This is an open access publisher version of an article that appears in:

BRITISH JOURNAL OF CANCER

The internet address for this paper is:
https://publications.icr.ac.uk/11708/

Please direct all emails to:
publications@icr.ac.uk

Institute of Cancer Research Repository
https://publications.icr.ac.uk

Published text:

S Barton, L Zabaglo, et al (2012), *Assessment of the contribution of the IHC4+C score to decision making in clinical practice in early breast cancer*, **British Journal of Cancer**, Vol. 106(11), 1760-1765

npg

Clinical Studies

Assessment of the contribution of the IHC4 + C score to decision making in clinical practice in early breast cancer

S Barton[1], L Zabaglo[2], R A'Hern[3], N Turner[1,4], T Ferguson[1], S O'Neill[1], M Hills[2], I Smith[1,5] and M Dowsett*[,2,4]

[1]Department of Medicine, Royal Marsden Hospital, Fulham Road, London SW3 6JJ, UK; [2]Academic Department of Biochemistry, Royal Marsden Hospital, Fulham Road, London SW3 6JJ, UK; [3]Cancer Research UK Clinical Trials & Statistics Unit, The Institute of Cancer Research, Sutton, UK; [4]Breakthrough Breast Cancer Centre, Institute of Cancer Research, Fulham Road, London SW3 6JB, UK; [5]Institute of Cancer Research, Fulham Road, London SW3 6JB, UK

BACKGROUND: The immunohistochemical (IHC) 4 + C score is a cost-effective prognostic tool that uses clinicopathologic factors and four standard IHC assays: oestrogen receptor (ER), PR, HER2 and Ki67. We assessed its utility in personalising breast cancer treatment in a clinical practice setting, through comparison with Adjuvant! Online (AoL) and the Nottingham Prognostic Index (NPI).
METHODS: We prospectively gathered clinicopathologic data for postmenopausal patients with hormone receptor-positive, HER2-negative, N0-3 resected early breast cancer treated consecutively at our institution. We retrospectively calculated and compared prognostic scores. The primary endpoint was the proportion of patients reclassified from AoL-defined intermediate-risk by application of the IHC4 + C score.
RESULTS: The median age of the 101 patients included in the analysis was 63. In all, 15 of the 26 patients classified as intermediate-risk by AoL were reallocated to a low-risk group by application of the IHC4 + C score and no patient was reclassified as high-risk group. Of the 59 patients classified as intermediate-risk group by the NPI, 24 were reallocated to a low-risk group and 13 to a high-risk group.
CONCLUSION: IHC4 + C reclassifies more than half of the patients stratified as being in intermediate-risk group by the AoL and NPI. The use of IHC4 + C may substantially improve decision-making on adjuvant chemotherapy.

The majority of patients with early breast cancer are recommended adjuvant systemic therapy after primary surgery to reduce risk of breast cancer recurrence and to increase the likelihood of cure. There is demonstrable benefit in the addition of adjuvant chemotherapy to endocrine therapy in women with oestrogen receptor (ER)-positive early breast cancer Early Breast Cancer Trialists' Collaborative Group (2005). However, patient selection is important because some women may be successfully treated with endocrine therapy alone, and spared the potentially serious and quality of life diminishing toxicities of chemotherapy. The proportion of patients in this group has increased markedly in recent years owing to the earlier diagnosis of breast cancer and the better prognosis disease that is associated with early detection. Endocrine therapy has side effects, but is comparatively well tolerated. The difficulty is in accurately stratifying which women with early ER-positive, HER2-negative breast cancer can be spared from chemotherapy. Combinations of prognostic and predictive clinicopathologic factors have been developed into tools that allocate patients into risk categories with the goal of improved stratification.

The Nottingham Prognostic Index (NPI) is one of the earliest tools to be developed (first presented in 1982, Haybittle et al, 1982; Lee and Ellis, 2008). It was designed through retrospective multivariate regression analysis in a study of 387 women with primary, operable breast cancer and has been independently and prospectively validated (Balslev et al, 1994; D'Eredita et al, 2001; Blamey et al, 2007). It uses tumour size, grade and nodal burden to classify patients into groups as good, moderate and poor prognosis, with NPI <3.4, 3.4–5.4 and >5.4, respectively. The original 15 year overall survival estimates for each prognostic group of 80%, 42% and 13%, respectively, were based on data from the 1980s (Galea et al, 1992). Updated survival estimates from the 1990s have been published together with additional divisions of the prognosis groups (Table 1) (Blamey et al, 2007).

Adjuvant! Online (AoL) is a popular prognostic and predictive tool freely available on the Internet (http://www.adjuvantonline.com). It uses the same factors as the NPI as well as ER status, age and comorbidity to project the likelihood of mortality and disease recurrence at 10 years, and the magnitude of benefit to be gained by adjuvant therapy for individual patients (Ravdin et al, 2001). The prognostic information is predominantly based on data from the Surveillance, Epidemiology and End-Results registry (http://www.seer.cancer.gov), and the projections of efficacy of adjuvant therapy are based on data from the Early Breast Cancer Trialists' Collaborative Group overviews of randomised clinical trials (Early Breast Cancer Trialists' Collaborative Group, 1998a, b, 2005). It has been independently validated using a large, population-based database (Olivotto et al, 2005). Its ease of use and the personalised information it generates, quantitatively estimating prognosis with and without adjuvant therapy, are reasons for its popularity.

Table I NPI 10-year breast cancer-specific survival (1990–1999; Blamey et al, 2007)

Prognostic group	NPI (0.2 × tumour size (cm) + grade + nodal status (pN0 = 1, pN1-3 = 2, pN ⩾ 4 = 3)	10-year survival (%)
Extremely good prognostic group (EPG)	2.08–2.4	96
Good prognostic group (GPG)	2.42–⩽ 3.4	93
Moderate prognostic group I (MPGI)	3.42–⩽ 4.4	81
Moderate prognostic group II (MPGII)	4.42–⩽ 5.4	74
Poor prognostic group (PPG)	5.42–6.4	50
Very Poor prognostic group (VPPG)	6.5–6.8	38

An important recent focus for translational researchers has been towards discovery of molecular markers that better predict individual residual risk after endocrine therapy in patients with ER-positive disease. A number of assays incorporating multiple molecular markers have been developed chiefly for this purpose. The Oncotype Dx or 21-gene recurrence score (RS) (Paik et al, 2004) is a multi-gene assay that tests for overexpression of 21 genes with reverse-transcriptase PCR (Paik et al, 2004). An algorithm is used to calculate the RS and determine risk category (low: < 18, intermediate: ⩾ 18–< 31 or high: ⩾ 31). The RS has been validated in ER-positive breast cancer; in node-negative and node-positive populations, and in tamoxifen-treated, AI-treated and chemotherapy-treated populations (Habel et al, 2006; Paik et al, 2006; Goldstein et al, 2008; Albain et al, 2010; Dowsett et al, 2010). Prospective validation of the predicted benefit from chemotherapy in the higher RS patients is taking place in the TAILORx trial (Zujewski and Kamin, 2008). The almost complete independence of the molecular risk information from RS from the clinicopathologic information encompassed in AoL means that more accurate risk estimates can be made by integrating the two types of data (Goldstein et al, 2008; Tang et al, 2011).

The immunohistochemical (IHC) 4 + C score is a prognostic tool based on quantitative values of four standard laboratory assays (ER, PR, HER2 and Ki67), and the clinicopathologic parameters of tumour grade, size, nodal burden, patient age, and treatment with AI or tamoxifen (hence IHC4 + clinical (IHC4 + C) score) (Cuzick et al, 2011). The IHC4 + C score was developed in a retrospective analysis from TransATAC after recognition that these IHC assays independently hold prognostic power in endocrine-treated patients (Dowsett et al, 2008; Dowsett et al, 2011b). IHC4 + C gives a prediction of the residual risk of distant recurrence at 9 years in postmenopausal women with node-negative, hormone receptor-positive disease treated with 5 years of adjuvant endocrine therapy (Cuzick et al, 2011). The score's performance was found to contain a comparable degree of prognostic information to the RS using the TransATAC data set, and was validated in an independent data set. Further assessment to establish broader applicability was advised (Cuzick et al, 2011).

Like many other institutions, we use AoL to assist in decision-making concerning adjuvant systemic therapy. For women with ER-positive breast cancer with good or intermediate prognosis, we calculate the predicted benefit of chemotherapy in addition to endocrine therapy and use this in making treatment recommendations. The objective of this study was to compare the prognostic information gained from IHC4 + C with that from the AoL and NPI, in a group of postmenopausal women with early breast cancer receiving treatment at our institution, to determine whether assessment of IHC4 + C assists in the selection of patients who may be safely treated with adjuvant hormone therapy alone and spared the side-effects of chemotherapy.

PATIENTS AND METHODS

Patient population

We prospectively recorded post-operative clinicopathologic data for women with early breast cancer consecutively presenting to our single institution during 1 year (February 2010 to February 2011). Patients were included if they had hormone receptor-positive, HER2-negative early breast cancer that was either node-negative or pathological stage N1 (1–3 axillary lymph nodes containing macrometastases (deposit ⩾ 2 mm)), and had undergone surgery to remove the breast tumour, with tumour-free surgical margins. Patients receiving neoadjuvant therapy were excluded. Patients with axillary lymph nodes containing micrometastases were classified as not involved. HER2-negative was defined as being HER2 0 or 1 + on IHC staining or IHC2 +, and either chromogenic or fluorescence in-situ-hybridisation (CISH or FISH) negative (Wolff et al, 2007). Only postmenopausal women were included, because the IHC4-C score has not yet been fully validated in premenopausal women. Postmenopausal was defined as women ⩾ 60 years of age or ⩾ 50, and amenorrhoeic for 12 months in the absence of hysterectomy or intrauterine contraceptive device (IUCD) or endocrine therapy, or women who had undergone bilateral oophorectomy, or in women aged 50–59 whose menopausal status was indeterminate such as because of hysterectomy or IUCD, in whom there was biochemical evidence (elevated FSH and suppressed oestradiol) of menopause. Patient age was recorded at the time when the decision on the patient's adjuvant management was made. Women over the age of 75 were excluded. Women who had incomplete axillary staging or no axillary staging surgery performed were excluded, as were women with bilateral breast cancer, due to anticipated unreliable prognostic estimations of both IHC4 + C and AoL in these settings.

Study design

The study was a retrospective comparison of the AoL, NPI and IHC4 + C score, to determine whether the IHC4 + C contributed to decision-making concerning adjuvant therapies. Clinicopathologic data was collected prospectively over the 1-year period, after which the IHC4 + C score, AoL and the NPI were retrospectively compared. The study was approved by our institution's clinical audit committee.

Study endpoints

The primary endpoint was the proportion of patients reallocated from AoL-defined intermediate risk of distant recurrence at 10 years, to either high or low risk, by application of the IHC4 + C score. The first secondary endpoint was the proportion of patients reallocated from NPI-defined moderate risk to either high or low risk, by application of the IHC4 + C score. Other secondary endpoints were correlation between AoL and IHC4 + C, correlation between the NPI and IHC4 + C, and correlation between the NPI and AoL.

The immunohistochemical (IHC)4 + C

We incorporated standard IHC tests on formalin-fixed paraffin-embedded breast tumour tissue in combination with clinicopathologic factors of tumour grade, size, nodal burden, patient age and AI treatment in calculating the IHC4 + C score for each patient, using the same laboratory methods that were described by the TransATAC research group in deriving the IHC4 other than for Ki67, where the MIB1 antibody was used instead of SP6 after establishing their close similarity (Zabaglo et al, 2010; Cuzick et al, 2011). The ER was quantified by the H-score, which is defined as the percentage of cells showing weak IHC staining, added to two

Clinical Studies

times the percentage of cells staining moderately, added to three times the percentage of cells staining intensely that was then divided by 30 to arrive at variable between 0–10 (ER_{10}). An *H*-score of more than 1 is positive. PgR was quantified by the percentage of cells staining positive, and this was divided by 10 to obtain a variable between 0 and 10 (PgR_{10}). HER2 assessment was given as either a positive (IHC3 + or IHC2 + and CISH or FISH positive) or negative result (IHC0 or 1 + or 2 + and CISH or FISH negative). Ki67 was quantified as the percentage of positively staining cancer cells, using the MIB1 antibody (Zabaglo *et al*, 2010).

The IHC4 score was calculated by the algorithm:

$$94.7 \times \{ -0.100\,ER_{10} - 0.079\,PgR_{10} + 0.586\,HER2 + 0.240\ln(1 + 10 \times Ki67) \}$$

As was described in the published methods for assessing IHC4 in the independent validation cohort (Cuzick *et al*, 2011), we also multiplied Ki67 by 0.4 to scale for the difference in scoring technique. This is because Ki67 scores are on average 2.5 times higher with manual reading than using the image analysis method from which this algorithm was derived.

The clinical score was calculated by the published algorithm:

$$100 \times \{ 0.417 N_{1-3} + 1.566 N_{4+} + 0.930\,(0.497 T_{1-2} + 0.882 T_{2-3} + 1.838 T_{>3} + 0.559 Gr_2 + 0.970 Gr_3 + 0.130 Age_{\geq 65} 0.149 Ana) \},$$

with N_j, T_j, Gr_j and Age_j representing categories of nodal status, tumour size, grade and patient age, respectively, and Ana representing treatment with anastrozole as opposed to tamoxifen. The IHC4 + C score was the sum of the IHC4 and the clinical score.

Adjuvant! Online

We prospectively calculated AoL predictions of mortality and recurrence for each patient and entered this into our database. We collected information regarding patient comorbidity, but owing to the large differences in breast cancer-specific mortality predicted by AoL when conditioned for competing mortality risk and also because neither the IHC4 + C score nor NPI corrects for competing mortality risk, we decided a *priori* to use an assumption of perfect health for all patients. We used 'adjuvant AI for 5 years' in assessing the predicted effectiveness of adjuvant endocrine therapy. We took the prediction of death from breast cancer at 10 years and subtracted the additional benefit conferred by 5 years adjuvant endocrine therapy to calculate the residual risk of breast cancer mortality at 10 years despite adjuvant endocrine therapy.

Nottingham Prognostic Index

The NPI was calculated for each patient postoperatively, and was based on operative pathological findings. The NPI score is $0.2 \times$ tumour size (cm) + grade + nodal status (pN0 = 1, pN1-3 = 2, pN $\geq$ 4 = 3). Patients with an NPI < 3.4 are considered to be at low risk, 3.4–5.4 at intermediate risk and > 5.4 at high risk.

Statistical analysis

The 9-year residual distant recurrence risk after adjuvant endocrine therapy estimate generated by IHC4 + C score was corrected to 10 years (assuming a constant recurrence event rate). The 10-year breast cancer-specific mortality after endocrine therapy that was generated from AoL was multiplied by 1.25, to arrive at a figure representing 10-year residual distant breast cancer recurrence risk after adjuvant endocrine therapy. The multiplier of 1.25 was derived through comparison of distant recurrence and breast cancer-specific survival outcomes in patients treated with 5 years of adjuvant tamoxifen or anastrozole monotherapy in the ATAC study, where the number of distant recurrence events at 10 years was 25% greater than the number of

breast cancer-specific mortality events at 10 years (Cuzick *et al*, 2010). The distant recurrence risk at 10 years after adjuvant endocrine therapy was chosen so the two tools can be directly compared. It also correlates with RS and meant that similar cutoffs between risk groups used could be applied to our results. We used cutoffs of 10 and 20% risk of distant recurrence at 10 years because these correspond with the 18 and 30 RS risk of distant recurrence at 10 years that is used to classify patients as being at intermediate risk, which has been used in previous comparisons (Goldstein *et al*, 2008; Dowsett *et al*, 2010; Tang *et al*, 2011).

The relationship between AoL and IHC4 + C was assessed on a continuous basis graphically and by Spearman's Rank correlation. The Stuart–Maxwell test of marginal homogeneity was used to test for significant differences between the prognostic scores using Stat 11.2 for windows. All patients meeting eligibility criteria were included in these analyses. The planned sample size was 100 patients, which was estimated to be a representative number and attainable within our single institution. It also meant that an odds ratio of approximately three could be detected, assuming there were 40% discordant pairs.

RESULTS

Patient and disease characteristics

Two hundred and four patients were considered potentially eligible for the study on the basis of having hormone receptor-positive, HER2-negative early breast cancer and $\leqslant 3$ axillary lymph nodes containing metastases. In all, 84 patients were excluded due to clinical eligibility factors (5 had bilateral breast cancer, 2 were male, 15 were aged > 75 years, 59 were premenopausal, 3 had unstaged axillary lymph nodes). One hundred and twenty patients met eligibility criteria for the study. In all, 19 patients were not able to be included due to difficulty in obtaining tissue from a different centre or IHC4 + C processing difficulty, leaving a total of 101 evaluable patients. The patient demographic and clinicopathologic characteristics are summarised in Table 2. The median patient age was 63, and the median tumour size was 20 mm.

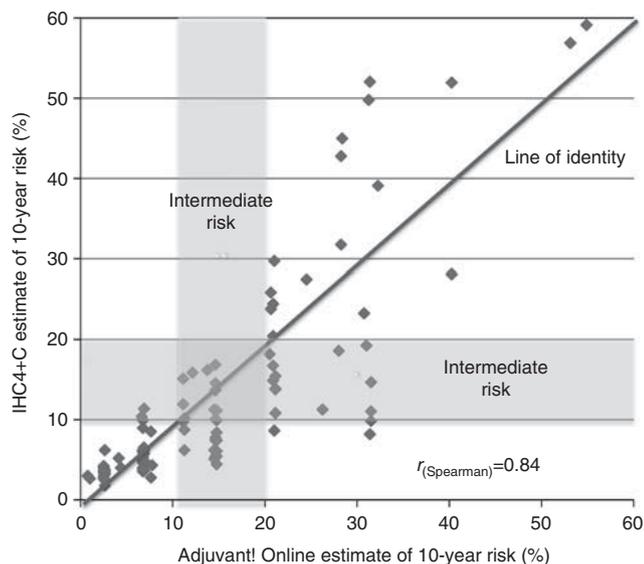### Comparison between AoL score and IHC4 + C score

Figure 1 shows the agreement between AoL and IHC4 + C estimates of risk of distant recurrence at 10 years. The Spearman's Rank correlation was 0.84. Overall AoL rated patients at higher risk than IHC4 + C. Agreement between the categories created for the IHC4 + C score and the AoL score was 68.3% (95% CI: 58.3–77.2)(Table 3). Within the AoL-defined intermediate-risk group, 58% of patients were reclassified by application of the IHC4 + C. Risk was lower by IHC4 + C, with 15/26 patients reclassified as low risk, 11/26 remaining classified as intermediate risk and no patient reclassified as high risk. IHC4 + C also frequently restratified patients from an AoL-defined high-risk group to a low-risk group. Forty-seven percent of the patients in the AoL-defined high-risk group were downgraded; 13/32 to intermediate-risk group and 3/32 to low-risk group by the use of IHC4 + C. In contrast, the risk stratification was found to be highly consistent between the low-risk groups, with 41/43 patients stratified by AoL as low-risk being also stratified as low-risk by IHC4 + C. Only two patients were reclassified from low to intermediate-risk by the use of IHC4 + C, and none were reclassified to high-risk.

Comparison between the NPI and IHC4 + C score

The correlation between the IHC4 + C and NPI score risk groups was 60.4% (95% CI: 50.2–70.0; Table 4). The IHC4 + C restratified the majority of the patients (37/59) in the NPI-defined intermediate-risk group to either a high- or low-risk group. Just over

Table 2 Demographic and clinicopathologic characteristics of the study population

Factor	Number of patients (N = 101)
Age	
50–59	35 (35%)
60–69	50 (49%)
70–75	16 (16%)
Tumour size (mm)	
<20	48 (48%)
20–49	46 (46%)
⩾50	7 (7%)
Tumour grade	
1	15 (15%)
2	62 (61%)
3	24 (24%)
Axillary nodal stage	
N0 (negative)	82 (81%)
N1 (1–3 nodes positive)	19 (19%)
ER status (H score)	
<10	1 (1%)
10–99	8 (8%)
100–199	49 (49%)
⩾200	43 (43%)
PR status (%)	
<5	22 (22%)
5–39	24 (24%)
40–79	25 (25%)
⩾80	30 (30%)
Ki67%	
<10	41 (41%)
10–19	36 (36%)
⩾20	24 (24%)

Figure 1 Spearman's Rank correlation between AoL and IHC4 + C estimates of risk of distant recurrence at 10 years.

one-third of these patients (24/59) were moved from an intermediate to a low-risk group, 13/59 were moved to high-risk and 22/59 remained in the intermediate-risk group. Only 3/38 patients in the NPI low-risk group were reclassified by IHC4 + C as intermediate-risk and none as high-risk. Only four patients were classified as high-risk group by NPI as none of our patients had ⩾4 nodes positive and few had grade 3 tumours, and in these circumstances only a very large tumour gives an NPI of >5.4 (high risk).

Table 3 Comparison of IHC4 + C score with Adjuvant! Online in assessing residual risk of distant recurrence at 10 years after adjuvant endocrine therapy

	IHC4-C		
	Low	**Intermediate**	**High**
Adjuvant! Online			
Low, 0–10.0%	41	2	0
Intermediate, >10.0–20.0%	15	11	0
High, >20.0%	3	12	17

Agreement: 68.3% (95% CI: 58.3–77.2), $P < 0.001$.

Table 4 Comparison of the IHC4 + C score with the NPI in assessing risk of breast cancer distant recurrence

	IHC4-C		
	Low, 0–10.0%	**Intermediate, >10.0–20.0%**	**High, >20.0%**
NPI			
Low (<3.4)	35	3	0
Intermediate (3.4–5.4)	24	22	13
High (>5.4)	0	0	4

Agreement: 60.4% (95% CI: 50.2–70.0), $P < 0.001$.

Table 5 Comparison of Adjuvant! Online with the NPI in assessing risk of distant recurrence

	Adjuvant! Online		
	Low, 0–10.0%	**Intermediate, >10.0–20.0%**	**High, >20.0%**
NPI			
Low (<3.4)	37	1	0
Intermediate (3.4–5.4)	6	25	28
High (>5.4)	0	0	4

Agreement: 65.3% (95% CI: 55.2–74.5), $P < 0.001$.

Correlation between AoL and NPI score

There was moderate correlation between the AoL and NPI at 65.3% (95% CI: 55.2–74.5; Table 5). The predominant category where there was disagreement was in the NPI-defined intermediate-risk group, with 28/55 patients classified as intermediate-risk by NPI being reallocated to a high-risk group by AoL. Only four patients were classified as high-risk group by both NPI and AoL. The greatest agreement was seen in the low-risk group, with 37 patients classified low-risk by both NPI and AoL.

DISCUSSION

We found in an oncology clinical practice setting that IHC4 + C provides additional stratification regarding the residual risk of distant recurrence in ER+ primary breast cancer patients to receive adjuvant endocrine therapy, supplementary to that provided by the AoL and NPI intermediate-risk groups. The IHC4 + C score downgraded more than half of the patients in the AoL-defined intermediate-risk group to a low-risk group. Nearly two-thirds in the NPI-defined intermediate-risk group were reallocated into either a low- or high-risk group, with risk stratification most often lowered. The risk category changes we observed indicate that the use of IHC4 + C in the clinical setting would often lead to a change in adjuvant treatment decisions with an overall increase in the number of patients who may safely be

spared unnecessary adverse effects of chemotherapy. This extra information based on biological phenotype would be valuable for decision-making concerning adjuvant chemotherapy in a substantial proportion of patients.

The correlation between two different tools has been evaluated previously to compare prognostic and predictive efficacy, and to determine whether one tool adds independent information to the other. Our results are similar to that seen when other risk-projecting tools have been equated with another. For instance, comparisons of Oncotype Dx with AoL have demonstrated that these prognostic tools each provide discrete prognostic information, with relatively poor congruency (Goldstein et al, 2008; Dowsett et al, 2010; Tang et al, 2011). In spite of this, experimentation with combining these two prognostic tools into one integrated tool has not yet been shown to be useful, although more information was gained in one study by the integration of common clinicopathologic factors with Oncotype Dx (Tang et al, 2011). The fact that the information contributed by clinicopathologic factors was complementary to that obtained by molecular analysis alone is relevant to the present study, as the IHC4 + C also combines surrogate information via IHC about the individual molecular profile of the tumour, and combines it with common clinicopathologic factors. The prognostic information provided by IHC4 + C on top of that from clinicopathologic features has previously been shown to be at least as great as the additional information provided by Oncotype Dx (Cuzick et al, 2011).

Our method of deriving predicted 10-year breast cancer-specific survival from AoL is different from that used by two other groups of authors (Dowsett et al, 2010; Tang et al, 2011). They used the AoL estimate of breast cancer-specific mortality without competing risks at 10 years follow-up (shown as '10-year risk'), which is not corrected for comorbidity. The mortality projections made by AoL are otherwise markedly affected by comorbidity (Ozanne et al, 2009). However, we accounted for the discrepancy by using 'perfect health' as the comorbidity status of all patients. Therefore, the only varying difference between the '10-year risk' figure and our estimate of breast cancer-specific mortality at 10 years is a small adjustment that AoL makes according to each patient's age. We observed that the 10-year breast cancer-specific survival estimates we calculated were very similar to the '10-year risk'. Our method of adjusting the AoL breast cancer-specific survival estimate into a distant recurrence risk estimate, using data from the ATAC study to calculate a simple multiplication factor, was a pragmatic approach to enabling direct comparison of the prognostic tools. We acknowledge that there may be minor variance between true AoL risk estimate and the AoL risk estimate we generated as a consequence of using this method, but would anticipate this to be relatively trivial.

A major advantage of IHC4 + C is its cost-effectiveness, being considerably less expensive than gene expression profiling tools such as Oncotype Dx. Another major advantage of IHC4 + C over Oncotype Dx is that it uses existing laboratory assays, and in principle it could therefore be performed at the majority of oncology clinical centres internationally. On the other hand, there are quality assurance issues with qualitative assessment of ER, PR, HER2 and Ki67% IHC, with the potential for interlaboratory variation in values. Although international groups are making efforts to standardise these assays (Wolff et al, 2007; Hammond et al, 2010), there remains justifiable concern. Ki67% in particular has caused apprehension due to variable methods of assessment and heterogeneity in results between different laboratories. However, following a focused working group meeting in which investigators with expertise in assessment of Ki67% made recommendations aimed towards standardising Ki67 methodology and reducing interlaboratory variability; the Ki67% assay results are likely to become better synchronised between laboratories (Dowsett et al, 2011a). The IHC4 + C score has been validated using a separate data set with IHC assays that were evaluated in a different laboratory.

However, for the IHC4 + C score to be validated comprehensively to ensure its reliability and reproducibility, it would need to be tested more widely, with assays conducted in several different laboratories (Cuzick et al, 2011). Our study adds to the literature with further evidence of Ki67 being reliably assayed with the results being utilised successfully in clinical practice.

Our results show that the IHC4 + C score does not appear to be additionally useful in patients classified by AoL or the NPI to be of low-risk, as there was high correlation between the two scores in these categories, and no patients were moved from AoL or NPI-defined low-risk to a high-risk group by IHC4 + C. IHC4 + C holds predictive information as well as prognostic, as by estimating residual risk after endocrine therapy it is partially predicting endocrine therapy benefit. However, even those patients who fail to gain substantial benefit from endocrine therapy are likely to remain at relatively low risk of recurrence, such that adjuvant chemotherapy risk would probably outweigh benefit. These results indicate that patients assessed as low-risk by AoL or NPI are unlikely to benefit from the supplementary prognostic information provided by IHC4 + C and should be recommended adjuvant endocrine therapy alone.

Judging by our findings, it is possible that IHC4 + C has clinical utility in the AoL high-risk group. IHC4 + C downgraded just less than half of the patients from a high-risk stratification by AoL to intermediate-risk or even, in a few, to a low-risk stratification. IHC4 correlates reasonably closely with Oncotype Dx (72%) (Cuzick et al, 2011), and for this reason it is likely to also be predictive of the magnitude of benefit from chemotherapy in the same way that Oncotype Dx is demonstrated to be (Paik et al, 2006; Albain et al, 2010). The additional information regarding risk stratification provided by IHC4 + C for patients at AoL-defined high-risk, but intermediate or low-risk by IHC4 + C, may be beneficial by predicting less chemosensitivity and/or greater benefit from endocrine therapy, which could contribute to decisions on adjuvant therapy.

It appears that AoL may overestimate risk compared with the IHC4 + C and NPI. A similar phenomenon was noticed in a comparison of AoL with Oncotype Dx (Dowsett et al, 2010). A reason why AoL may predict a higher level of risk compared with IHC4 + C is that it is based on clinical data that was mainly collected in the era preceding breast screening, and may disproportionately convey risk statistics associated with breast cancers that are biologically more aggressive, whereas IHC4 + C is based on outcomes from a low-risk breast cancer population.

There was a low frequency of patients stratified as high-risk by NPI compared with AoL and IHC4 + C. This is likely to be because our study population was confined to patients with ≤3 lymph nodes, and hormone receptor-positive disease (which is less often grade 3 than triple-negative breast cancer). In this context, tumour size must be large for NPI risk category to be high. IHC-4 and AoL are influenced by a greater number of variables, resulting in wider division of risk group allocation in this low- and intermediate-risk hormone receptor-positive population.

In conclusion, we found that IHC4 + C assisted with risk stratification in the challenging group of patients with early breast cancer classified as being of intermediate-risk by the AoL and NPI. More than half of the patients classified as being of intermediate-risk by the AoL and NPI were reclassified by IHC4 + C. This study is important for providing evidence that IHC4 + C may be useful in guiding decision-making regarding adjuvant chemotherapy in an oncology clinical practice setting.

REFERENCES

Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, Ravdin P, Bugarini R, Baehner FL, Davidson NE, Sledge GW, Winer EP, Hudis C, Ingle JN, Perez EA, Pritchard KI, Shepherd L, Gralow JR, Yoshizawa C, Allred DC, Osborne CK, Hayes DF (2010) Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. *Lancet Oncol* 11: 55–65

Balslev I, Axelsson CK, Zedeler K, Rasmussen BB, Carstensen B, Mouridsen HT (1994) The Nottingham Prognostic Index applied to 9149 patients from the studies of the Danish Breast Cancer Cooperative Group (DBCG). *Breast Cancer Res Treat* 32: 281–290

Blamey RW, Ellis IO, Pinder SE, Lee AH, Macmillan RD, Morgan DA, Robertson JF, Mitchell MJ, Ball GR, Haybittle JL, Elston CW (2007) Survival of invasive breast cancer according to the Nottingham Prognostic Index in cases diagnosed in 1990–1999. *Eur J Cancer* 43: 1548–1555

Cuzick J, Dowsett M, Pineda S, Wale C, Salter J, Quinn E, Zabaglo L, Mallon E, Green AR, Ellis IO, Howell A, Buzdar AU, Forbes JF (2011) Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the genomic health recurrence score in early breast cancer. *J Clin Oncol* 29: 4273–4278

Cuzick J, Sestak I, Baum M, Buzdar A, Howell A, Dowsett M, Forbes JF (2010) Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 10-year analysis of the ATAC trial. *Lancet Oncol* 11: 1135–1141

D'Eredita G, Giardina C, Martellotta M, Natale T, Ferrarese F (2001) Prognostic factors in breast cancer: the predictive value of the Nottingham Prognostic Index in patients with a long-term follow-up that were treated in a single institution. *Eur J Cancer* 37: 591–596

Dowsett M, Allred C, Knox J, Quinn E, Salter J, Wale C, Cuzick J, Houghton J, Williams N, Mallon E, Bishop H, Ellis I, Larsimont D, Sasano H, Carder P, Cussac AL, Knox F, Speirs V, Forbes J, Buzdar A (2008) Relationship between quantitative estrogen and progesterone receptor expression and human epidermal growth factor receptor 2 (HER-2) status with recurrence in the arimidex, tamoxifen, alone or in combination trial. *J Clin Oncol* 26: 1059–1065

Dowsett M, Cuzick J, Wale C, Forbes J, Mallon EA, Salter J, Quinn E, Dunbier A, Baum M, Buzdar A, Howell A, Bugarini R, Baehner FL, Shak S (2010) 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. *J Clin Oncol* 28: 1829–1834

Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, Ellis M, Henry NL, Hugh JC, Lively T, McShane L, Paik S, Penault-Llorca F, Prudkin L, Regan M, Salter J, Sotiriou C, Smith IE, Viale G, Zujewski JA, Hayes DF (2011a) Assessment of Ki67 in breast cancer: recommendations from the international Ki67 in breast cancer working group. *J Natl Cancer Inst* 103: 1656–1664

Dowsett M, Salter J, Zabaglo L, Mallon E, Howell A, Buzdar AU, Forbes J, Pineda S, Cuzick J (2011b) Predictive algorithms for adjuvant therapy: TransATAC. *Steroids* 76: 777–780

Early Breast Cancer Trialists' Collaborative Group (1998a) Polychemotherapy for early breast cancer: an overview of the randomised trials. *Lancet* 352: 930–942

Early Breast Cancer Trialists' Collaborative Group (1998b) Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 351: 1451–1467

Early Breast Cancer Trialists' Collaborative Group (2005) Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 365: 1687–1717

Galea MH, Blamey RW, Elston CE, Ellis IO (1992) The Nottingham Prognostic Index in primary breast cancer. *Breast Cancer Res Treat* 22: 207–219

Goldstein LJ, Gray R, Badve S, Childs BH, Yoshizawa C, Rowley S, Shak S, Baehner FL, Ravdin PM, Davidson NE, Sledge Jr GW, Perez EA, Shulman LN, Martino S, Sparano JA (2008) Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features. *J Clin Oncol* 26: 4063–4071

Habel LA, Shak S, Jacobs MK, Capra A, Alexander C, Pho M, Baker J, Walker M, Watson D, Hackett J, Blick NT, Greenberg D, Fehrenbacher L, Langholz B, Quesenberry CP (2006) A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients. *Breast Cancer Res* 8: R25

Hammond ME, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, Fitzgibbons PL, Francis G, Goldstein NS, Hayes M, Hicks DG, Lester S, Love R, Mangu PB, McShane L, Miller K, Osborne CK, Paik S, Perlmutter J, Rhodes A, Sasano H, Schwartz JN, Sweep FC, Taube S, Torlakovic EE, Valenstein P, Viale G, Visscher D, Wheeler T, Williams RB, Wittliff JL, Wolff AC (2010) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Oncol* 28: 2784–2795

Haybittle JL, Blamey RW, Elston CW, Johnson J, Doyle PJ, Campbell FC, Nicholson RI, Griffiths K (1982) A prognostic index in primary breast cancer. *Br J Cancer* 45: 361–366

Lee AH, Ellis IO (2008) The Nottingham Prognostic Index for invasive carcinoma of the breast. *Pathol Oncol Res* 14: 113–115

Olivotto IA, Bajdik CD, Ravdin PM, Speers CH, Coldman AJ, Norris BD, Davis GJ, Chia SK, Gelmon KA (2005) Population-based validation of the prognostic model ADJUVANT! for early breast cancer. *J Clin Oncol* 23: 2716–2725

Ozanne EM, Braithwaite D, Sepucha K, Moore D, Esserman L, Belkora J (2009) Sensitivity to input variability of the Adjuvant! Online breast cancer prognostic model. *J Clin Oncol* 27: 214–219

Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 351: 2817–2826

Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, Cronin M, Baehner FL, Watson D, Bryant J, Costantino JP, Geyer Jr. CE, Wickerham DL, Wolmark N (2006) Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 24: 3726–3734

Ravdin PM, Siminoff LA, Davis GJ, Mercer MB, Hewlett J, Gerson N, Parker HL (2001) Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 19: 980–991

Tang G, Shak S, Paik S, Anderson SJ, Costantino JP, Geyer Jr. CE, Mamounas EP, Wickerham DL, Wolmark N (2011) 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* 127: 133–142

Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, McShane LM, Paik S, Pegram MD, Perez EA, Press MF, Rhodes A, Sturgeon C, Taube SE, Tubbs R, Vance GH, van de Vijver M, Wheeler TM, Hayes DF (2007) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* 25: 118–145

Zabaglo L, Salter J, Anderson H, Quinn E, Hills M, Detre S, A'Hern R, Dowsett M (2010) Comparative validation of the SP6 antibody to Ki67 in breast cancer. *J Clin Pathol* 63: 800–804

Zujewski JA, Kamin L (2008) Trial assessing individualized options for treatment for breast cancer: the TAILORx trial. *Future Oncol* 4: 603–610

Clinical Studies