype DX (RCT re-analyses; endocrine monotherapy)

Nodal status	**Oncotype DX risk group**	**Percent of patients distant recurrence-free at 10 years (95% CI)**				
		TransATAC data request[2]	**TransATAC (Sestak 2016 SABCS[4])**	**TransATAC (Dowsett 2010;[5] 9yr recurrence)**	**B14 (Paik 2004,[6] Tang 2011a[8])**	**B20 (Paik 2006[7])**
LN0	ODX low	94.9	94.1	96 (93 to 97)	93.2 (90.4, 96.0)	96.8 (93.7, 99.9)
LN0	ODX int	87.7	83.3	88 (82 to 92)	85.7 (79.7, 91.7)	90.9 (82.5, 99.4)
LN0	ODX high	77.2	72.8	75 (66 to 83)	69.5 (62.6, 76.4)	60.5 (46.2, 74.8)
		LN1-3 only	**Incl LN4+**	**Incl LN4+**		
LN+	ODX low	81.8 (72.7-88.0)	73.8	83 (76 to 88)		
LN+	ODX int	75.4 (63.0-84.2)	65.3	72 (61 to 80)		
LN+	ODX high	68.6 (44.7-83.9)	51.2	51 (36 to 65)		

Data from Table 12 in EAG report. No additional RCTs of endocrine monotherapy reported distant recurrence in LN+ patients.

2.5. Consistency of outcomes across studies: Oncotype low-risk patients subgrouped by clinical risk

There are several comments referring to the 10-year distant recurrence rate of 15% in the LN0 Oncotype DX low-risk group in the TransATAC analysis (i.e. 85.4% distant recurrence-free). It is vital to point out that this does not represent the Oncotype DX low-risk group as a whole (see response 2.1 and Table 2 for the whole Oncotype DX low risk group). Instead, it represents the LN0 NPI>3.4 subgroup (i.e. LN0 and clinically intermediate-risk).

Table 3 shows distant recurrence-free rates at 10 years for LN0 patients, subgrouped by clinical risk. For TransATAC, these were subgrouped according to NPI score (which includes nodal status, tumour grade and tumour size). For the Oncotype DX low-risk, clinically intermediate subgroup (NPI>3.4), the distant recurrence-free rate at 10 years was 85.4%. We could not identify any other studies subgrouping by NPI score. However, the B14 analysis subgrouped by various other measures of clinical risk: tumour size, grade and Adjuvant! Online (AOL).[6, 8] B14 results appeared consistent with TransATAC, with similar 10-year distant recurrence-free rates for Oncotype DX low-risk, clinically intermediate-risk patients (tumour >4cm, 87%; grade poor-differentiated, 86%; AOL intermediate-risk, 86.6%, AOL high-risk, 95.0%). Outcomes for other Oncotype DX risk groups sub-grouped by clinical status were also consistent across studies (Table 3).

Table 3: 10-year distant recurrence for Oncotype DX by clinical risk group (RCT re-analyses)

Oncotype DX risk group	**Clinical risk**	**TransATAC data request[2] LN0**		**B14 (Paik 2004,[6] Tang 2011a[8]) LN0**	
		Definition of clinical risk	**% DRF at 10yr (95% CI)**	**Definition of clinical risk**	**% DRF at 10yr**
ODX low	Clinical low	NPI≤3.4	98.3 (96.3-99.2)	Tumour <1cm	100
				Grade well-diff	96
				AOL low-risk	94.4
	Clinical intermediate	NPI>3.4	85.4 (77.6-90.7)	Tumour >4cm	87
				Grade poor-diff	86
				AOL int-risk	86.6
				AOL high-risk	95.0
ODX int	Clinical low	NPI≤3.4	93.1 (86.7-96.5)	Tumour <1cm	87
				Grade well-diff	91
				AOL low-risk	90.0
	Clinical intermediate	NPI>3.4	79.8 (69.4-86.9)	Tumour >4cm	88
				Grade poor-diff	76
				AOL int-risk	86.1
				AOL high-risk	76.6
ODX high	Clinical low	NPI≤3.4	83.8 (57.7-94.5)	Tumour <1cm	83
				Grade well-diff	69
				AOL low-risk	81.8
	Clinical intermediate	NPI>3.4	74.9 (59.8-85.1)	Tumour >4cm	47
				Grade poor-diff	60
				AOL int-risk	56.8
				AOL high-risk	68.5

TransATAC data from Table 124 in EAG report. B14 data by size/grade estimated from graphs in Paik 2004.[6] DRF, distant recurrence-free

2.6. Consistency of Oncotype 5yr outcomes between TransATAC and observational studies

There were several comments suggesting that the TransATAC recurrence rates used in the EAG model were less favourable than the recurrence rates from observational studies of Oncotype DX. Table 4 shows outcomes at 5 years for TransATAC and for observational studies of Oncotype DX (no 5-year data were available for other reanalyses of RCTs). Outcomes at 5 years were similar between

TransATAC and observational studies of Oncotype DX. It should be noted that some patients in the observational studies received chemotherapy; this may have improved observed outcomes.

The differences between the TransATAC recurrence rates used in the EAG model and the recurrence rates reported in observational studies appear to be due to: (a) the model data being stratified by clinical risk (those with NPI >3.4 had less favourable outcomes), and (b) the observational data being reported at a 5-year rather than 10-year follow-up.

Table 4: 5-year outcomes for Oncotype DX (RCTs and observational studies; some chemotherapy use)

Oncotype DX risk group			**LN0-mic**						**LN0-3, clin high risk**
	TransATAC data request[2] (LN0) N=829		**CT use in obs. studies**	TAILORx (Sparano 2015[9]) N=1626	MD Anderson (Le Du 2015[10]) N=1030	Clalit (Stemmer 2016[11]) N=1594	Memorial Sloan Kettering (Wen 2017[12]) N=1406	SEER (Petkov 2016,[13] Roberts 2016[14]) N=38,568	WSG PlanB (Nitz 2017 [15-17]) N=2646
	CT use	**DRFI 5yr**		**DRFS 5yr**	**DRFS 5yr**	**DRFI 5yr**	**DRFI 5yr**	**BCSS 5yr**	**IDFS 5yr**
ODX very low (<11/12)	None		0%	99.3 (98.7, 99.6)			99.9%	99.6 (99.4, 99.8)	94.2 (91.2, 97.3)
ODX low (RS<18)	None	99.1	1-12%	-	95.9 (93.0, 97.6)	99.5 (98.4, 99.8)	99.6%	99.6 (99.4, 99.7)	
ODX int (RS 18-30)	None	94.0	26-43%		-	98.8 (97.2, 99.4)		98.6 (98.3, 98.9)	94.3 (92.8, 95.8) **(RS 12-25)**
ODX high (RS >30)	None	88.9	89-90%		76.4 (59.2, 87.1)	93.1 (87.1, 96.3)		95.6 (94.4, 96.6)	84.2 (80.6, 87.8) **(RS ≥25)**

Data from Table 26 in EAG report. CT, chemotherapy; DRFS, distant recurrence-free survival; DRFI, distant recurrence-free interval; IDFS, invasive disease-free survival; BCSS, breast cancer-specific survival

3. Ability of Oncotype DX to predict differential relative benefit from adjuvant chemotherapy

3.1. Clarification on the difference between absolute and relative benefit

A key issue for clinical and cost-effectiveness of tumour profiling tests is whether the **relative** benefit from chemotherapy differs between test risk groups. It is important to note that this relates to relative rather than absolute benefit. We concluded in our EAG report that all the tests have additional prognostic ability over clinicopathological factors, at least in LN0 patients, i.e. that recurrence rates are higher in higher-risk groups. This means that the **absolute** benefit of chemotherapy is also higher in higher-risk groups. However, this does not necessarily mean that the relative benefit differs between groups.

As an example, if distant recurrence rates in the test high-risk group were 30% without chemotherapy and 20% with chemotherapy, the absolute benefit of chemotherapy would be 10%. Likewise, if distant recurrence rates in the test low-risk group were 3% without chemotherapy and 2% with chemotherapy, the absolute benefit of chemotherapy would be 1% (i.e. much smaller). However, the relative benefit would be the same in both groups (relative risk of 0.67, i.e. chemotherapy reduces recurrence by one-third).

3.2. Summary of data on the ability of Oncotype DX to predict benefit from chemotherapy

Data on ability of Oncotype DX to predict differential relative chemotherapy benefit is summarised in this section. Limitations of the chemotherapy benefit studies are summarised in Section 3.3. The EAG's overall view on chemotherapy benefit data is provided in Section 3.4.

Data on the ability of Oncotype DX to predict chemotherapy benefit comes mainly from two re-analyses of RCTs: one in LN0 patients (NSABP-B20; Paik 2006,[7] Tang 2011a[8]) and one in LN+ (SWOG-8814, Albain 2010[7, 8, 18]). In both, patients were randomised to endocrine monotherapy or endocrine plus chemotherapy. Summary results are provided in Table 5.

Relative and absolute benefit per risk group (adjusted and unadjusted): Both studies showed that unadjusted HRs for the effect of chemotherapy vs. no chemotherapy on survival and recurrence outcomes were most favourable in the higher-risk groups. HRs were generally statistically significant in high-risk groups but not in low- or intermediate-risk (). In the B20[7, 8] study (LN0), unadjusted HRs for 10-year distant recurrence-free interval (DRFI) in the low, intermediate and high-risk groups were 1.31, 0.61 and 0.26. HRs restricted to HER2- patients (adjusted and unadjusted) showed the same pattern (Table 5; not reported in journal article - provided via personal communication with Dr Tang via NICE). However, it is interesting to note that absolute differences (for chemotherapy vs. no chemotherapy) were very small in the low and intermediate-risk groups (1.1% and 1.8%, both favouring no chemotherapy), though greater in the high-risk group (27.6% favouring chemotherapy).

In SWOG-8814 (LN+),[18] DRFI was not reported. HRs for 10-year disease-free survival (DFS) for low, intermediate and high-risk groups, adjusted for number of positive nodes, were 1.02, 0.72 and 0.59.

Interaction tests (adjusted and unadjusted): Interaction tests indicate whether the difference in chemotherapy effect for a change in RS score is statistically significant. In B20 (LN0), the unadjusted interaction test for 10-year DRFI (for continuous RS score by chemotherapy) was reported as p=0.031[8] or p=0.038,[7] indicating a statistically significant difference in chemotherapy benefit as RS changes (Table 5). Interaction tests adjusted for clinicopathological factors were borderline significant for the full cohort (p=0.035, p=0.039 and p=0.068; difference due to method of assessing grade), while for the HER2- subgroup they were statistically significant (p=0.007, p=0.018 and p=0.022). The EAG report stated that it was unclear whether all factors were adjusted for simultaneously in B20; however, personal communication with the biostatistician (via NICE) confirms that this was the case.

In SWOG-8814 (LN+), the interaction test for 10-year DFS (for continuous RS score by chemotherapy; adjusted for number of nodes) was p=0.053 for all years and p=0.029 for years 0-5. Interaction tests adjusted individually for each of age, ethnicity, tumour size, grade, PR, P53 and HER2 were also statistically significant (p=not reported). Initially, the EAG interpreted this as a model including all clinicopathological variables; however, clarification from the authors in a personal communication to the EAG stated that each variable was included in a separate model. However, an interaction test adjusted for Allred-scored ER status was not significant (p=0.15). No interaction test was available that included all clinicopathological variables together.

Observational studies: Three observational studies had some data on chemotherapy benefit: two studies in patients with LN0 disease (MD Anderson[10, 19] and SEER[14, 20]) and one study in patients with LN+ disease (Clalit Health[21, 22]). Evidence was mixed and at high risk from confounding, since receipt of chemotherapy was influenced by Oncotype DX score, and patients receiving chemotherapy were likely to be at higher risk. Only one study (SEER) reported an interaction test; this was statistically significant (p=0.03), but only adjusted for grade, tumour size, age and race (omitting ER and PR).[13, 14] The other two studies only reported HRs for chemotherapy versus no chemotherapy in intermediate (MD Anderson and Clalit Health)[10, 11, 19, 21, 22] and high-risk patients (MD Anderson),[10, 19] and these were statistically non-significant, even after adjustment for confounders in one study.[10, 19]

Table 5: Prediction of chemotherapy benefit by Oncotype DX – Reanalyses of RCT data

Study	Outcome	% recurrence-free; absolute benefit			Hazard ratio for CT vs no CT (95% CI)			Interaction tests	Adjusted interaction tests
		Low	Intermediate	High	Low	Intermediate	High		
NSABP-B20 **LN0** ER+ N=651 Paik 2006[7] Tang 2011a[8] Personal comm.	**DRFI 10yr Unadjusted** **HER2- Unadjusted** **HER2- Adjusted[a]**	CT: 95.6% No CT: 96.8% Abs diff -1.1%	CT: 89.1% No CT: 90.9% Abs diff -1.8%	CT: 88.1% No CT: 60.5% Abs diff 27.6%	1.31 (0.46, 3.78), *p*=0.61 1.21 (0.41, 3.55), *p*=0.73 1.18 (0.40, 3.53), *p*=0.76[a]	0.61 (0.24, 1.59), *p*=0.39 0.78 (0.29, 2.11), *p*=0.62 0.67 (0.24, 1.87), *p*=0.44[a]	**0.26 (0.13, 0.53), p<0.001** **0.21 (0.08, 0.53), p<0.001** **0.20 (0.07, 0.52), p=0.001[a]**	Interaction (continuous RS) ***p*=0.031** or ***p*=0.038** (Tang 2011a[8] and Paik 2006[7])	Interaction[a] (continuous RS) adjusted for age, tumour size, grade, ER and PR: - All pts: ***p*=0.035, 0.039**, 0.068[b] - HER2-: ***p*=0.007, 0.018, 0.022** [b]
	DFS 10yr				0.91 (0.57, 1.45)	0.79 (0.43, 1.47)	**0.41 (0.23, 0.71)**	p=0.082	
	OS 10yr				1.37 (0.63, 3.01)	0.94 (0.4, 2.25)	**0.31 (0.16, 0.60)**	**p=0.011**	
SWOG-8814 **LN+** HR+ HER2+/- N=367 Albain 2010[18]	**DFS 10yr**	CT: 64% No CT: 60% Abs diff 4%		CT: 55% No CT: 43% Abs diff 12%	1·02 (0·54, 1·93); *p*=0·97 [c]	0·72 (0·39, 1·31); *p*=0·48[c]	0·59 (0·35, 1·01); ***p*=0·033**[c]		- Interaction (continuous RS) adjusted for positive nodes: All years: *p*=0.053 [c] 0-5 years: ***p*=0.029** [c] 5-10 years: *p*=0.58 [c] - Interaction (continuous RS) adjusted for each of age, ethnicity, size, grade, PR, P53, HER2: significant (p=NR). - Interaction adjusted for Allred-scored ER: p=0·15
	BCSS 10yr			CT: 73% No CT: 54% Abs diff 19%	*p*=0.56	*p*=0.89	***p*=0.033** [c]		
	OS 10yr			CT: 68% No CT: 51% Abs diff 17%	1·18 (0·55, 2·54, p=0·68)[c] p=0.63 log-rank	0·84 (0·40, 1·78, p=0·65) [c] p=0.85 log-rank	0·56 (0·31, 1·02, *p*=0·057) [c] ***p*=0.027** log-rank		Interaction (continuous RS) [c] All yrs: ***p*=0.026** 0-5 yrs: ***p*=0.016** 5-10 yrs: *p*=0.87

Data from Table 22 in EAG report. [a]Adjusted for age, tumour size, grade, ER and PR.[b]p-values correspond to analyses using different assessments of tumour grade. [C]Adjusted for number of positive nodes (1 to 3 vs. 4 or more)

3.3. Key limitations of studies assessing chemotherapy benefit

a) *Lack of data on chemotherapy benefit for the clinically intermediate-risk group:* NICE currently recommends Oncotype DX only for patients who are clinically intermediate-risk, for whom the chemotherapy decision is uncertain. This is a key subgroup for the economic modelling (defined as NPI>3.4). There are no data on the chemotherapy effect in patients who are Oncotype DX low-risk but clinically intermediate-risk. It is plausible that even if there is no chemotherapy benefit for clinically-low Oncotype DX-low patients, there could be benefit for clinically-intermediate (NPI>3.4) Oncotype DX-low patients.

b) *Statistical significance of interaction tests:* Most unadjusted interaction tests were statistically significant (Table 5). In terms of adjusted interaction tests, these were significant or borderline significant in B20 (LN0); and more clearly significant for the new HER2- subgroup (personal communication via NICE). One of the key concerns in the EAG report was that it was unclear whether all factors were adjusted for simultaneously in B20; however, personal communication with the biostatistician confirms that this was the case. This, along with the new HER2- subgroup analysis, provides stronger evidence for an interaction than presented in the EAG report.

However, in SWOG-8814 (LN+), it is now apparent after clarification from the lead biostatistician that interaction tests were adjusted for each clinicopathological factor individually (not all together, as initially thought by the EAG). All were individually significant except for the interaction test adjusted for Allred-scored ER status (p=0.15). As such, it remains unclear whether the interaction test would remain significant after adjustment for all relevant clinicopathological variables.

This also raises an interesting point as to whether results should be adjusted for ER status. On the one hand, test results should be adjusted to account for the effect of clinicopathological factors for which data are available in routine practice. On the other hand, it is not clear to what extent quantitative ER results are routinely available in UK practice, or their level of analytic validity; the SWOG-8814 author noted in his personal communication that performance of the Allred score is subject to some variability between pathologists. The author further stated that *"It is certainly possible that by including other measures of HER2, ER degree, Ki-67, grade, nodal size etc that one could make the interaction nonsignificant. However … you do get the benefit of most of those in a single well controlled measure (RS) rather than relying on separate assays for each with high known variability."* In other words, the benefit of Oncotype DX could be more accurate prognosis, rather than the prediction of chemotherapy benefit.

c) *Possible overestimation of chemotherapy benefit due to B20 being derivation study:* Patients from the no-chemotherapy arm of B20 were used to derive the Oncotype DX score. Therefore, Oncotype DX

may be overfitted in this study arm (i.e. recurrence rates may be artificially low in Oncotype low-risk patients and artificially high in Oncotype DX high-risk patients). This could lead to an overestimate of chemotherapy benefit since the chemotherapy arm was not used in derivation, therefore recurrence rates in this arm may show less separation between the low and high risk groups.

B14 (Paik 2004)[6] is a validation study of Oncotype DX (tamoxifen only; no chemotherapy arm). Comment 162 notes that the prognostic effect of Oncotype DX in the no-chemotherapy arm of B20 is greater than that in B14. As shown in Table 6, in the absence of chemotherapy, there is greater separation in B20 than B14; in other words, low-risk patients have a better 10-year recurrence-free rate in B20 (96.8%) than B14 (93.2%), while high-risk patients have a worse recurrence-free rate in B20 (60.5%) than B14 (69.5%).

In terms of prediction of chemotherapy benefit, B20 has a worse recurrence-free rate in the chemotherapy arm in low-risk patients (95.6% with chemotherapy vs. 96.8% without). This is counter-intuitive, and gives a corresponding HR greater than 1 (HR=1.31). However, comparing the chemotherapy arm of B20 (95.6% recurrence-free) with the no-chemotherapy arm of B14 (93.2% recurrence-free) indicates a small benefit in low-risk patients, though this breaks randomisation and may be affected by population differences between trials.

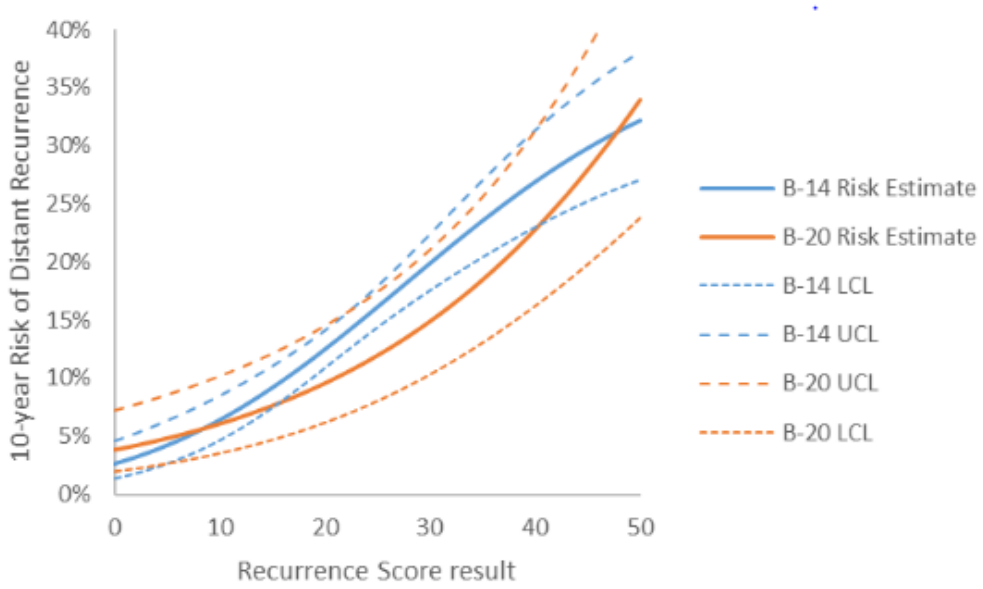
Additional data (personal communication with Dr Tang) compares the recurrence rates for a range of Oncotype DX scores in B14 and B20 (Figure 1). This analysis (which uses continuous Oncotype DX scores) is interpreted by Dr Tang as suggesting that the range of distant recurrence risk estimates, and slopes, are very similar between B20 and B14. However, the EAG still note that recurrence rates per risk group do appear to show greater separation in B20 than B14 (Table 6).

Table 6: Comparison of Oncotype prognostic ability in B14 and B20

Oncotype risk group	NSABP-B14 (Paik 2004)[6]		NSABP-B20 (Paik 2006[7])			
	Tamoxifen		Tamoxifen		Tamoxifen + chemotherapy	
	% patients per risk group (n)	% recurrence-free 10yr	% patients per risk group (n)	% recurrence-free 10yr	% patients per risk group (n)	% recurrence-free 10yr
Low	51% (388)	93.2%	60% (135)	96.8%	51% (218)	95.6%
Intermediate	22% (149)	85.7%	20% (45)	90.9%	21% (89)	89.1%
High	27% (181)	69.5%	21% (47)	60.5%	28% (117)	88.1%

Data from Table 12 in EAG report (also comment 161a in Comments on Diagnostics Consultation Document)

Figure 1: 10yr risk of distant recurrence in tamoxifen-alone groups: B20 and B14 (personal communication with Dr Tang, B20 study)



d) *Clinical relevance of chemotherapy benefit is unclear for the Oncotype DX intermediate-risk group:* Hazard ratios for chemotherapy benefit are available for this group, but it is unclear how they should be interpreted in clinical practice, i.e., would patients be treated, not treated, or would other clinicopathological variables be taken into consideration when making a decision?

e) *The number of events per subgroup is relatively low*, particularly for the B20 study (Table 7). Confidence intervals for the hazard ratios in low-risk and intermediate-risk groups are very wide in both B20 and SWOG-8814 (Table 5).

Table 7: Event rates for B14, B20 and SWOG-8814

Oncotype risk group	**Treatment**	**N events / N patients**		
		B14 (Paik 2004)[6] LN0	**B20 (Paik 2006)[7] LN0**	**SWOG-8814 (Albain 2010),[18] LN+**
Low	Chemo	-	10 / 218	26 / 91
Low	No chemo	28 / 338	5 / 135	15 / 55
Intermediate	Chemo	-	9 / 89	20 / 57
Intermediate	No chemo	25 / 149	7 / 45	22 / 46
High	Chemo	-	13 / 117	28 / 71
High	No chemo	56 / 181	18 / 47	26 / 47

3.4. EAG summary of evidence and limitations for prediction of chemotherapy benefit by Oncotype

Both B20 (LN0) and SWOG-8814 (LN+) showed that hazard ratios for chemotherapy vs. no chemotherapy were most favourable in the higher-risk groups, and were generally statistically significant in high-risk groups but not in low- or intermediate-risk groups. Unadjusted interaction tests

were statistically significant. Adjusted interaction tests were borderline significant in B20 (significant in HER2- patients), while in SWOG-8814 they were significant when adjusted for some clinicopathological variables individually, but not when adjusting for ER determined by Allred status.

Considering the limitations discussed above, the EAG considers that there remains uncertainty surrounding whether Oncotype DX is associated with a predictive benefit of chemotherapy (i.e. a difference in relative effect by genomic risk group), and if so, that there is uncertainty in the likely magnitude of this predictive effect within the clinical subgroups considered in this appraisal.

3.5. Observational studies showing low recurrence rates in test low-risk groups: to what extent does this bypass the issue of whether tests are predictive for chemotherapy benefit?

Some comments have noted the low recurrence rates within Oncotype low-risk groups in large observational studies. These are summarised in Table 4. LN0 patients with RS<18 have been reported as having a 5-year DRFS of 95.9%[10] and a 5-year DRFI of 99.5-99.6%.[10-14] For LN0-mic patients with RS<11/12, reported rates of 5-year DRFS, DRFI and BCSS range from 99.3-99.9%.[9, 12-14] The fact that TAILORx has not yet reported final results also indicates that recurrence rates are likely to be low.

Some commentators question whether these low recurrence rates in low-risk patients bypass the issue of whether tests are predictive for chemotherapy benefit. This is an important consideration. However, the EAG consider the following points to be important here:

a) The low-risk RS cut-off is currently 18 rather than 11 or 12, according to the NICE scope, the manufacturers, UK clinical practice, and NHS England Access Scheme data. Despite this, data using the RS<11/12 cut-point were included in the EAG clinical review for completeness.

b) NICE currently recommends Oncotype DX only for patients who are clinically intermediate-risk, for whom the chemotherapy decision is uncertain. This clinically-intermediate subgroup is a key subgroup for the economic modelling (defined as NPI>3.4). Conversely, the observational studies (as well as the reanalyses of RCTs) include a range of clinically low- and intermediate-risk patients. Patients who are RS low-risk but clinically intermediate-risk have a higher recurrence rate than the wider RS low-risk group, as shown in both TransATAC and B14 (see Table 3). The observational evidence may include patients who would not require an Oncotype DX test in UK clinical practice due to their low clinical risk, and may mask a subgroup of clinically-intermediate risk patients with higher recurrence rates.

c) The issue of predictive performance remains important for the modelling, because whether to accept the very different **relative** chemotherapy benefits between high-risk and low-risk patients (e.g. from the B20 study, with its limitations as discussed above) has a large impact on cost-effectiveness.

4. Risk of recurrence after 5 years

As noted in the EAG report, the assumptions employed in the model regarding the long-term risk of distant recurrence and the impact of chemotherapy are based on the earlier model reported by Ward *et al*[23] used in NICE DG10.[24] These assumptions are also applied in the Genomic Health model. As noted in the EAG's response to consultation on the assessment report, whilst there is some evidence which suggests that for some patients with particular disease subtypes, recurrence rates remain approximately constant between 5 and 20-years, there is also uncertainty surrounding the duration over which the benefit of chemotherapy is sustained, hence constraining recurrence at 15-years reduces the likelihood of overestimating this benefit of chemotherapy. We undertook sensitivity analyses in which the risk tapering assumption is removed (see EAG report, Tables 139, 142, 145, 148 and 151); these sensitivity analyses indicate that removing the assumption of capped recurrence risk does not significantly impact upon the conclusions drawn from the analysis.

5. Adverse effects of chemotherapy

5.1 Additional EAG sensitivity analysis - Inclusion of additional adverse events

In response to the DCD, several commentators have criticised the EAG model for excluding long-term adverse events (AEs) associated with chemotherapy, for example, chronic heart failure (CHF), permanent alopecia and peripheral neuropathy. As noted in the original EAG report, CHF was excluded from the EAG model due to a lack of evidence on the joint survival impact of CHF and metastatic breast cancer.

Within this addendum, the EAG has undertaken exploratory analyses to assess the potential impact of including these potential late effects of chemotherapy on the cost-effectiveness of the tumour profiling tests.

Estimated lifetime QALY losses and costs associated with CHF were obtained from a re-analysis of the model previously developed as part of the OPTIMA-Prelim study (Hall *et al*[25]); this was one of a minority of studies identified within the EAG's review which included this late effect of chemotherapy. The lifetime impact of CHF was estimated using the Hall *et al* model by comparing two scenarios: (i) all patients receive adjuvant chemotherapy (including excess CHF risk), and; (ii) the excess CHF risk is set equal to zero (although background levels of CHF are still included).

In addition, the EAG has included additional disutilities associated with permanent alopecia and peripheral neuropathy, based on studies identified within a systematic review of studies reporting utility values associated with AEs of chemotherapy (Shabaruddin *et al*[26]). Of the range of potentially relevant disutilities reported in the review, studies were considered potentially relevant for inclusion in the exploratory analysis if they: (a) included a counterfactual state for comparison (i.e. the same state

without the AE), and (b) if the valuations were elicited from the general public (rather than from patients experiencing the AE or from health care practitioners acting as proxy for patients). The selected disutility for alopecia was based on a general population time trade-off (TTO) study of lung cancer states reported by Nafees *et al.*[27] The disutility for peripheral neuropathy was based on a general population TTO study of colorectal cancer states reported by Shiroiwa *et al.*[28]

These additional HRQoL and cost impacts were included in the EAG's model, based on the assumptions set out in Table 8. The results of the analysis are shown in Table 9.

Table 8: Additional assumptions included in EAG's sensitivity analysis

Adverse event	Incidence	Health loss	Cost
Acute myeloid leukaemia (AML)	0.49% at 10-years (Wolff *et al*[29])	Health state utility = 0.26	Lifetime cost £10,400
CHF	Based on excess CHF risk relative to that of the general population	Net lifetime QALY loss -0.0385 QALYs (Hall *et al*[25])	Net lifetime cost -£2 (Hall *et al*[25])
Alopecia	15% of all patients receiving chemotherapy (commentator opinion)	Disutility = -0.04495 (Nafees *et al*[27])	Cost not included in analysis
Peripheral neuropathy	12% of all patients receiving chemotherapy (commentator opinion)	Disutility = -0.02 (Shiroiwa *et al*[28])	Cost not included in analysis

Table 9: Central estimates of cost-effectiveness

Test	Scenario	NPI≤3.4	NPI>3.4	LN+ (1-3 nodes)
Oncotype DX	EAG base case	£120,144	Dominated	Dominated
	Additional AEs included	£121,270	£548,524	Dominated
IHC4+C	EAG base case	£2,752	Dominating	Dominating
	Additional AEs included	£1,735	Dominating	Dominating
Prosigna	EAG base case	£89,693	£25,857	£28,666
	Additional AEs included	£88,114	£25,277	£31,807
EPClin	EAG base case	£141,848	£46,482	£21,489
	Additional AEs included	£350,042	£46,310	£19,911
Test	**Scenario**	**MINDACT ITT**	**MINDACT high-risk**	**MINDACT low-risk**
MammaPrint	EAG base case	£134,059	Dominated	£399,182
	Additional AEs included	£59,193	Dominated	£848,869

As shown in Table 9, the economic conclusions drawn from the analyses are largely unchanged by the inclusion of these additional AEs, although the inclusion of alternative disutilities may lead to different results. The EAG has a number of concerns regarding the reliability of this additional exploratory analysis:

- The QALY losses and costs associated with CHF have been derived from a separate model (Hall *et al*[25]).
- The baseline health state utilities for the relapse-free and post-relapse states included in the EAG model (taken from Lidgren *et al*[30]) may already include a proportion of patients who are experiencing AEs at the time of HRQoL assessment.
- The Lidgren *et al* study[30] and the AE utility studies identified from the Shabaruddin *et al* review[26] relate to different hypothetical populations; the selected utility estimates for peripheral neuropathy and alopecia do not relate to breast cancer states.
- The available AE utility studies[26] typically use stated preference elicitation techniques (standard gamble or time trade-off), hence both the measurement and valuation of AEs within these studies are from individuals who do not have breast cancer and who have not experienced the AE under consideration. This is not ideal.
- As they are based on comparisons of hypothetical health state scenarios, it is unlikely that the disutilities from the AE utility studies include the possibility of amelioration or resolution of the AE under consideration. It is also unclear how to quantify the distribution of severity of the AEs resulting from chemotherapy within the analysis.

5.2 QALY shortfall analysis

In light of the uncertainties associated with the analysis presented in Section 5.1, the EAG undertook a further analysis which presents the QALY shortfall associated with each test achieving an ICER of £20,000 and £30,000 per QALY gained, based on the deterministic version of the EAG model (see Table 10, Table 11, Table 12, Table 13, Table 14 and Table 15). Other things being equal, this additional analysis may further inform the Appraisal Committee's deliberations around whether other factors which cannot be reliably quantified might have a sufficient impact on the ICERs of the tumour profiling tests to change the interpretation of the model results.

Within each analysis, the QALY shortfall represents the additional number of incremental QALYs that would need to be accrued, given the currently quantified estimates of the incremental QALYs gained for the test and its incremental cost, in order for each test to achieve an ICER at a particular threshold (λ=£20,000 per QALY gained or λ=£30,000 per QALY gained). In health economic terms, this QALY shortfall is equivalent to net clinical benefit. The Committee may find it useful to consider whether the expected magnitude of the health losses avoided by reducing chemotherapy use via tumour profiling tests which are not captured in the EAG model is likely to be equal to or greater than this estimated QALY shortfall. It should be noted that this analysis is predicated on the commentators' assumption that the adverse effects of chemotherapy have been underestimated in the EAG's model. However, the EAG model suggests that with the exception of IHC4+C, all tests increase chemotherapy use at least in

some subgroups (see EAG report, Appendix 7); where this is the case, changing the balance of the net health gains and losses of chemotherapy will produce less favourable ICERs for the tumour profiling tests. It should also be noted that any potential underestimation of QALY losses only apply to those patients who would have received chemotherapy and who would have experienced associated late effects who now do not receive chemotherapy due to the tumour profiling test result and thus avoid these late effects.

The QALY shortfall analysis operates as follows. As shown in Table 10, within the LN0 NPI>3.4 group, Oncotype DX (assuming prognostic benefit only) is estimated to lead to -0.02 QALYs and additional costs of £869 compared with no testing, hence it is expected to be dominated by no testing. In this subgroup, Oncotype DX would need to make up a further 0.06 QALYs in order to achieve an ICER of £20,000 per QALY gained given its incremental cost (£869 / [0.06+-0.02] = £20,000). Within this subgroup, the EAG model suggests that the probability of receiving chemotherapy is reduced by 16% due to the use of Oncotype DX. Assuming that 25% of these patients experience late effects of chemotherapy which are not accounted for within the EAG model, this means that 4% (0.16 x 0.25) of those forgoing chemotherapy will avoid late effects. Given the overall QALY shortfall of 0.06 QALYs and the probability of avoiding late effects of 0.04, this means that each patient who would have experienced a late effect of chemotherapy would have had to have lost 1.49 QALYs (0.06/0.04) due to that AE in order for Oncotype DX to be cost-effective at a threshold of £20,000 per QALY gained.

The results for this analysis are summarised below.

Oncotype DX (prognostic benefit assumed) – refer to Table 10

LN0, NPI≤3.4 – Analysis not relevant as more patients receive chemotherapy in the test group.

LN0, NPI>3.4 – Each patient who avoids chemotherapy and avoids experiencing a late AE not quantified in the EAG model would have to save 1.49 QALYs due to the unquantified AE in order for Oncotype DX to have an ICER of £20,000 per QALY gained. Assuming a threshold of £30,000 per QALY gained, the equivalent value is 1.12 QALYs per patient.

LN+ (1-3 nodes) – Each patient who avoids chemotherapy and avoids experiencing a late AE not quantified in the EAG model would have to save 1.44 QALYs due to the unquantified AE in order for Oncotype DX to have an ICER of £20,000 per QALY gained. Assuming a threshold of £30,000 per QALY gained, the equivalent value is 1.29 QALYs per patient.

Oncotype DX (predictive benefit assumed) – refer to Table 11

LN0, NPI≤3.4 – Analysis not relevant as more patients receive chemotherapy in the test group.

LN0, NPI>3.4 – Analysis not relevant as test dominates.

LN+ (1-3 nodes) – Analysis not relevant as test dominates.

IHC4+C – refer to Table 12

LN0, NPI≤3.4 – Analysis not relevant as ICER already below £20,000 per QALY gained.

LN0, NPI>3.4 – Analysis not relevant as test dominates.

LN+ (1-3 nodes) – Analysis not relevant as test dominates.

Prosigna – refer to Table 13

LN0, NPI≤3.4 – Analysis not relevant test increases chemotherapy use.

LN0, NPI>3.4 – Analysis not relevant test increases chemotherapy use.

LN+ (1-3 nodes) – Analysis not relevant test increases chemotherapy use.

EPClin – refer to Table 14

LN0, NPI≤3.4 – Analysis not relevant test increases chemotherapy use.

LN0, NPI>3.4 – Analysis not relevant test increases chemotherapy use.

LN+ (1-3 nodes) – Analysis not relevant at threshold of £30,000 per QALY gained as ICER is below this. Each patient who avoids chemotherapy and avoids experiencing a late AE not quantified in the EAG model would have to save 0.69 due to the unquantified AE in order for EPClin to have an ICER of £20,000 per QALY gained.

MammaPrint – refer to Table 15

MINDACT ITT - Each patient who avoids chemotherapy and avoids experiencing a late AE not quantified in the EAG model would have to save 2.03 QALYs due to the unquantified AE in order for MammaPrint DX to have an ICER of £20,000 per QALY gained. Assuming a threshold of £30,000 per QALY gained, the equivalent value is 1.23 QALYs per patient.

MINDACT high-risk - Each patient who avoids chemotherapy and avoids experiencing a late AE not quantified in the EAG model would have to save 1.39 QALYs due to the unquantified AE in order for MammaPrint to have an ICER of £20,000 per QALY gained. Assuming a threshold of £30,000 per QALY gained, the equivalent value is 1.11 QALYs per patient.

MINDACT low-risk - Analysis not relevant test increases chemotherapy use.

Table 10: QALY shortfall analysis - Oncotype DX (prognostic benefit only)

Oncotype DX (prognostic)	**LN0, NPI<3.4**	**LN0, NPI>3.4**	**LN+ (1-3 nodes)**
Inc. QALYs	0.01	-0.02	-0.07
Inc. costs	£1,317	£869	£647
ICER	£120,144	Dominated	Dominated
QALY shortfall to achieve ICER=£20,000/QALY gained	**0.05**	**0.06**	**0.10**
QALY shortfall to achieve ICER=£30,000/QALY gained	**0.03**	**0.04**	**0.09**
Proportion patients avoiding chemo due to testing	0.00	0.16	0.29
Proportion patients unaccounted AEs (assumption based on consultation responses)	0.25	0.25	0.25
Proportion patients tested avoiding chemo with unaccounted AEs	n/a - more get chemo in test group	0.04	0.07
QALY loss for patients avoiding chemo with unaccounted AEs required to achieve shortfall at λ=£20,000/QALY	**n/a - more get chemo in test group**	**1.49**	**1.44**
QALY loss for patients avoiding chemo with unaccounted AEs required to achieve shortfall at λ=£30,000/QALY	**n/a - more get chemo in test group**	**1.12**	**1.29**

Table 11: QALY shortfall analysis - Oncotype DX (predictive benefit)

Oncotype DX (predictive)	**LN0, NPI<3.4**	**LN0, NPI>3.4**	**LN+ (1-3 nodes)**
Inc. QALYs	0.04	0.27	0.09
Inc. costs	£1,211	-£364	-£68
ICER	£34,245	Dominating	Dominating
QALY shortfall to achieve ICER=£20,000/QALY gained	**0.03**	**n/a - ICER already below threshold**	**n/a - ICER already below threshold**
QALY shortfall to achieve ICER=£30,000/QALY gained	**0.01**	**n/a - ICER already below threshold**	**n/a - ICER already below threshold**
Proportion patients avoiding chemo due to testing	0.00	0.16	0.29
Proportion patients unaccounted AEs (assumption based on consultation responses)	0.25	0.25	0.25
Proportion patients tested avoiding chemo with unaccounted AEs	n/a - more get chemo in test group	0.04	0.07
QALY loss for patients avoiding chemo with unaccounted AEs required to achieve shortfall at λ=£20,000/QALY	**n/a - more get chemo in test group**	**n/a - ICER already below threshold**	**n/a - ICER already below threshold**
QALY loss for patients avoiding chemo with unaccounted AEs required to achieve shortfall at λ=£30,000/QALY	**n/a - more get chemo in test group**	**n/a - ICER already below threshold**	**n/a - ICER already below threshold**

Table 12: QALY shortfall analysis - IHC4+C

IHC4+C	**LN0, NPI<3.4**	**LN0, NPI>3.4**	**LN+ (1-3 nodes)**
Inc. QALYs	0.01	0.01	0.05
Inc. costs	£22	-£89	-£269
ICER	£2,752	Dominating	Dominating
QALY shortfall to achieve ICER=£20,000/QALY gained	**n/a - ICER already below threshold**	**n/a - ICER already below threshold**	**n/a - ICER already below threshold**
QALY shortfall to achieve ICER=£30,000/QALY gained	**n/a - ICER already below threshold**	**n/a - ICER already below threshold**	**n/a - ICER already below threshold**
Proportion patients avoiding chemo due to testing	0.04	0.08	0.07
Proportion patients unaccounted AEs (assumption based on consultation responses)	0.25	0.25	0.25
Proportion patients tested avoiding chemo with unaccounted AEs	0.01	0.02	0.02
QALY loss for patients avoiding chemo with unaccounted AEs required to achieve shortfall at λ=£20,000/QALY	**n/a - ICER already below threshold**	**n/a - ICER already below threshold**	**n/a - ICER already below threshold**
QALY loss for patients avoiding chemo with unaccounted AEs required to achieve shortfall at λ=£30,000/QALY	**n/a - ICER already below threshold**	**n/a - ICER already below threshold**	**n/a - ICER already below threshold**

Table 13: QALY shortfall analysis - Prosigna

Prosigna	**LN0, NPI<3.4**	**LN0, NPI>3.4**	**LN+ (1-3 nodes)**
Inc. QALYs	0.02	0.07	0.07
Inc. costs	£1,891	£1,713	£1,967
ICER	£89,693	£25,857	£28,666
QALY shortfall to achieve ICER=£20,000/QALY gained	**0.07**	**0.02**	**0.03**
QALY shortfall to achieve ICER=£30,000/QALY gained	**0.04**	**n/a - ICER already below threshold**	**n/a - ICER already below threshold**
Proportion patients avoiding chemo due to testing	0.00	-0.01	-0.08
Proportion patients unaccounted AEs (assumption based on consultation responses)	0.25	0.25	0.25
Proportion patients tested avoiding chemo with unaccounted AEs	n/a - more get chemo in test group	n/a - more get chemo in test group	n/a - more get chemo in test group
QALY loss for patients avoiding chemo with unaccounted AEs required to achieve shortfall at λ=£20,000/QALY	**n/a - more get chemo in test group**	**n/a - more get chemo in test group**	**n/a - more get chemo in test group**
QALY loss for patients avoiding chemo with unaccounted AEs required to achieve shortfall at λ=£30,000/QALY	**n/a - more get chemo in test group**	**n/a - more get chemo in test group**	**n/a - more get chemo in test group**

Table 14: QALY shortfall analysis - EPClin

EPClin	**LN0, NPI<3.4**	**LN0, NPI>3.4**	**LN+ (1-3 nodes)**
Inc. QALYs	0.01	0.03	0.06
Inc. costs	£1,686	£1,401	£1,185
ICER	£141,848	£46,482	£21,489
QALY shortfall to achieve ICER=£20,000/QALY gained	**0.07**	**0.04**	**0.00**
QALY shortfall to achieve ICER=£30,000/QALY gained	**0.04**	**0.02**	**n/a - ICER already below threshold**
Proportion patients avoiding chemo due to testing	-0.07	-0.01	0.02
Proportion patients unaccounted AEs (assumption based on consultation responses)	0.25	0.25	0.25
Proportion patients tested avoiding chemo with unaccounted AEs	n/a - more get chemo in test group	n/a - more get chemo in test group	0.01
QALY loss for patients avoiding chemo with unaccounted AEs required to achieve shortfall at λ=£20,000/QALY	**n/a - more get chemo in test group**	**n/a - more get chemo in test group**	**0.69**
QALY loss for patients avoiding chemo with unaccounted AEs required to achieve shortfall at λ=£30,000/QALY	**n/a - more get chemo in test group**	**n/a - more get chemo in test group**	**n/a - ICER already below threshold**

Table 15: QALY shortfall analysis - MammaPrint

MammaPrint	**MINDACT ITT**	**MINDACT high-risk**	**MINDACT low-risk**
Inc. QALYs	0.01	-0.04	0.01
Inc. costs	£1,757	£1,380	£2,415
ICER	£134,059	Dominated	£399,182
QALY shortfall to achieve ICER=£20,000/QALY gained	**0.07**	**0.11**	**0.11**
QALY shortfall to achieve ICER=£30,000/QALY gained	**0.05**	**0.09**	**0.07**
Proportion patients avoiding chemo due to testing	0.15	0.33	-0.03
Proportion patients unaccounted AEs (assumption based on consultation responses)	0.25	0.25	0.25
Proportion patients tested avoiding chemo with unaccounted AEs	0.04	0.08	n/a - more get chemo in test group
QALY loss for patients avoiding chemo with unaccounted AEs required to achieve shortfall at λ=£20,000/QALY	**2.03**	**1.39**	**n/a - more get chemo in test group**
QALY loss for patients avoiding chemo with unaccounted AEs required to achieve shortfall at λ=£30,000/QALY	**1.23**	**1.11**	**n/a - more get chemo in test group**

6. Probability of having chemotherapy

Several commentators have suggested other potentially relevant decision impact studies could or should have been included in the EAG report. However, the studies suggested are either already included in the EAG report, or were excluded from the report with justification. The only exception to this is a study reported by Rodriguez *et al*; this study was not identified by the EAG searches, however, the results appear to be consistent with other Prosigna decision impact studies already included in the EAG review.

7. EAG systematic review and meta-analysis

All major comments relating to this theme are discussed in the EAG's table of responses.

8. EAG economic model

8.1. Re-analysis of MammaPrint by Agendia within the EAG model

Agendia have undertaken a re-analysis of the cost-effectiveness of MammaPrint using the EAG model *"with corrected usage of available MammaPrint data in those instances where we [Agendia] strongly disagree with the chosen inputs in the current model."* With respect to this analysis, the company claims that on the basis of altered model inputs, the ICER for MammaPrint is now less than £30,000 per QALY gained. However, the EAG notes that within the company's re-analysis, chemotherapy is assumed to be associated with no additional benefit in terms of DRFS for any patient population (including those with clinical-high MammaPrint-high risk). If this was the case, genomic testing would have no value as clinicians would never give chemotherapy to any patient. The EAG considers Agendia's re-analysis of the EAG model to be inappropriate and believes that the results are not meaningful.

8.2. Additional EAG sensitivity analysis - Cost-effectiveness of adjuvant chemotherapy by subgroup

During the consultation on the EAG report and the DCD, it has been suggested that the EAG model is predisposed to find giving chemotherapy to all patients a clinically effective and cost-effective use of resources. This interpretation of the model is inaccurate. In the interests of clarity, Table 16 presents the results of an analysis comparing 100% chemotherapy versus 0% chemotherapy using the EAG model. As shown in the table, the strategy involving the indiscriminate use of chemotherapy is dominated by the no chemotherapy option for patients with NPI≤3.4 (i.e. chemotherapy generates fewer QALYs at a greater cost). Chemotherapy appears to have a favourable clinical and cost-effectiveness profile within the LN0, NPI>3.4 and LN+ subgroups.

Table 16: Cost-effectiveness of chemotherapy versus no chemotherapy

Subgroup	Option	QALYs	Costs	Inc. QALYs	Inc. costs	ICER
LN0, NPI≤3.4	100% chemotherapy	13.83	£7,454	-0.04	£3,670	**Dominated**
	No chemotherapy	13.87	£3,784	-	-	-
LN0, NPI>3.4	100% chemotherapy	12.85	£11,700	0.27	£2,316	**£8,449**
	No chemotherapy	12.58	£9,384	-	-	-
LN+	100% chemotherapy	12.63	£12,668	0.35	£2,011	**£5,787**
	No chemotherapy	12.28	£10,658	-	-	-

8.3. Additional EAG sensitivity analysis - Alternative estimates of chemotherapy benefit within clinical risk subgroups

Several commentators have raised issues regarding the estimated relative risk of distant recurrence associated with chemotherapy. The original EAG report acknowledged that there is uncertainty around this estimate and notes that the estimated relative risk of 0.76 was calculated using the most relevant data reported within the EBCTCG 2011 meta-analysis paper[31] (data specifically relating to distant recurrence). The EAG notes that it is possible that the relative benefit of chemotherapy could be different between clinical risk groups, although the EBCTCG meta-analysis does not provide sufficient information to determine the relative risk of distant recurrence within each of the three model subgroups (LN-, NPI≤3.4; LN- NPI>3.4, and LN+[1-3 nodes]). Tables 139, 142, 145, 148 and 151 of the EAG report presented sensitivity analyses using values of 0.70 and 0.80 to explore the impact of this uncertainty on the cost-effectiveness of the tests; these limits are similar to reported rate ratios for any recurrence (including local and regional) for ER+ patients with N0/N- and N1-3 within the EBCTCG meta-analysis paper.

Within this addendum, the EAG has expanded this existing sensitivity analysis to reflect a broader range of relative risk estimates. As shown in Table 17, the economic conclusions drawn from the model for Oncotype DX, IHC4+C and MammaPrint are unaffected by these alternative values. Conversely, within the scenarios in which chemotherapy is assumed to be less favourable, the ICERs for Prosigna and EPClin are markedly less favourable in the LN0 NPI>3.4 and LN+ subgroups.

Table 17: Additional EAG sensitivity analysis - Alternative estimates of chemotherapy benefit within subgroups

Test	Scenario	ICER (per QALY gained)		
		LN0 NPI≤3.4	**LN0 NPI>3.4**	**LN+**
Oncotype	**Chemotherapy RR = 0.76 (EAG base case)**	**£120,144**	**Dominated**	**Dominated**
	Chemotherapy RR = 0.6	£69,967	Dominated	Dominated
	Chemotherapy RR = 0.7	£94,920	Dominated	Dominated
	Chemotherapy RR = 0.8	£145,102	Dominated	Dominated
	Chemotherapy RR = 0.9	£297,925	£201,602	Dominated
IHC4+C	**Chemotherapy RR = 0.76 (EAG base case)**	**£2,752**	**Dominating**	**Dominating**
	Chemotherapy RR = 0.6	£1,326	Dominating	Dominating
	Chemotherapy RR = 0.7	£2,138	Dominating	Dominating
	Chemotherapy RR = 0.8	£3,223	Dominating	Dominating
	Chemotherapy RR = 0.9	£4,745	Dominating	Dominating
Prosigna	**Chemotherapy RR = 0.76 (EAG base case)**	**£89,693**	**£25,857**	**£28,666**
	Chemotherapy RR = 0.6	£52,504	£13,975	£14,678
	Chemotherapy RR = 0.7	£71,107	£19,926	£21,508
	Chemotherapy RR = 0.8	£107,875	£31,645	£36,018
	Chemotherapy RR = 0.9	£214,907	£65,467	£87,917
EPClin	**Chemotherapy RR = 0.76 (EAG base case)**	**£141,848**	**£46,482**	**£21,489**
	Chemotherapy RR = 0.6	£65,750	£26,202	£11,702
	Chemotherapy RR = 0.7	£99,445	£36,317	£16,663
	Chemotherapy RR = 0.8	£195,508	£56,485	£26,089
	Chemotherapy RR = 0.9	£2,680,967	£116,586	£50,984
MammaPrint	**Scenario**	**MINDACT ITT**	**mAOL High risk**	**mAOL Low risk**
	Chemotherapy RR = 0.76 (EAG base case)	**£134,059**	**Dominated**	**£399,182**
	Chemotherapy RR = 0.6	£176,352	Dominated	£113,124
	Chemotherapy RR = 0.7	£148,424	Dominated	£161,338
	Chemotherapy RR = 0.8	£127,971	Dominated	£276,670
	Chemotherapy RR = 0.9	£112,346	£216,964	£920,361

RR – relative risk

9. Company economic models - new model submitted by Agendia

In response to the diagnostic consultation document, Agendia submitted a revised version of their model based on the MINDACT trial. The EAG has scrutinised this new analysis. The EAG notes that the model trace shows that the proportion of patients remaining alive and recurrence-free increases over time, whilst the proportion of the modelled cohort who are dead is allowed to decrease over time (see Figure 2 and Figure 3); this is clearly incorrect and as such the model lacks any face validity. In addition, whilst the company states that extrapolation has now been included in the model in order to account for longer-term costs and health impacts (assuming a constant event rate), the model trace indicates that no additional events occur between years 7 and 10. This also indicates major programming errors. On the basis of these errors, the EAG does not consider the company's new analyses to be reliable.

Figure 2: Probability of being alive, genomic test group, new Agendia model

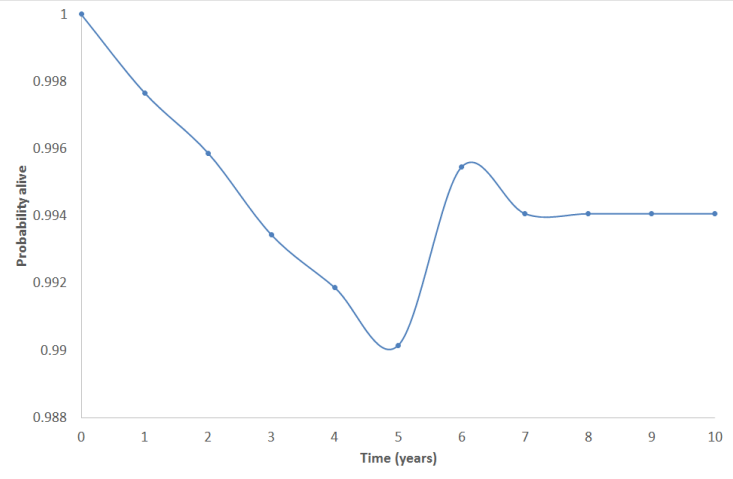
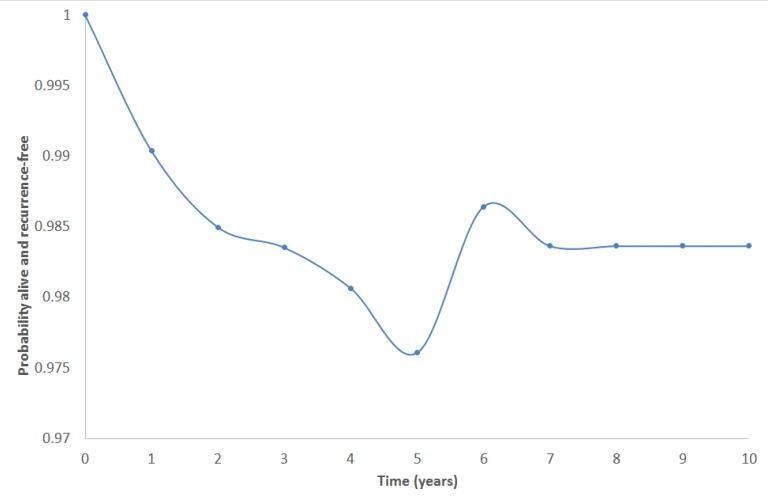


Figure 3: Probability of being alive and recurrence-free, genomic test group, new Agendia model



10. New commercial access schemes

Analyses based on company access proposals are included in a confidential addendum.

REFERENCES

1. National Institute for Health and Care Excellence. Diagnostics Consultation Document: Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer.; 2018.
2. Sestak I, Dowsett M, Cuzick J. NICE request - TransATAC data analysis. In; 2017.
3. Sestak I, Buus R, Cuzick J, et al. Comparison of the performance of 6 prognostic signatures for estrogen receptor–positive breast cancer: A secondary analysis of a randomized clinical trial. *JAMA Oncology* 2018; 10.1001/jamaoncol.2017.5524.

4. Sestak I, Buus R, Cuzick J, Dubsky P, Kronenwett R, Ferree S, *et al.* Comprehensive comparison of prognostic signatures for breast cancer recurrence in TransATAC. *San Antonio Breast Cancer Symposium*, abstract no. 5480.

5. Dowsett M, Cuzick J, Wale C, Forbes J, Mallon EA, Salter J, *et al.* Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: A TransATAC study. *Journal of Clinical Oncology* 2010;28:1829-34.

6. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, *et al.* A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *New England Journal of Medicine* 2004;351:2817-26.

7. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, *et al.* Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *Journal of Clinical Oncology* 2006;24:3726-34.

8. Tang G, Shak S, Paik S, Anderson SJ, Costantino JP, Geyer CE, Jr., *et al.* Comparison of the prognostic and predictive utilities of the 21-gene Recurrence Score assay and Adjuvant! for women with node-negative, ER-positive breast cancer: results from NSABP B-14 and NSABP B-20. *Breast Cancer Res Treat* 2011a;127:133-42.

9. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, *et al.* Prospective validation of a 21-gene expression assay in breast cancer. *New England Journal of Medicine* 2015;373:2005-14.

10. Le Du F, Gonzalez-Angulo AM, Park M, Liu DD, Hortobagyi GN, Ueno NT. Effect of 21-gene RT-PCR assay on adjuvant therapy and outcomes in patients with stage I breast cancer. *Clinical Breast Cancer* 2015;15:458-66.

11. Stemmer SM, Steiner M, Rizel S, Soussan-Gutman L, Geffen DB, Nisenbaum B, *et al.* Real-life analysis evaluating 1594 N0/Nmic breast cancer patients for whom treatment decisions incorporated the 21-gene recurrence score result: 5-year KM estimate for breast cancer specific survival with recurrence score results <30 is >98%. *Cancer Research Conference: 38th Annual CTRC AACR San Antonio Breast Cancer Symposium San Antonio, TX United States Conference Start* 2016;76.

12. Wen HY, Krystel-Whittemore M, Patil S, Pareja F, Bowser ZL, Dickler MN, *et al.* Breast carcinoma with an Oncotype Dx recurrence score <18: Rate of distant metastases in a large series with clinical follow-up. *Cancer* 2017;123:131-7.

13. Petkov VI, Miller DP, Howlader N, Gliner N, Howe W, Schussler N, *et al.* Breast-cancer-specific mortality in patients treated based on the 21-gene assay: a SEER population-based study. *NPJ Breast Cancer* 2016;2:16017.

14. Roberts M, Petkov VI, Miller DP, Shak S, Howlader N, Cronin K, *et al.* Breast cancer specific survival in patients with node-positive hormone receptor positive invasive breast cancer and Oncotype DX recurrence score results in the SEER database. *Journal of Clinical Oncology Conference* 2016;34.

15. Gluz O, Nitz U, Chrlstgen M, Kates RE, Clemens M, Kraemer S, *et al.* Prognostic impact of 21 gene recurrence score, IHC4, and central grade in high-risk HR+/HER2-early breast cancer (EBC): 5-year results of the prospective Phase III WSG PlanB trial. *Journal of Clinical Oncology Conference* 2016a;34.

16. Gluz O, Nitz UA, Christgen M, Kates RE, Shak S, Clemens M, *et al.* West German Study Group Phase III PlanB trial: First prospective outcome data for the 21-gene recurrence score assay and concordance of prognostic markers by central and local pathology assessment. *Journal of Clinical Oncology* 2016b;34:2341-9.

17. Nitz U, Gluz O, Christgen M, Kates RE, Clemens M, Malter W, *et al.* Reducing chemotherapy use in clinically high-risk, genomically low-risk pN0 and pN1 early breast cancer patients: five-

year data from the prospective, randomised phase 3 West German Study Group (WSG) PlanB trial. *Breast Cancer Research and Treatment* 2017; 10.1007/s10549-017-4358-6.

18. Albain K, Barlow W, Shak S, Hortobagyi G, Livingston R, Yeh IT*, et al.* Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. *Lancet Oncology* 2010;11:55-65.

19. Barcenas CH, Raghavendra A, Sinha AK, Syed MP, Hsu L, Patangan MG, Jr.*, et al.* Outcomes in patients with early-stage breast cancer who underwent a 21-gene expression assay. *Cancer* 2017;15:15.

20. Petkov VI, Miller DP, Howlader N, Gliner N, Howe W, Schussler NC*, et al.* Outcome disparities by age and 21-gene recurrence score (RS) in hormone receptor positive (HR+) breast cancer (BC). *Journal of Clinical Oncology Conference* 2016;34.

21. Stemmer S, Steiner M, Rizel S, Soussan-Gutman L, Ben-Baruch N, Bareket-Samish A*, et al.* Clinical outcomes in early-stage breast cancer treated according to the recurrence score results: Evidence from a large prospectively-designed cohort study. In. Academic-In-Confidence; 2017.

22. Stemmer SM, Steiner M, Rizel S, Geffen D, Nisenbaum B, Peretz T*, et al.* First prospectively-designed outcome study in estrogen receptor (ER)+ breast cancer (BC) patients (pts) with N1mi or 1-3 positive nodes in whom treatment decisions in clinical practice incorporated the 21-gene recurrence score (RS) result. *Annals of Oncology Conference: 41st European Society for Medical Oncology Congress, ESMO* 2016;27.

23. Ward S, Scope A, Rafia R, Pandor A, Harnan S, Evans P*, et al.* Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in breast cancer management: a systematic review and cost-effectiveness analysis. *Health Technology Assessment* 2013;17:1-302.

24. National Institute for Health and Care Excellence. Gene expression profiling and expanded immunohistochemistry tests for guiding adjuvant chemotherapy decisions in early breast cancer management: MammaPrint, Oncotype DX, IHC4 and Mammostrat. Diagnostics Guidance 10. London; 2013.

25. Hall P., Smith A, Hulme C, Vargas-Palacios A, Makris A, Hughes-Davies L*, et al.* Value of information analysis of multiparameter tests for chemotherapy in early breast cancer: the OPTIMA-prelim trial. *Value in Health (epub ahead of print)* 2017.

26. Shabaruddin FH, Chen LC, Elliott RA, Payne K. A systematic review of utility values for chemotherapy-related adverse events. *PharmacoEconomics* 2013;31:277-88.

27. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. *Health & Quality of Life Outcomes* 2008;6.

28. Shiroiwa T, Fukuda T, Tsutani K. Health utility scores of colorectal cancer based on societal preference in Japan. *Quality of Life Research* 2009;18:1095-103.

29. Wolff AC, Blackford AL, Visvanathan K, Rugo HS, Moy B, Goldstein LJ. Risk of marrow neoplasms after adjuvant breast cancer therapy: the National Comprehensive Cancer Network experience. *Journal of Clinical Oncology* 2015;33:340-9.

30. Lidgren M, Wilking N, Jonsson B, Rehnberg C. Health related quality of life in different states of breast cancer. *Quality of Life Research* 2007;16:1073-81.

31. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Comparisons between diff erent polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100 000 women in 123 randomised trials. *Lancet* 2012;379:432-44.