Validity of the online PREDICT tool in older patients with breast cancer: a population-based study

N A de Glas[1,2], E Bastiaannet[1,2], C C Engels[1], A J M de Craen[2], H Putter[3], C J H van de Velde[1], A Hurria[4], G J Liefers[1] and J E A Portielje*,[5]

[1]Department of Surgery, Leiden University Medical Centre, PO Box 9600, 2300RC Leiden, The Netherlands; [2]Department of Gerontology and Geriatrics, Leiden University Medical Centre, PO Box 9600, 2300RC Leiden, The Netherlands; [3]Department of Medical Statistics, Leiden University Medical Centre, PO Box 9600, 2300RC Leiden, The Netherlands; [4]Cancer and Ageing Research Program, City of Hope, 1500 E Duarte Road, Duarte, CA 91010, USA and [5]Department of Medical Oncology, Haga Hospital The Hague, Leyweg 275, 2545 CH Den Haag, The Netherlands

Background: Predicting breast cancer outcome in older patients is challenging, as it has been shown that the available tools are not accurate in older patients. The PREDICT tool may serve as an alternative tool, as it was developed in a cohort that included almost 1800 women aged 65 years or over. The aim of this study was to assess the validity of the online PREDICT tool in a population-based cohort of unselected older patients with breast cancer.

Methods: Patients were included from the population-based FOCUS-cohort. Observed 5- and 10-year overall survival were estimated using the Kaplan–Meier method, and compared with predicted outcomes. Calibration was tested by composing calibration plots and Poisson Regression. Discriminatory accuracy was assessed by composing receiver-operator-curves and corresponding c-indices.

Results: In all 2012 included patients, observed and predicted overall survival differed by 1.7%, 95% confidence interval (CI) = − 0.3–3.7, for 5-year overall survival, and 4.5%, 95% CI = 2.3–6.6, for 10-year overall survival. Poisson regression showed that 5-year overall survival did not significantly differ from the ideal line (standardised mortality ratio (SMR) = 1.07, 95% CI = 0.98–1.16, $P = 0.133$), but 10-year overall survival was significantly different from the perfect calibration (SMR = 1.12, 95% CI = 1.05–1.20, $P = 0.0004$). The c-index for 5-year overall survival was 0.73, 95% CI = 0.70–0.75, and 0.74, 95% CI = 0.72–0.76, for 10-year overall survival.

Conclusions: PREDICT can accurately predict 5-year overall survival in older patients with breast cancer. Ten-year predicted overall survival was, however, slightly overestimated.

Adjuvant systemic treatment in early breast cancer is aimed at reducing the risk of distance recurrence and breast cancer mortality (Early Breast Cancer Trialists' Collaborative Group (EBCTCG), 2005). There are several prediction tools available that can aid in the decision to treat patients with systemic therapy, based on clinical and pathological factors (Engelhardt *et al*, 2013). Prediction tools that can aid in individualising treatment decisions

are especially important for older patients, as older patients comprise a heterogeneous group because of differences in comorbidity and functional status (Braithwaite *et al*, 2010; Barnett *et al*, 2012). Currently, the Dutch guideline specifically advices to use the prediction tool 'Adjuvant! Online' in older patients with early breast cancer (NABON, 2012). Furthermore, the guideline of the International Society for Geriatric Oncology

states that although not specifically validated in older patients, Adjuvant! Online can be used older patients with breast cancer to weigh the risks and benefits of adjuvant treatment (Wildiers *et al*, 2007). Adjuvant! Online is a tool that predicts absolute 10-year overall survival, breast cancer survival and recurrence, and provides clinicians with the expected benefits of endocrine therapy and chemotherapy (Ravdin *et al*, 2001). It is the most well-known and most widely used clinical prediction model available (Ravdin *et al*, 2001; Engelhardt *et al*, 2014).

Recently, we demonstrated that Adjuvant! Online does not accurately predict overall survival and recurrence in older women with early breast cancer (de Glas *et al*, 2014c). This was most likely explained by the fact that the model was developed in a cohort of women with a maximum age of 69 years (Ravdin *et al*, 2001). Therefore, alternative models that can accurately predict breast cancer outcome in older patients are required, especially as the number of older patients with breast cancer is rapidly increasing because of the ageing of Western societies (DeSantis *et al*, 2011, 2013; Biganzoli *et al*, 2012).

The more recently developed PREDICT tool may serve as an alternative prediction tool for older patients with breast cancer, as the patient cohort in which it was developed included almost 1800 women aged 65 years or over (Wishart *et al*, 2010). PREDICT calculates expected overall survival at 5 and 10 years, based on several patient and tumour characteristics(Wishart *et al*, 2010, 2012). In addition, the model provides expected benefits of chemotherapy, endocrine treatment, and trastuzumab. The model was developed in a population-based cohort in the United Kingdom, and validated in two separate cohorts (Wishart *et al*, 2010, 2011). It includes all variables that are also incorporated in the Adjuvant! Online model except for comorbidity status, but it additionally uses detection-mode (screen-detected or interval cancer), and recently HER2 status and KI67 status were added to the model (Wishart *et al*, 2012; Eastern Cancer Registry and Information Centre and Cambridge University, 2015).

The aim of this study was to assess the validity of the online PREDICT tool by studying its predictions for 5- and 10-year overall survival, using a population-based cohort of unselected older patients with breast cancer.

MATERIALS AND METHODS

Patients. For this study, we used data from the population-based FOCUS-cohort (Female breast cancer in the elderly; Optimizing Clinical guidelines USing clinico-pathological and molecular data; de Glas *et al*, 2013; Kiderlen *et al*, 2013; de Glas *et al*, 2014c). All consecutive patients with breast cancer aged 65 years and older in the geographically defined region West of the Comprehensive Cancer Center, the Netherlands, between 1997 and 2004 were included in this cohort ($N = 3672$). Information on patient- and tumour characteristics, detection-mode, treatments, adverse events, and recurrences was registered from medical charts. Follow-up on survival status was available until 31 December 2013 through linkage of patients with municipal population registries.

For this study, all patients with unilateral, unicentric, invasive, local adenocarcinoma were included, provided that they received adequate locoregional treatment (either mastectomy or lumpectomy with radiotherapy and axillary lymph node staging) and they did not receive neoadjuvant treatment.

Immunohistochemistry for Ki67 and HER-2 status. All tumour material of patients who were included was collected. Stainings for Ki67 and HER-2 status were performed according to the standard protocols (de Kruijf *et al*, 2010). HER-2-positive breast cancer as $2+$ or $3+$ by immunohistochemistry, whereas 0–1 were defined as HER-2-negative breast cancer. FISH staining was not available.

Statistical analyses. All patients were manually entered into the online PREDICT tool via the PREDICT website (Eastern Cancer Registry and Information Centre and Cambridge University (2015)). If information on one of the variables was missing, patients were not excluded but the 'unknown' category was used. Predicted 5-year and 10-year survival outcomes were registered, taking the treatment that patients actually received into account.

All analyses were performed in IBM SPSS Statistics version 20.0 (Armonk, NY, USA) and R version 3.1.0 (The R Foundation, Vienna, Austria). A *P*-value of 0.05 was defined as statistically significant.

First, we calculated 5- and 10-year observed overall survival for all subcategories of patients using the Kaplan–Meier method. These observed outcomes were compared with predicted 5- and 10-year survival outcomes of patients using one-sampled *T*-tests, with predicted outcomes as 'fixed' value and observed outcomes as the evaluated variable. These analyses were performed for the whole cohort as well as for all subgroups that were defined by PREDICT. In addition, we calculated observed and predicted outcomes for subgroups based on number of comorbidities according to the ICD-10 classification (World Health Organisation, 1992; 0, 1, 2, 3, and 4 or more) and treatment characteristics (most extensive surgery and systemic treatment).

Next, we assessed the calibration of the model by plotting the observed and predicted 5-year and 10-year overall survival outcomes. For this, the cohort was divided into 10% intervals of the predicted values by PREDICT. Conform methods that were used in previous studies (Olivotto *et al*, 2005; Mook *et al*, 2009; de Glas *et al*, 2014c), intervals that contained less than 100 patients were combined. Calibration was assessed by comparing the calibration plot with the 'ideal line' ($x = y$), using Poisson Regression (van Houwelingen and Putter, 2012).

Finally, discrimination was assessed by composing receiver-operator curves (ROC) and calculating corresponding c-indices. These are calculated by determining the area under the curve (AUC) of the ROC. A c-index can be interpreted as follows: a c-index of 1.0 means that the model has perfect discrimination, whereas a c-index of 0.5 indicates that the model predicts just as well as flipping a coin (Akobeng, 2007).

RESULTS

Overall, 2012 patients from the FOCUS-cohort fulfilled the inclusion criteria and were included. The inclusion flow diagram has been previously published (de Glas *et al*, 2014c). Patient characteristics are presented in Table 1. Median age was 75.0 years (inter quartile range = 69.0–79.0). The majority of patients (1527 out of 2012 patients, 75.9%) had one or more comorbidities.

Observed 5-year overall survival was 74.2% (s.e. 1.0). Predicted 5-year overall survival was 75.9%, which means that the observed and predicted 5-year overall survival differed by 1.7%, 95% confidence interval (CI) = − 0.3 to 3.7, $P = 0.089$ (Table 1). The difference between predicted and observed 5-year overall survival was largest in the oldest patients (11.3%, 95% CI = 3.8–18.8, $P = 0.003$ in patients ⩾85 years). Overall survival was *underestimated* in patients without comorbidity (predicted: observed overall survival = − 3.7%, 95% CI = − 7.2 to − 0.2, $P = 0.040$), and *overestimated* in patients with ⩾4 comorbidities (predicted: observed overall survival 11.8%, 95% CI = 6.9–16.7, $P < 0.0001$). In addition, 5-year overall survival was *underestimated* in patients who received breast-conserving surgery (predicted: observed overall survival = − 3.2%, 95% CI = − 5.6 to − 0.8, $P = 0.008$), and *overestimated* in patients who received mastectomy as most extensive surgery (predicted: observed 5-year overall survival 4.9%, 95% CI = 2.4–7.5, $P = 0.0002$). Finally, PREDICT *overestimated* 5-year overall survival in patients who received endocrine therapy

Table 1. Baseline characteristics and observed and expected overall 5- and 10-year overall survival

		5-Year overall survival			10-Year overall survival		
	n (%)	PREDICT predicted	Observed (s.e.)	Predicted observed (95% CI)	PREDICT predicted	Observed (s.e.)	Predicted-observed (95% CI)
All patients	2012 (100)	75.9	74.2 (1.0)	1.7 (−0.3 to 3.7)	57.9	53.4 (1.1)	4.5 (2.3 to 6.6)
Age (years)							
65–69	560 (27.8)	85.1	86.0 (1.5)	−0.9 (−3.8 to 2.0)	72.2	73.0 (1.9)	−0.8 (−4.5 to 2.9)
70–74	544 (27.0)	81.1	83.0 (1.6)	−1.9 (−5.0 to1.2)	65.6	66.6 (2.0)	−1.0 (−4.9 to 2.9)
75–79	417 (20.7)	75.5	70.3 (2.2)	5.2 (0.9 to 9.5)	55.5	45.2 (2.5)	10.3 (5.4 to 15.2)
80–84	322 (16.0)	61.2	62.0 (2.7)	−0.8 (−6.1 to 4.5)	41.1	29.2 (2.6)	11.9 (6.8 to17.0)
⩾85	169 (8.4)	50.0	38.7 (3.8)	11.3 (3.8 to 18.8)	24.1	12.1 (2.6)	12.0 (6.9 to 17.1)
Number of comorbidities							
0	485 (24.1)	77.1	80.8 (1.8)	−3.7 (−7.2 to −0.2)	59.9	65.0 (2.2)	−5.1 (−9.4 to 0.8)
1	476 (23.7)	76.4	79.4 (1.9)	−3.0 (−6.7 to 0.7)	58.9	61.5 (2.2)	−2.6 (−6.9 to 1.7)
2	398 (19.8)	74.8	72.4 (2.2)	2.4 (−1.9 to 6.7)	56.8	50.0 (2.5)	6.8 (1.9 to 11.7)
3	275 (13.7)	76.4	71.6 (2.7)	4.8 (−0.5 to 10.2)	57.9	49.9 (3.0)	8.0 (1.7 to 14.3)
⩾4	378 (18.8)	74.5	62.7 (2.5)	11.8 (6.9 to 16.7)	55.4	34.7 (2.5)	20.7 (15.8 to 25.6)
Mode of detection							
Symptomatic	1040 (51.7)	71.2	67.7 (1.5)	3.5 (0.6 to 6.4)	51.2	44.4 (1.6)	6.8 (3.7 to 9.9)
Screen-detected	586 (29.1)	85.9	89.6 (1.3)	−3.7 (−6.3 to −1.1)	73.0	74.1 (1.8)	−1.1 (−4.6 to 2.4)
Unknown	386 (19.2)	73.2	68.1 (2.4)	5.1 (0.4 to 9.8)	53.2	46.4 (2.6)	6.8 (1.7 to 11.9)
Tumour size (cm)							
0.1–1.0	299 (14.9)	85.0	87.0 (1.9)	−2.0 (−5.7 to 1.7)	71.8	72.7 (2.6)	−0.9 (−6.0 to 4.2)
1.1–2.0	762 (37.9)	79.3	79.9 (1.5)	−0.6 (−3.5 to 2.3)	62.8	48.9 (1.8)	13.9 (10.4 to 17.4)
2.1–3.0	540 (26.8)	72.3	67.4 (2.0)	4.9 (1.0 to 8.8)	52.1	45.3 (2.2)	6.8 (2.5 to 11.1)
3.1–5.0	265 (13.2)	65.2	60.0 (3.0)	5.2 (−0.7 to 11.1)	43.3	37.0 (3.0)	6.3 (0.4 to 12.2)
>5	60 (3.0)	58.6	53.3 (6.4)	5.3 (−7.5 to 18.1)	34.4	32.6 (6.1)	1.8 (−10.4 to 14.0)
Unknown	86 (4.3)	80.9	80.1 (4.3)	0.8 (−7.8 to 9.4)	64.3	54.0 (5.4)	10.3 (−0.4 to 21.0)
Tumour grade							
1	293 (14.6)	82.9	78.8 (2.4)	4.1 (−0.6 to 8.8)	68.4	57.1 (2.9)	11.3 (5.6 to 17.0)
2	682 (33.9)	78.3	76.4 (1.6)	1.9 (−1.2 to 5.0)	60.4	54.7 (1.9)	5.7 (2.0 to 9.4)
3	459 (22.8)	66.4	67.1 (2.2)	−0.7 (−5.0 to 3.6)	46.3	47.3 (2.3)	−1.0 (−5.5 to −3.5)
Unknown	578 (28.7)	77.0	74.8 (1.8)	2.2 (−1.4 to 5.7)	58.9	54.9 (2.1)	4.0 (−0.1 to 8.1)
Positive nodes (n)							
0	553 (27.5)	80.5	77.2 (1.8)	3.3 (−0.2 to 6.8)	64.8	56.1 (2.1)	8.7 (4.6 to 12.8)
1–3	213 (10.6)	71.0	65.1 (3.2)	5.9 (−0.4 to 12.2)	50.2	39.8 (3.3)	10.4 (3.9 to 16.9)
4–9	93 (4.6)	63.8	52.7 (5.2)	11.1 (0.8 to 21.4)	41.2	24.4 (4.5)	16.8 (7.9 to 25.7)
>9	51 (2.5)	53.2	47.1 (7.0)	6.1 (−8.0 to 20.2)	26.8	29.3 (6.4)	−2.5 (−15.4 to 10.4)
Unknown	1102 (54.8)	76.6	77.5 (1.3)	−0.9 (−3.5 to 1.7)	58.8	58.3 (1.5)	0.5 (−2.4 to 3.4)
Oestrogen receptor status							
Negative	370 (18.4)	62.2	65.0 (2.5)	−2.8 (−7.7 to 2.1)	47.6	48.5 (2.6)	−0.9 (−6.0 to 4.2)
Positive	1404 (69.8)	78.7	75.8 (1.1)	2.9 (0.7 to 5.1)	59.8	53.3 (1.3)	6.5 (4.0 to 0.1)
Unknown	238 (11.8)	80.8	79.0 (2.6)	1.8 (−3.3 to 6.9)	62.8	61.7 (3.2)	1.1 (−5.2 to 7.4)
HER2 status							
Negative	942 (46.8)	76.5	75.6 (1.4)	0.9 (−1.8 to 3.6)	58.3	52.2 (1.6)	6.1 (3.0 to 9.2)
Positive	258 (12.8)	70.0	70.1 (2.9)	−0.1 (−5.8 to 5.6)	51.8	50.3 (3.1)	1.5 (−4.6 to 7.6)
Unknown	812 (40.4)	77.0	75.0 (1.5)	2.0 (−0.9 to 4.9)	59.4	55.8 (1.8)	3.6 (0.1 to 7.1)
KI67 status							
Negative (<10%)	1021 (50.7)	76.0	73.7 (1.4)	2.3 (−0.4 to 5.0)	57.8	52.3 (1.6)	5.5 (2.4 to 8.6)
Positive (⩾10%)	103 (5.1)	68.2	67.9 (4.6)	0.3 (−8.8 to 9.4)	50.6	50.1 (5.0)	0.5 (−9.4 to 10.4)
Unknown	888 (44.1)	76.7	75.4 (1.4)	1.3 (−1.4 to 4.0)	58.9	55.2 (1.7)	3.7 (0.4 to 7.0)
Most extensive surgery							
Breast conserving surgery	781 (38.8)	82.9	86.1 (1.2)	−3.2 (−5.6 to −0.8)	67.9	70.7 (1.6)	−2.8 (−5.9 to 0.3)
Mastectomy	1231 (61.2)	71.4	66.5 (1.3)	4.9 (2.4 to 7.5)	51.6	42.5 (1.4)	9.1 (6.4–11.8)
Systemic treatment							
None	1133 (56.3)	76.4	77.0 (1.3)	0.6 (−3.2 to 2.0)	59.8	57.6 (1.5)	2.2 (−0.7 to 5.1)
Endocrine therapy only	778 (38.7)	74.9	69.6 (1.7)	5.3 (2.0 to 8.6)	54.7	46.8 (1.8)	7.9 (4.4 to 11.4)
Chemotherapy only	52 (2.6)	73.2	75.0 (6.0)	−1.8 (−13.9 to 10.3)	60.4	63.3 (6.7)	−2.9 (−16.4–10.6)
Endocrine and chemotherapy	49 (2.4)	81.9	81.6 (5.5)	0.3 (−10.8 to 11.4)	65.6	51.0 (7.1)	14.6 (0.3 to 28.9)

but no chemotherapy (predicted: observed 5-year overall survival 5.3%, 95% CI = 2.0–8.6, P = 0.002).

Ten-year overall survival was predicted less well than 5-year overall survival, especially in older patients (aged 75 years or over) and patients with multiple comorbidities (5-year predicted and observed overall survival in patients with 4 or more comorbidities differed 11.8%, 95% CI = 6.7–16.7, whereas 10-year predicted and observed overall survival in this group differed 20.7%, 95% CI = 15.8–25.6).

Predicted and observed 10-year overall survival differed by 4.5%, 95% CI = 2.3–6.6, $P < 0.0001$ (Table 1). Again, the largest differences were observed in the oldest patients (predicted: observed 10-year overall survival in patients ≥ 85 12.0%, 95% CI = 6.9–17.1, $P < 0.0001$), and in patients with many comorbidities (predicted: observed 10-year overall survival in patients with 4 or more comorbidities 20.7%, 95% CI = 15.8–25.6, $P < 0.0001$). Furthermore, 10-year overall survival was overestimated in patients who received mastectomy as most extensive surgery (predicted: observed 10-year overall survival 9.1%, 95% CI = 6.4–11.8, $P < 0.0001$) and in patients who received both endocrine therapy and chemotherapy (predicted: observed 10-year overall survival 14.6%, 95% CI = 0.3–28.9, $P = 0.045$).

Next, calibration was assessed by composing calibration plots for 5-year and 10-year overall survival (Figure 1). Visually, both models resembled the 'ideal line' ($x = y$) reasonably well. Statistical analyses showed that 5-year overall survival did not significantly differ from the ideal line (standardised mortality ratio (SMR) = 1.07, 95% CI = 0.98–1.16, $P = 0.133$). However, 10-year overall survival was significantly different from the perfect calibration (SMR = 1.12, 95% CI = 1.05–1.20, $P = 0.0004$).

ROCs and corresponding c-indices are presented in Figure 2A and B. The c-index of the predicted 5-year overall survival was 0.73, 95% CI = 0.70–0.75, and 0.74, 95% CI = 0.72–0.76, for 10-year overall survival.

DISCUSSION

This study shows that the PREDICT tool can accurately predict 5-year overall survival in older patients with breast cancer. Ten-year overall survival estimates were slightly overestimated and less well calibrated than 5-year overall survival estimates, but still much more accurate than the predictions by Adjuvant! Online.

These findings are mostly in line with two previous validation studies that were performed in the United Kingdom, in which a good performance of the model was observed (Wishart et al, 2010, 2011). However, a recent validation study in younger patients (< 40) showed that the model did not accurately predict 5-year overall survival, whereas 10-year overall survival was well calibrated (Maishman et al, 2015), again showing that prediction models cannot be easily extrapolated to other study populations that were not represented in the original study cohort. In contrast, we showed that 10-year overall survival was overestimated by PREDICT by 4.5%, which was much more accurate than 10-year overall survival as predicted by Adjuvant! Online (de Glas et al, 2014c). This superior performance of PREDICT in older patients can most likely be explained by the fact that its predictions were based on a large cohort that did include a large number of older
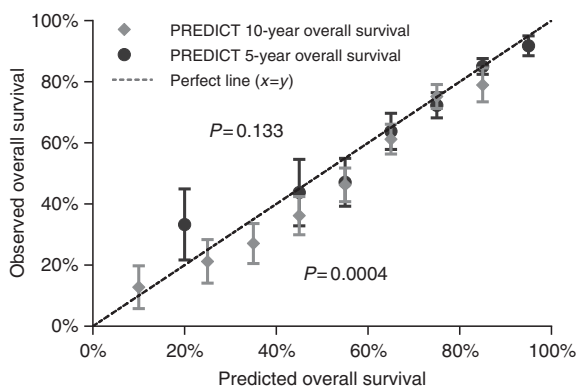
patients, whereas Adjuvant! Online was developed in a data set of patients with a maximum age of 69 years (Ravdin et al, 2001; Wishart et al, 2010). In addition, the unclearly defined comorbidity classification that was used by Adjuvant! Online strongly influenced its predictions (de Glas et al, 2014c). Furthermore, PREDICT has incorporated additional biomarkers (HER2 and Ki67), which may improve its predictions (Wishart et al, 2012).

Still, PREDICT has several limitations that should be addressed before the tool can be implemented in older patients. The main limitation of PREDICT is that only overall survival estimates are reported (Eastern Cancer Registry and Information Centre and Cambridge University (2015)), despite the fact that the original model also predicted breast cancer mortality (Wishart et al, 2010). Breast cancer-specific mortality is an important end point because competing causes of mortality can drive breast cancer treatment decisions in older adults (Wildiers et al, 2013).

Furthermore, PREDICT did not accurately predict overall survival (both at 5 and 10 years) for the oldest patients (≥ 85) and in patients with many comorbidities. These findings suggest that these patients were underrepresented in the original patient cohort in which the model was developed. In addition, PREDICT does not incorporate comorbidity status as a predictor, in contrast with Adjuvant! Online. Our data suggest that this is a limitation of the model, as overall survival was underestimated in patients with no comorbidity, whereas it was overestimated in patients with 3 or more comorbidities. This shows that the older breast cancer population is highly heterogeneous because of large differences in comorbidity status. It has been previously shown that comorbidity is a strong predictor for competing mortality (Kiderlen et al, 2014). Therefore, we suggest that future prediction models should incorporate a standardised comorbidity scale in their predictions. In addition, there are large differences between patients in terms of functional status, cognitive status, and physiological reserves.

There are several other prediction models besides PREDICT and Adjuvant! Online that can predict breast cancer outcomes. A recent systematic review identified 20 available prediction models of which Oncotype Dx is probably the most well-known (Engelhardt et al, 2013). Most of these models were based on genetic risk scores, and none of the prediction models besides Adjuvant! Online included comorbidity status (Engelhardt et al, 2013). Furthermore, none of the prediction models has been adequately validated in older patients (Engelhardt et al, 2013), despite the fact that ∼35–45% of patients with breast cancer are older than 65 years at diagnosis in Western countries, and this number is expected to increase in upcoming decades because of ageing of Western societies (DeSantis et al, 2013). Interestingly, the currently running ASTER 70s phase 3 trial (EudraCT 2011–004744-22) investigates the role of genomic grade index to guide indication of adjuvant chemotherapy in patients aged 70 years and older with luminal B breast cancer (Brain, 2014).

New prediction models or improvements of currently available prediction models that can be used in older patients are therefore required, especially as breast cancer mortality increases with age (van de Water et al, 2012). In contrast with younger patients, breast cancer mortality has not improved in older patients in the past decades (Bastiaannet et al, 2011; de Glas et al, 2014b). This may be explained by the fact that current guidelines are mostly based on studies that were performed in younger patients (Wildiers et al, 2007), or in highly selected older populations (van de Water et al, 2014). To improve outcome of older patients with breast cancer, it is essential to individualise treatment based on tumour characteristics, comorbidity status, but also functional status and the patients' preference. We propose that future models aimed at older patients specifically should use PREDICT as a starting point. Ideally, a new prediction tool specifically designed for older patients with breast cancer should incorporate comorbidity and functional performance. Such a tool should not only predict overall
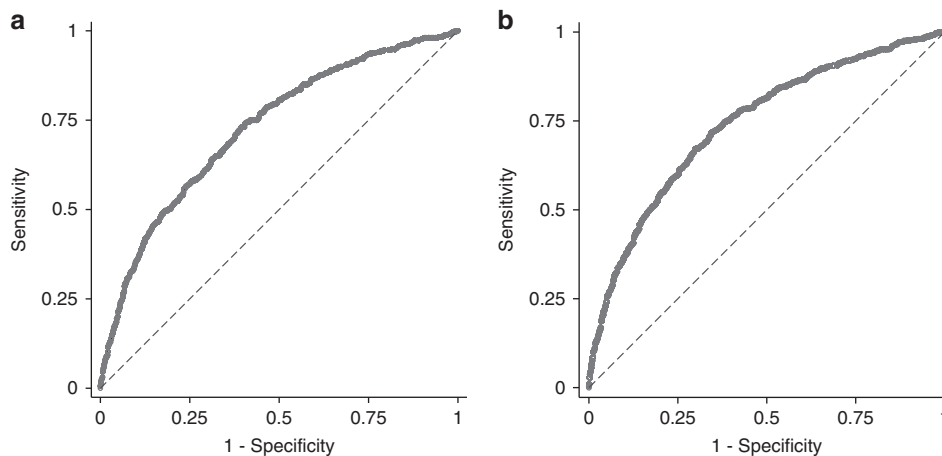

Figure 1. Calibration plot: observed vs predicted overall survival. *P*-values were calculated using Poisson regression.

Figure 2. ROC for 5- and 10-year overall survival. (**A**) Receiver-operating curve (ROC) for 5-year overall survival. The calculated area under the curve (AUC) was 0.73, 95% CI = 0.70–0.75. (**B**) ROC for 10-year overall survival. The calculated AUC was 0.74, 95% CI = 0.72–0.76.

survival, but also toxicity of treatment, quality of life, and functional decline. Both the International Society for Geriatric Oncology and the EORTC state that these so-called patient-related end points should be incorporated in clinical studies for older patients, in order to balance treatment efficacy with adverse events and toxicity in older patients (Wildiers *et al*, 2013). However, current randomised clinical breast cancer trials rarely incorporate these end points in their design (de Glas *et al*, 2014a). Large, prospective observational studies will therefore become increasingly important for the older breast cancer population, in order to study these end points and serve as the basis for new prediction tools that can aid in clinical decision making and individualised treatment.

Strengths and limitations. The main strength of this study is the use of a large population-based cohort, with nearly complete 10-year follow-up, that includes all consecutive patients in a geographically defined area in the Netherlands. In addition, we were able to perform central revision of pathological biomarkers in order to avoid intra-observer variation between hospitals. To our knowledge, this is the first study in which the validity of PREDICT was studied in a large cohort of older patients with breast cancer. This study also has its limitations. First, we were not able to perform FISH for HER2-status because of budget reasons, which is a limitation of the study as it may have resulted in overcall of HER2 positivity. Also, there was a large proportion of missing data concerning Her2 and Ki67 status, because of limited amounts of tumour material. Ki67 status remains controversial because of different cutoff points and reliability. We were restricted to the way the model uses Ki67-status (as a continuous variable). Furthermore, there was a large proportion of missing data on detection mode and Tumour Grade. However, the predictions of the model did not strongly differ in patients with missing data, which suggests that data were truly missing completely at random. In addition, despite these missing data, PREDICT still predicted 5-year overall survival reasonably well. Second, the effect of treatment can never be disentangled from the predictions of the model, as validation studies always appear in a retrospective setting, which means that patients already received treatment independent of the predictions of the model. Also, the percentage of 5% of patients receiving chemotherapy is fairly low, which may suggest that the validity of PREDICT may be different in populations where patients receive more chemotherapy. However, it must be noted that differences between observed and predicted survival outcomes did not strongly differ in several subgroups of treatment, which indicates that this has not strongly influenced our analyses. Third, we did not have information on the specific chemotherapy regimen that was used.

However, only 5% of patients received chemotherapy, and therefore it is unlikely that this has influenced the results. Finally, it must be noted that comparing observed and predicted outcomes using One-Sampled *t*-test is strongly influenced by sample size. Therefore, large subgroups may result in significant differences while small subgroups may not. Hence, one should interpret the absolute differences rather than the corresponding *P*-values.

In conclusion, PREDICT can accurately predict 5-year overall survival in older patients with breast cancer. Ten-year predicted overall survival was, however, overestimated. New prediction tools or improvements of currently available tools that can be used in older patients are needed, taking additional comorbidity into account and focusing on relevant end points for older patients with breast cancer.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

Dr A Hurria received research support from Celgene and GSK. In addition, she worked as a consultant for GTx Inc. and Boehringer Inhelheim Pharmaceuticals. The remaining authors declare no conflict of interest.

REFERENCES

Akobeng AK (2007) Understanding diagnostic tests 3: Receiver operating characteristic curves. *Acta Paediatr* **96**(5): 644–647.

Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B (2012) Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* **380**(9836): 37–43.

Bastiaannet E, Portielje JE, van de Velde CJ, de Craen AJ, van d V, Kuppen PJ, van der Geest LG, Janssen-Heijnen ML, Dekkers OM, Westendorp RG, Liefers GJ (2011) Lack of survival gain for elderly women with breast cancer. *Oncologist* **16**(4): 415–423.

Biganzoli L, Wildiers H, Oakman C, Marotti L, Loibl S, Kunkler I, Reed M, Ciatto S, Voogd AC, Brain E, Cutuli B, Terret C, Gosney M, Aapro M, Audisio R (2012) Management of elderly patients with breast cancer: updated recommendations of the International Society of Geriatric Oncology (SIOG) and European Society of Breast Cancer Specialists (EUSOMA). *Lancet Oncol* **13**(4): e148–e160.

Brain E. Breast cancer in older women: predicting adjuvant benefit (2014) *Lancet Oncol* **15**(7): 672–674.

Braithwaite D, Satariano WA, Sternfeld B, Hiatt RA, Ganz PA, Kerlikowske K, Moore DH, Slattery ML, Tammemagi M, Castillo A, Melisko M, Esserman L, Weltzien EK, Caan BJ (2010) Long-term prognostic role of functional limitations among women with breast cancer. *J Natl Cancer Inst* **102**(19): 1468–1477.

de Glas NA, Hamaker ME, Kiderlen M, de Craen AJ, Mooijaart SP, van de Velde CJ, van Munster BC, Portielje JE, Liefers GJ, Bastiaannet E (2014a) Choosing relevant endpoints for older breast cancer patients in clinical trials: an overview of all current clinical trials on breast cancer treatment. *Breast Cancer Res Treat* **146**(3): 591–597.

de Glas NA, Jonker JM, Bastiaannet E, de Craen AJ, van de Velde CJ, Siesling S, Liefers GJ, Portielje JE, Hamaker ME (2014b) Impact of omission of surgery on survival of older patients with breast cancer. *Br J Surg* **101**(11): 1397–1404.

de Glas NA, Kiderlen M, Bastiaannet E, de Craen AJ, van de Water W, van de Velde CJ, Liefers GJ (2013) Postoperative complications and survival of elderly breast cancer patients: a FOCUS study analysis. *Breast Cancer Res Treat* **138**(2): 561–569.

de Glas NA, van de Water W, Engelhardt EG, Bastiaannet E, de Craen AJ, Kroep JR, Putter H, Stiggelbout AM, Weijl NI, van de Velde CJ, Portielje JE, Liefers GJ (2014c) Validity of Adjuvant! Online program in older patients with breast cancer: a population-based study. *Lancet Oncol* **15**(7): 722–729.

de Kruijf EM, van Nes JG, Sajet A, Tummers QR, Putter H, Osanto S, Speetjens FM, Smit VT, Liefers GJ, van de Velde CJ, Kuppen PJ (2010) The predictive value of HLA class I tumor cell expression and presence of intratumoral Tregs for chemotherapy in patients with early breast cancer. *Clin Cancer Res* **16**(4): 1272–1280.

DeSantis C, Ma J, Bryan L, Jemal A (2013) Breast cancer statistics, 2013. *CA Cancer J Clin* **64**(1): 52–62.

DeSantis C, Siegel R, Bandi P, Jemal A (2011) Breast cancer statistics, 2011. *CA Cancer J Clin* **61**(6): 409–418.

Early Breast Cancer Trialists' Collaborative Group (EBCTCG) (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**(9472): 1687–1717.

Eastern Cancer Registry and Information Centre and Cambridge University (2015) Predict (accessed on 02 February 2015). Available from http://www.predict.nhs.uk.

Engelhardt EG, Garvelink MM, de Haes JH, van der Hoeven JJ, Smets EM, Pieterse AH, Stiggelbout AM (2013) Predicting and communicating the risk of recurrence and death in women with early-stage breast cancer: a systematic review of risk prediction models. *J Clin Oncol* **32**(3): 238–250.

Engelhardt EG, Pieterse AH, van Duijn-Bakker N, Kroep JR, de Haes HC, Smets EM, Stiggelbout AM (2014) Breast cancer specialists' views on and use of risk prediction models in clinical practice: a mixed methods approach. *Acta Oncol* **54**(3): 361–367.

Kiderlen M, de Glas NA, Bastiaannet E, Engels CC, van de Water W, de Craen AJ, Portielje JE, van de Velde CJ, Liefers GJ (2013) Diabetes in relation to breast cancer relapse and all-cause mortality in elderly breast cancer patients: a FOCUS study analysis. *Ann Oncol* **24**(12): 3011–3016.

Kiderlen M, de Glas NA, Bastiaannet E, van de Water W, de Craen AJ, Guicherit OR, Merkus JW, Extermann M, van de Velde CJ, Liefers GJ (2014) Impact of comorbidity on outcome of older breast cancer patients: a FOCUS cohort study. *Breast Cancer Res Treat* **145**(1): 185–192.

Maishman T, Copson E, Stanton L, Gerty S, Dicks E, Durcan L, Wishart GC, Pharoah P, Eccles D (2015) An evaluation of the prognostic model PREDICT using the POSH cohort of women aged 40 years at breast cancer diagnosis. *Br J Cancer* **112**(6): 983–991.

Mook S, Schmidt MK, Rutgers EJ, van de Velde AO, Visser O, Rutgers SM, Armstrong N, van't Veer LJ, Ravdin PM (2009) Calibration and discriminatory accuracy of prognosis calculation for breast cancer with the online Adjuvant! program: a hospital-based retrospective cohort study. *Lancet Oncol* **10**(11): 1070–1076.

NABON (2012) Richtlijn Mammacarcinoom versie 2.0. Last updated on 13 February 2012 (accessed on 31 January 2015). Available from http://www.oncoline.nl/mammacarcinoom.

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**(12): 2716–2725.

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**(4): 980–991.

van de Water W, Kiderlen M, Bastiaannet E, Siesling S, Westendorp RG, van de Velde CJ, Nortier JW, Seynaeve C, de Craen AJ, Liefers GJ (2014) External validity of a trial comprised of elderly patients with hormone receptor-positive breast cancer. *J Natl Cancer Inst* **106**(4): dju051.

van de Water W, Markopoulos C, van de Velde CJ, Seynaeve C, Hasenburg A, Rea D, Putter H, Nortier JW, de Craen AJ, Hille ET, Bastiaannet E, Hadji P, Westendorp RG, Liefers GJ, Jones SE (2012) Association between age at diagnosis and disease-specific mortality among postmenopausal women with hormone receptor-positive breast cancer. *JAMA* **307**(6): 590–597.

van Houwelingen H, Putter H (2012) *Dynamic Prediction in Clinical Survival Analysis*. CRC Press: Boca Raton.

Wildiers H, Kunkler I, Biganzoli L, Fracheboud J, Vlastos G, Bernard-Marty C, Hurria A, Extermann M, Girre V, Brain E, Audisio RA, Bartelink H, Barton M, Giordano SH, Muss H, Aapro M (2007) Management of breast cancer in elderly individuals: recommendations of the International Society of Geriatric Oncology. *Lancet Oncol* **8**(12): 1101–1115.

Wildiers H, Mauer M, Pallis A, Hurria A, Mohile SG, Luciani A, Curigliano G, Extermann M, Lichtman SM, Ballman K, Cohen HJ, Muss H, Wedding U (2013) End points and trial design in geriatric oncology research: a joint European organisation for research and treatment of cancer-alliance for clinical trials in oncology-international society of geriatric oncology position article. *J Clin Oncol* **31**(29): 3711–3718.

Wishart GC, Azzato EM, Greenberg DC, Rashbass J, Kearins O, Lawrence G, Caldas C, Pharoah PD (2010) PREDICT: a new UK prognostic model that predicts survival following surgery for invasive breast cancer. *Breast Cancer Res* **12**(1): R1.

Wishart GC, Bajdik CD, Azzato EM, Dicks E, Greenberg DC, Rashbass J, Caldas C, Pharoah PD (2011) A population-based validation of the prognostic model PREDICT for early breast cancer. *Eur J Surg Oncol* **37**(5): 411–417.

Wishart GC, Bajdik CD, Dicks E, Provenzano E, Schmidt MK, Sherman M, Greenberg DC, Green AR, Gelmon KA, Kosma VM, Olson JE, Beckmann MW, Winqvist R, Cross SS, Severi G, Huntsman D, Pylkas K, Ellis I, Nielsen TO, Giles G, Blomqvist C, Fasching PA, Couch FJ, Rakha E, Foulkes WD, Blows FM, Begin LR, van't Veer LJ, Southey M, Nevanlinna H, Mannermaa A, Cox A, Cheang M, Baglietto L, Caldas C, Garcia-Closas M, Pharoah PD (2012) PREDICT Plus: development and validation of a prognostic model for early breast cancer that includes HER2. *Br J Cancer* **107**(5): 800–807.

World Health Organisation (1992) *The ICD-10 Classification of Mental and Behavioural Disorders*, vol 10. pp 1–267. Available from http://www.who.int/substance_abuse/terminology/ICD10ResearchDiagnosis.pdf?ua=1.