Breast Cancer Research and Treatment
https://doi.org/10.1007/s10549-019-05399-2

EPIDEMIOLOGY

Validation of the online prediction model CancerMath in the Dutch breast cancer population

Liza A. Hoveling[1] · Marissa C. van Maaren[1] · Tom Hueting[2] · Luc J. A. Strobbe[3] · Mathijs P. Hendriks[4] · Gabe S. Sonke[5] · Sabine Siesling[1,6]

Received: 15 April 2019 / Accepted: 6 August 2019

Abstract

Purpose CancerMath predicts the expected benefit of adjuvant systemic therapy on overall (OS) and breast cancer-specific survival (BCSS). Here, CancerMath was validated in Dutch breast cancer patients.

Methods All operated women diagnosed with stage I–III primary invasive breast cancer in 2005 were identified from the Netherlands Cancer Registry. Calibration was assessed by comparing 5- and 10-year predicted and observed OS/BCSS using χ^2 tests. A difference > 3% was considered as clinically relevant. Discrimination was assessed by area under the receiver operating characteristic (AUC) curves.

Results Altogether, 8032 women were included. CancerMath underestimated 5- and 10-year OS by 2.2% and 1.9%, respectively. AUCs of 5- and 10-year OS were both 0.77. Divergence between predicted and observed OS was most pronounced in grade II, patients without positive nodes, tumours 1.01–2.00 cm, hormonal receptor positive disease and patients 60–69 years. CancerMath underestimated 5- and 10-year BCSS by 0.5% and 0.6%, respectively. AUCs were 0.78 and 0.73, respectively. No significant difference was found in any subgroup.

Conclusion CancerMath predicts OS accurately for most patients with early breast cancer although outcomes should be interpreted with care in some subgroups. BCSS is predicted accurately in all subgroups. Therefore, CancerMath can reliably be used in (Dutch) clinical practice.

Keywords CancerMath · Breast cancer · Prediction model · Validation · Overall survival · Breast cancer-specific survival

Introduction

Adjuvant systemic therapy reduces recurrence risks and improves long-term survival in early breast cancer [1–3]. However, its effects vary due to tumour heterogeneity and differences in treatment response [4–6]. Besides, it may be associated with severe side effects and high costs [7, 8].

Several prediction models have been developed to help personalise adjuvant systemic therapy use in breast cancer patients. A commonly used online prediction model is PREDICT [9], which has been validated in many countries [9–18]. However, PREDICT does not predict accurately in ER-negative disease, patients < 40 years, patients > 79 years, T3 tumours and patients treated with both endocrine therapy and chemotherapy [15–17]. Another existing prediction model that assesses the benefits of adjuvant systemic therapy is the CancerMath therapy calculator (hereafter abbreviated as CancerMath) [19]. CancerMath, in contrast to PREDICT, includes histological type, but lacks information on

Liza A. Hoveling and Marissa C. van Maaren contributed equally to this work.

✉ Marissa C. van Maaren
m.vanmaaren@iknl.nl

[1] Department of Research, Netherlands Comprehensive Cancer Organisation, P.O. Box 19079, 3501 DB Utrecht, The Netherlands

[2] Evidencio Medical Decision Support, Haaksbergen, The Netherlands

[3] Department of Surgical Oncology, Canisius Wilhelmina Hospital, Nijmegen, The Netherlands

[4] Department of Medical Oncology, Northwest Clinics, Alkmaar, The Netherlands

[5] Department of Medical Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands

[6] Department of Health Technology and Services Research, Faculty of Behavioural, Management and Social Sciences, Technical Medical Centre, University of Twente, Enschede, The Netherlands

Published online: 30 August 2019

Table 1 Predicted and observed 5- and 10-year overall survival by subgroups of predictors in CancerMath

Predictors	Categories	*N* (%)	5-year OS				10-year OS			
			Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value	Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value
Entire cohort		8032 (100.0)	6988 (87.0)	7167 (89.2) (88.5–89.9)	−179 (−2.2)	0.03	5952 (74.1)	6101 (76.0) (75.0–76.9)	−149 (−1.9)	0.05
Age (years)	< 40	420 (5.2)	394 (93.8)	387 (92.1) (89.1–94.4)	7 (1.7)	0.73	369 (87.8)	353 (84.1) (80.2–87.4)	16 (3.7)	0.41
	40–49	1526 (19.0)	1434 (94.0)	1436 (94.1) (92.8–95.2)	−2 (−0.1)	1.00	1343 (88.0)	1332 (87.3) (85.5–88.9)	11 (0.7)	0.76
	50–59	2229 (27.8)	2055 (92.2)	2088 (93.7) (92.6–94.6)	−33 (−1.5)	0.47	1868 (83.8)	1931 (86.6) (85.1–88.0)	−63 (−2.8)	0.14
	60–69	1929 (24.0)	1703 (88.3)	1782 (92.4) (91.1–93.5)	−79 (−4.1)	0.06	1449 (75.1)	1531 (79.4) (77.5–81.2)	**−82 (−4.3)**	**0.03**
	70–79	1312 (16.3)	1050 (80.0)	1104 (84.2) (82.1–86.1)	−54 (−4.2)	0.09	760 (57.9)	811 (61.8) (59.1–64.4)	−51 (−3.9)	0.06
	> 79	616 (7.7)	356 (57.8)	370 (60.1) (56.1–64.0)	−14 (−2.3)	0.46	165 (26.8)	143 (23.2) (19.9–26.8)	22 (3.6)	0.09
Tumour diameter (cm)	< 0.11	39 (0.5)	36 (93.5)	38 (97.4) (86.5–99.9)	−2 (−3.9)	0.81	33 (84.9)	30 (76.9) (60.7–88.9)	3 (8.0)	0.59
	0.11–0.50	256 (3.2)	236 (92.3)	241 (94.1) (90.5–96.7)	−5 (−1.8)	0.76	212 (82.8)	224 (87.5) (82.8–91.3)	−12 (−4.7)	0.41
	0.51–1.00	1211 (15.1)	1107 (91.4)	1159 (95.7) (94.4–96.8)	−52 (−4.3)	0.12	981 (81.0)	1044 (86.2) (84.1–88.1)	**−63 (−5.2)**	**0.04**
	1.01–2.00	3678 (45.8)	3251 (88.4)	3382 (92.0) (91.0–92.8)	**−131 (−3.6)**	**0.02**	2806 (76.3)	2913 (79.2) (77.9–80.5)	−107 (−2.9)	0.04
	2.01–3.00	2088 (26.0)	1758 (84.2)	1770 (84.8) (83.2–86.3)	−12 (−0.6)	0.78	1451 (69.5)	1453 (69.6) (67.6–71.6)	−2 (−0.1)	1.00
	3.01–4.00	585 (7.3)	467 (79.9)	458 (78.3) (74.7–81.6)	9 (1.6)	0.66	366 (62.6)	353 (60.3) (56.2–64.3)	13 (2.3)	0.49
	4.01–5.00	175 (2.2)	133 (76.2)	119 (68.0) (60.5–74.8)	14 (8.2)	0.21	100 (57.2)	84 (48.0) (40.4–55.7)	16 (9.2)	0.11
Number of positive nodes	0	5291 (65.9)	4645 (87.8)	4843 (91.5) (90.8–92.3)	**−198 (−3.7)**	**0.00**	3984 (75.3)	4184 (79.1) (78.0–80.2)	**−200 (−3.8)**	**0.00**
	1–3	2180 (27.1)	1884 (86.4)	1894 (86.9) (85.4–88.3)	−10 (−0.5)	0.81	1598 (73.3)	1587 (72.8) (70.9–74.7)	11 (0.5)	0.79
	4–9	532 (6.6)	437 (82.2)	410 (77.1) (73.3–80.6)	27 (5.1)	0.19	354 (66.5)	315 (59.2) (54.9–63.4)	**39 (7.3)**	**0.04**
	9–10	29 (0.4)	21 (73.8)	20 (69.0) (49.2–84.7)	1 (4.8)	0.76	15 (53.1)	15 (51.7) (32.5–70.6)	0 (0.0)	1.00

Table 1 (continued)

Predictors	Categories	*N* (%)	5-year OS				10-year OS			
			Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value	Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value
ER status	Unknown	116 (1.4)	103 (88.4)	107 (92.2) (85.8–96.4)	−4 (−3.8)	0.66	89 (76.7)	90 (77.6) (68.9–84.8)	−1 (−0.9)	0.92
	Positive	6535 (81.4)	5718 (87.5)	5949 (91.0) (90.3–91.7)	**−231 (−3.5)**	**0.00**	4882 (74.7)	5055 (77.4) (76.3–78.4)	−173 (−2.7)	0.01
	Negative	1381 (17.2)	1172 (84.9)	1111 (80.5) (78.3–82.5)	61 (4.4)	0.07	978 (70.8)	956 (69.2) (66.7–71.7)	22 (1.6)	0.49
PR status	Unknown	507 (6.3)	439 (86.5)	450 (88.8) (85.7–91.4)	−11 (−2.3)	0.58	371 (73.2)	375 (74.0) (69.9–77.7)	−4 (−0.8)	0.84
	Positive	5109 (63.6)	4496 (88.0)	4682 (91.6) (90.8–92.4)	**−186 (−3.6)**	**0.01**	3868 (75.7)	3997 (78.2) (77.1–79.4)	−129 (−2.5)	0.04
	Negative	2416 (30.1)	2056 (85.1)	2035 (84.2) (82.7–85.7)	21 (0.9)	0.65	1711 (70.8)	1729 (71.6) (69.7–73.4)	−18 (−0.8)	0.65
HER2 status	Unknown	1413 (17.6)	1214 (85.9)	1239 (87.7) (85.9–89.4)	−25 (−1.8)	0.47	1016 (71.9)	1032 (73.0) (70.6–75.3)	−16 (−1.1)	0.62
	Positive	1000 (12.5)	857 (85.7)	892 (89.2) (87.1–91.1)	−35 (−3.5)	0.23	722 (72.2)	774 (77.4) (74.7–80.0)	−52 (−5.2)	0.05
	Negative	5619 (70.0)	4922 (87.6)	5036 (89.6) (88.8–90.4)	−114 (−2.0)	0.10	4214 (75.0)	4295 (76.4) (75.3–77.5)	−81 (−1.4)	0.21
HR/HER2 status	Unknown	1417 (18.3)	1264 (85.9)	1285 (87.4) (85.7–89.1)	−21 (−1.4)	0.55	1059 (72.0)	1071 (72.8) (70.5–75.1)	−12 (−0.8)	0.71
	HR+/HER2−	4827 (60.1)	4237 (87.8)	4402 (91.2) (90.4–92.0)	**−165 (3.4)**	**0.01**	3633 (72.3)	3754 (77.8) (76.6–78.9)	−121 (−2.5)	0.04
	HR+/HER2+	646 (8.0)	565 (87.5)	597 (92.4) (90.4–94.4)	−32 (4.9)	0.18	487 (75.4)	514 (79.6) (76.4–82.7)	−27 (−4.2)	0.22
	HER2 positive	341 (4.3)	281 (92.3)	287 (84.2) (80.2–88.1)	−6 (1.8)	0.72	226 (66.3)	252 (73.9) (69.2–78.6)	−26 (−7.6)	0.08
	Triple negative	747 (9.3)	644 (86.3)	596 (79.8) (76.9–92.7)	48 (6.5)	0.06	546 (73.1)	510 (68.3) (64.9–71.6)	36 (4.8)	0.12
Histological type	Ductal	6526 (81.3)	5678 (87.0)	5821 (89.2) (88.4–90.0)	−143 (−2.2)	0.06	4836 (74.1)	4950 (75.9) (74.8–76.9)	−114 (−1.8)	0.10
	Lobular	790 (9.8)	680 (86.1)	707 (89.5) (87.1–91.5)	−27 (−3.4)	0.30	570 (72.1)	605 (76.6) (73.4–79.4)	−35 (−4.5)	0.14
	Ductal and lobular	289 (3.6)	257 (88.9)	266 (92.0) (88.3–94.9)	−9 (−3.1)	0.57	223 (77.2)	228 (78.9) (73.7–83.4)	−5 (−1.7)	0.74
	Other	427 (5.3)	375 (87.9)	373 (87.4) (83.8–90.4)	2 (0.5)	0.92	323 (75.6)	318 (74.5) (70.1–78.5)	5 (1.1)	0.79

Table 1 (continued)

Predictors	Categories	*N* (%)	5-year OS				10-year OS			
			Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value	Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value
Grade	Unknown	317 (3.9)	273 (86.0)	280 (88.3) (84.3–91.6)	−7 (−2.3)	0.65	229 (72.3)	239 (75.4) (70.3–80.0)	−10 (−3.1)	0.52
	I	1854 (23.1)	1667 (89.9)	1742 (94.0) (92.8–95.0)	−75 (−4.1)	0.07	1459 (78.7)	1523 (82.2) (80.3–83.9)	−64 (−3.5)	0.09
	II	3488 (43.4)	3024 (86.7)	3155 (90.5) (89.4–91.4)	**−131 (−3.8)**	**0.02**	2560 (73.4)	2690 (77.1) (75.7–78.5)	**−130 (−3.7)**	**0.01**
	III	2373 (29.5)	2029 (85.5)	1990 (83.9) (82.3–85.3)	39 (1.6)	0.39	1701 (71.7)	1649 (69.5) (67.6–71.3)	52 (2.2)	0.20
Hormonal therapy	None	4564 (56.8)	3994 (87.5)	4058 (88.9) (88.0–89.8)	−65 (−1.4)	0.31	3409 (74.7)	3517 (77.1) (75.8–78.3)	−108 (−2.4)	0.07
	Tamoxifen	853 (10.6)	740 (86.8)	788 (92.4) (90.4–94.1)	−48 (−5.6)	0.08	628 (73.6)	673 (74.7) (71.6–77.6)	−45 (−1.1)	0.07
	Aromatase inhibitor	1724 (21.5)	1445 (83.8)	1456 (84.5) (82.7–86.1)	−11 (−0.7)	0.76	1184 (68.7)	1181 (68.5) (66.3–70.7)	3 (0.2)	0.92
	Tamoxifen to aromatase inhibitor	659 (8.2)	592 (89.8)	637 (96.7) (95.0–97.9)	−45 (−6.9)	0.06	523 (79.4)	560 (85.0) (82.0–87.6)	−37 (−5.6)	0.11
	Ovarian ablation	3 (0.04)	3 (100.0)	3 (100.0) (n.a.)	0 (0.0)	1.00	3 (0.0)	3 (100.0) (n.a.)	0 (0.0)	1.00
	Ovarian ablation and tamoxifen	229 (2.9)	216 (94.3)	225 (98.3) (95.6–99.5)	−9 (−4.0)	0.54	203 (88.7)	203 (88.7) (83.8–92.4)	0 (0.0)	1.00
Chemotherapy	None	5363 (66.8)	4553 (84.9)	4735 (88.3) (87.4–89.1)	**−182 (−3.4)**	**0.01**	3749 (69.9)	3894 (72.6) (71.4–73.8)	−145 (−2.7)	0.02
	Generation 1	1158 (14.4)	1043 (90.1)	1022 (88.3) (86.3–90.1)	21 (1.8)	0.51	930 (80.3)	935 (80.7) (78.4–83.0)	−5 (−0.4)	0.86
	Generation 2	673 (8.4)	627 (93.1)	638 (94.8) (92.8–96.4)	−11 (−1.7)	0.65	578 (85.9)	578 (85.9) (83.0–88.4)	0 (0.0)	1.00
	Generation 3	151 (1.9)	140 (92.4)	139 (92.1) (86.5–95.8)	1 (0.3)	1.00	128 (84.7)	126 (83.4) (76.5–89.0)	2 (1.3)	0.86
	Generation unknown	687 (8.6)	627 (91.2)	633 (92.1) (89.9–94.0)	−6 (−0.9)	0.79	565 (82.2)	568 (82.7) (79.6–85.4)	−3 (−0.5)	0.89

Table 1 (continued)

Predictors	Categories	*N* (%)	5-year OS				10-year OS			
			Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value	Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value
Adjuvant systemic therapy	None	3712 (46.2)	3244 (87.4)	3359 (90.5) (89.5–91.4)	**− 115 (− 3.1)**	**0.04**	2758 (74.3)	2881 (77.6) (76.2–78.9)	**− 123 (− 3.3)**	**0.02**
	Both	1817 (22.6)	1684 (92.7)	1733 (95.4) (94.3–96.3)	− 49 (− 2.7)	0.00	1546 (85.1)	1571 (86.5) (84.8–88.0)	− 25 (− −1.4)	0.00
	Only hormonal therapy	1651 (20.6)	1313 (79.5)	1376 (83.3) (81.5–85.1)	− 63 (− 3.8)	0.08	994 (60.2)	1013 (61.4) (59.0–63.7)	− 19 (− 1.2)	0.54
	Only chemotherapy	852 (10.6)	752 (88.3)	699 (82.0) (79.3–84.6)	53 (6.3)	0.05	654 (76.8)	636 (74.7) (71.6–77.5)	18 (2.1)	0.48
Targeted therapy	No	7609 (94.7)	6613 (86.9)	6771 (89.0) (88.3–89.7)	− 158 (− 2.1)	0.05	5615 (73.8)	5737 (75.4) (74.4–76.4)	− 122 (− 1.6)	0.10
	Yes	423 (5.3)	379 (89.6)	396 (96.6) (91.3–96.0)	**− 17 (− 4.0)**	**0.03**	336 (79.4)	364 (86.1) (82.7–89.4)	− 28 (− 6.7)	0.13

Differences and accompanying *p*-values indicated in bold are considered to be significant

N total number, *OS* overall survival, *CI* confidence interval, *ER* oestrogen receptor, *PR* progesterone receptor, *HR* hormonal receptor, *HER2* human epidermal growth factor receptor 2, *n.a* not applicable due to low number of patients

screening and Ki-67, and expresses the therapy benefit in gained life years. Multiple studies report that many patients still have an inaccurate perception of adjuvant treatment effects [20, 21], which may in part be explained by presentation of expected benefits.

CancerMath was released in 2009 and predicts overall (OS) and breast cancer-specific survival (BCSS) until 15 years from diagnosis, using the following predictors: age at diagnosis, tumour size, grade, oestrogen (ER) and progesterone receptor (PR) status, human epidermal growth factor receptor 2 (HER2) status, nodal status and histological type. The parameters were derived from the Surveillance, Epidemiology, and End Results (SEER) database [19]. The expected benefit of adjuvant systemic therapy was based on the reductions in death found by meta-analyses of randomised trials [1, 2]. This study aimed to validate the online prediction model CancerMath in the Dutch breast cancer population by assessing its prognostic performance in different subgroups.

Methods

Study design

Data on patient-, tumour- and treatment-related characteristics were obtained from the Netherlands Cancer Registry (NCR). Topography and morphology were coded according to the International Classification of Diseases for Oncology (ICD-O), 3rd edition [22]. Staging was coded according to the tumour, node and metastasis (TNM) classification system, 6th edition [23]. Data on recurrences were actively obtained from patient files. Vital status and date of death at 5 and 10 years following diagnosis were obtained through linkage with the Municipal Personal Records database.

Study population

All operated women diagnosed with stage I–III primary invasive breast cancer in 2005 in the Netherlands, who had a pathologically established tumour, were identified (none of the included patients received neoadjuvant systemic therapy). Women with an unknown tumour size, a tumour > 5 cm, an unknown number of positive lymph nodes or > 10 positive lymph nodes were excluded to match the CancerMath criteria.

Outcomes and definitions

Primary outcomes were OS and BCSS. OS was defined as the time between diagnosis and date of death or last observation. Due to lacking data on cause of death, we classified patients as 'died due to breast cancer' if they experienced

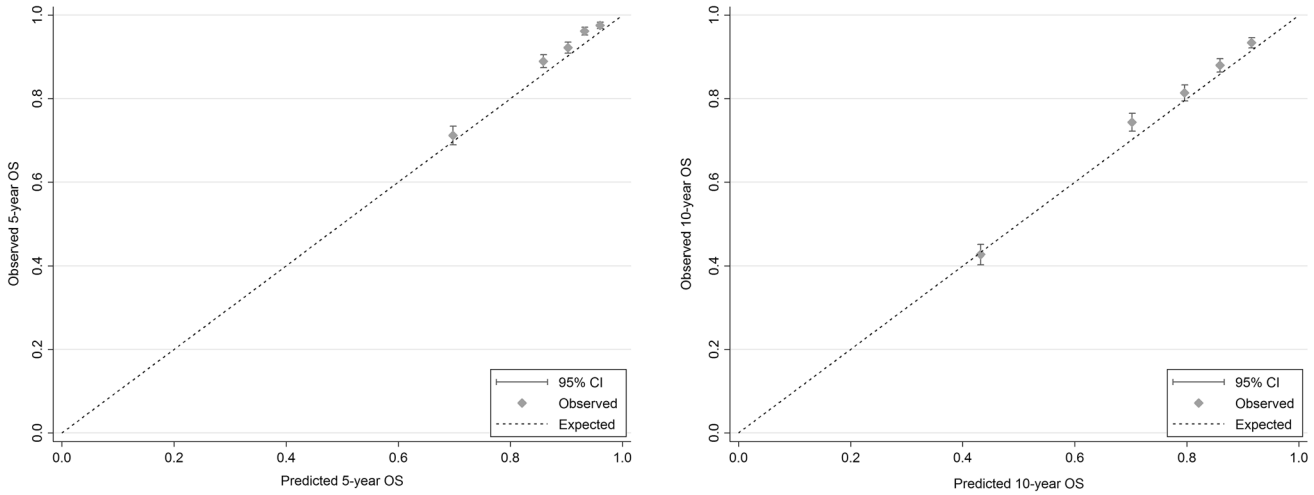

Fig. 1 Observed and predicted 5-year (left) and 10-year (right) overall survival in Dutch non-metastatic breast cancer patients diagnosed in 2005

a distant metastasis and died within 5 or 10 years from diagnosis.

Statistical analysis

For 687 of 2669 (25.7%) women receiving chemotherapy, it was unknown whether they received first- or second-generation chemotherapy. To determine the classification of these patients, calibration and discrimination analyses were performed in these women classified as first generation, second generation, a combination of first and second generation (by using the mean coefficient of first and second generation in the predictions) or they were excluded. Since results were similar (data not shown), it was decided to classify these women as a combination of first and second generation in further analyses.

Five- and 10-year OS and BCSS were calculated for different subgroups based on age (< 40, 40–49, 50–59, 60–69, 70–79, > 79 years), tumour size, number of positive lymph nodes, histological type (ductal, lobular, a combination or other), tumour grade, ER, PR and HER2 status, endocrine therapy, chemotherapy and targeted therapy. Model calibration was quantified by comparing observed and predicted outcomes using a χ^2 test. A significant difference of > 3% was regarded as clinically relevant, since a 10-year survival benefit of 3–5% is generally considered to be an indication for adjuvant systemic therapy [10]. Calibration was graphically presented by plotting averages of observed and predicted outcomes, grouped by quintiles based on the predicted outcomes. The slope of the fitted line was compared with the slope indicating a perfect relationship ($y = x$). Discriminatory accuracy was assessed by generating receiver operating characteristic (ROC) curves. The discriminatory accuracy was quantified by the area under the ROC curve (AUC), which was calculated for the entire population as well as all predefined subgroups. An AUC of 1 indicates excellent discriminative performance while an AUC of 0.5 indicates no discriminative performance.

All underlying statistical equations of CancerMath were freely available as HTML code [24]. To obtain predicted outcomes for all included patients, the HTML code was translated into a R script (R version 3.5.1). All further analyses were performed in Stata 14.1 (StataCorp LP, College Station, TX, USA). P values < 0.05 were considered to be statistically significant.

Results

In total, 8772 patients were included. Women with an unknown tumour size ($n = 218$), a tumour > 5 cm ($n = 435$), an unknown number of positive lymph nodes ($n = 94$) or > 10 positive lymph nodes ($n = 305$) were excluded, resulting in a final study population of 8032 patients. None of the included patients were treated with neoadjuvant systemic therapy. The median age at diagnosis was 59 years (interquartile range: 50–69 years). After 5 years, 865 (10.8%) women had died. Of these women, 398 (5% of entire population) developed distant metastases and were considered to have died of breast cancer. After 10 years, 1,931 (24.0%) women had died. Of these, 764 (9.5% of entire population) developed distant metastases and were considered to have died of breast cancer. Most women had no positive lymph nodes, presented with a ductal tumour, a grade II tumour, were hormonal receptor positive or HER2 negative. Almost

Table 2 Predicted and observed 5- and 10-year breast cancer-specific survival by subgroups of predictors in CancerMath

Predictors	Categories	*N* (%)	5-year BCSS				10-year BCSS			
			Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value	Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value
Entire cohort		8032 (100.0)	7590 (94.5)	7634 (95.0) (94.5–95.5)	−44 (−0.5)	0.62	7221 (89.9)	7268 (90.5) (89.8–91.1)	−47 (−0.6)	0.58
Age (years)	< 40	420 (5.2)	396 (94.3)	391 (93.1) (90.2–95.3)	5 (1.2)	0.81	374 (89.1)	360 (85.7) (82.0–88.9)	14 (3.4)	0.46
	40–49	1526 (19.0)	1453 (95.2)	1449 (95.0) (93.7–96.0)	4 (0.2)	0.92	1386 (90.8)	1369 (89.7) (88.1–91.2)	17 (1.1)	0.65
	50–59	2229 (27.8)	2111 (94.7)	2140 (96.0) (95.1–96.8)	−29 (−1.3)	0.53	2006 (90.0)	2059 (92.4) (91.2–93.4)	−53 (−2.4)	0.24
	60–69	1929 (24.0)	1823 (94.5)	1855 (96.2) (95.2–97.0)	−32 (−1.7)	0.45	1730 (89.7)	1766 (91.6) (90.2–92.8)	−36 (−1.9)	0.39
	70–79	1312 (16.3)	1237 (94.3)	1230 (93.8) (92.3–95.0)	7 (0.5)	0.84	1179 (89.9)	1163 (88.6) (86.8–90.3)	16 (1.3)	0.63
	> 79	616 (7.7)	572 (92.8)	569 (92.4) (90.0–94.3)	3 (0.4)	0.92	547 (88.8)	551 (89.5) (86.7–91.8)	−4 (−0.7)	0.86
Tumour diameter (cm)	< 0.1	39 (0.5)	39 (100.0)	39 (100.0) (91.0–1)	0 (0.0)	1.00	39 (100.0)	39 (100.0) (91.0–1.00)	0 (0.0)	1.00
	0.11–0.5	256 (3.2)	252 (98.5)	254 (99.2) (97.2–99.9)	−2 (−0.7)	0.92	249 (97.2)	251 (98.1) (95.5–99.4)	−2 (−0.9)	0.89
	0.51–1.00	1211 (15.1)	1178 (97.3)	1195 (98.7) (97.9–99.2)	−17 (−1.4)	0.62	1150 (95.0)	1169 (96.5) (95.3–97.5)	−19 (−1.5)	0.58
	1.01–2.00	3678 (45.8)	3498 (95.1)	3545 (96.4) (95.7–97.0)	−47 (−1.3)	0.42	3347 (91.0)	3382 (92.0) (91.0–92.8)	−35 (−1.0)	0.54
	2.01–3.00	2088 (26.0)	1940 (92.9)	1929 (92.4) (91.2–93.5)	11 (0.5)	0.81	1819 (87.1)	1808 (86.6) (85.1–88.0)	11 (0.5)	0.81
	3.01–4.00	585 (7.3)	527 (90.1)	520 (88.9) (86.1–91.3)	7 (1.2)	0.75	480 (82.1)	482 (82.4) (79.1–85.4)	−2 (−0.3)	0.92
	4.01–5.00	175 (2.2)	155 (88.3)	152 (86.9) (80.9–91.5)	3 (1.4)	0.84	138 (79.1)	137 (78.3) (71.4–84.2)	1 (0.8)	0.92
Number of positive nodes	0	5291 (65.9)	5053 (95.5)	5121 (96.8) (96.3–97.2)	−68 (−1.3)	0.34	4857 (91.8)	4944 (93.4) (92.7–94.1)	−87 (−1.6)	0.21
	1–3	2180 (27.1)	2038 (93.5)	2034 (93.3) (92.2–94.3)	4 (0.2)	0.92	1921 (88.1)	1913 (87.8) (86.3–89.1)	8 (0.3)	0.86
	4–9	532 (6.6)	473 (88.9)	455 (85.5) (82.2–88.4)	18 (3.4)	0.41	423 (79.6)	392 (73.7) (69.7–77.4)	31 (5.9)	0.13
	9–10	29 (0.4)	25 (85.3)	24 (82.8) (64.2–94.2)	1 (2.5)	0.89	21 (73.8)	19 (65.5) (45.7–82.1)	2 (8.3)	0.60

Table 2 (continued)

Predictors	Categories	*N* (%)	5-year BCSS				10-year BCSS			
			Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value	Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value
ER status	Unknown	116 (1.4)	112 (96.4)	115 (99.1) (95.3–100.0)	− 3 (− 2.7)	0.76	108 (93.4)	112 (96.6) (91.4–99.1)	− 4 (− 3.2)	0.73
	Positive	6535 (81.4)	6221 (95.2)	6324 (96.8) (96.3–97.2)	− 103 (− 1.6)	0.19	5966 (91.3)	6023 (92.2) (91.5–92.8)	− 57 (− 0.9)	0.46
	Negative	1381 (17.2)	1257 (91.0)	1195 (86.5) (84.6–88.3)	62 (4.5)	0.08	1150 (83.3)	1133 (82.0) (79.9–84.0)	17 (1.3)	0.61
PR status	Unknown	507 (6.3)	482 (95.0)	486 (95.9) (93.7–97.4)	− 4 (− 0.9)	0.84	461 (90.9)	457 (90.1) (87.2–92.6)	4 (0.8)	0.86
	Positive	5109 (63.6)	4874 (95.4)	4962 (97.1) (96.6–97.6)	− 88 (− 1.7)	0.21	4685 (91.7)	4751 (93.0) (92.3–93.7)	− 66 (− 1.3)	0.33
	Negative	2416 (30.1)	2232 (92.4)	2186 (90.5) (89.2–91.6)	46 (1.9)	0.33	2080 (86.1)	2060 (85.3) (83.8–86.7)	20 (0.8)	0.65
HER2 status	Unknown	1413 (17.6)	1342 (95.0)	1339 (94.8) (93.5–95.9)	3 (0.2)	0.92	1286 (91.0)	1280 (90.6) (88.9–92.1)	6 (0.4)	0.86
	Positive	1000 (12.5)	915 (91.5)	932 (93.2) (91.5–94.7)	− 17 (− 1.7)	0.57	842 (84.2)	872 (87.2) (85.0–89.2)	− 30 (− –3.0)	0.30
	Negative	5619 (70.0)	5332 (94.9)	5363 (95.4) (94.9–96.0)	− 31 (− 0.5)	0.67	5096 (90.7)	5116 (91.1) (90.3–91.8)	− 20 (− 0.4)	0.78
HR/HER2 status	Unknown	1417 (18.3)	1396 (94.9)	1391 (94.6) (93.4–95.7)	5 (0.3)	0.89	1335 (90.8)	1329 (90.4) (88.8–91.9)	6 (0.4)	0.86
	HR+/HER2−	4827 (60.1)	4606 (95.4)	4679 (96.9) (96.4–97.4)	− 73 (− 1.5)	0.28	4423 (91.6)	4464 (92.5) (91.7–93.2)	− 41 (− 0.8)	0.54
	HR+/HER2+	646 (8.0)	601 (93.1)	620 (96.0) (94.5–97.5)	− 19 (− 2.9)	0.44	563 (87.2)	579 (89.6) (87.3–92.0)	− 16 (− 2.4)	0.50
	HER2 positive	341 (4.3)	302 (88.4)	301 (88.3) (84.8–91.7)	1 (0.1)	1.00	268 (78.7)	282 (82.7) (78.7–86.7)	− 14 (− 4.0)	0.39
	Triple negative	747 (9.3)	686 (91.8)	643 (86.1) (83.6–88.6)	43 (5.7)	0.10	634 (84.8)	614 (82.2) (79.4–84.9)	20 (2.6)	0.43
Histological type	Ductal	6526 (81.3)	6154 (94.3)	6186 (94.8) (94.2–95.3)	− 32 (− 0.5)	0.68	5841 (89.5)	5876 (90.0) (89.3–90.8)	− 35 (− 0.5)	0.65
	Lobular	790 (9.8)	748 (94.7)	758 (96.0) (94.3–97.2)	− 10 (− 1.3)	0.72	713 (90.3)	726 (91.9) (89.8–93.7)	− 13 (− 1.6)	0.64
	Ductal and lobular	289 (3.6)	276 (95.5)	281 (97.2) (94.6–98.8)	− 5 (− 1.7)	0.76	265 (91.8)	266 (92.0) (88.3–94.9)	− 1 (− 0.2)	1.00
	Other	427 (5.3)	414 (97.0)	409 (95.8) (93.4–97.5)	5 (1.2)	0.79	404 (94.6)	400 (93.7) (90.9–95.8)	4 (0.9)	0.84

Table 2 (continued)

Predictors	Categories	*N* (%)	5-year BCSS				10-year BCSS			
			Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value	Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value
Grade	Unknown	317 (3,9)	300 (94.6)	303 (95.6) (92.7–97.6)	− 3 (− 1.0)	0.86	286 (90.1)	289 (91.2) (87.5–94.1)	− 3 (− 1.1)	0.84
	I	1854 (23.1)	1804 (97.3)	1835 (99.0) (98.4–99.4)	− 31 (− 1.7)	0.47	1763 (95.1)	1801 (97.1) (96.3–97.9)	− 38 (− 2.0)	0.37
	II	3488 (43.4)	3303 (94.7)	3376 (96.8) (96.1–97.3)	− 73 (− 2.1)	0.20	3150 (90.3)	3217 (92.2) (91.3–93.1)	− 67 (− 1.9)	0.23
	III	2373 (29.5)	2186 (92.1)	2120 (89.3) (88.0–90.6)	66 (2.8)	0.16	2027 (85.4)	1961 (82.6) (81.1–84.1)	66 (2.8)	0.15
Hormonal therapy	None	4564 (56.8)	4327 (94.8)	4331 (94.9) (94.2–95.6)	− 4 (− 0.1)	1.00	4126 (90.4)	4183 (91.7) (90.8–92.4)	− 57 (− 1.3)	0.37
	Tamoxifen	853 (10.6)	806 (94.5)	796 (93.3) (91.4–94.9)	10 (1.2)	0.72	768 (90.0)	747 (87.6) (85.2–89.7)	21 (2.4)	0.45
	Aromatase inhibitor	1724 (21.5)	1614 (93.6)	1638 (95.0) (93.9–96.0)	− 24 (− 1.4)	0.54	1524 (88.4)	1529 (88.7) (87.1–90.1)	− 5 (− 0.3)	0.89
	Tamoxifen to aromatase inhibitor	659 (8.2)	624 (94.7)	641 (97.3) (95.7–98.4)	− 17 (− 2.6)	0.50	594 (90.1)	601 (91.2) (88.8–93.2)	− 7 (− 1.1)	0.76
	Ovarian ablation	3 (0.04)	3 (100.0)	3 (100.0) (0.29–1.00)	0 (0.00)	1.00	3 (100.0)	3 (100.0) (29.2–1.00)	0 (− 5.8)	0.92
	Ovarian ablation and tamoxifen	229 (2.9)	218 (95.3)	225 (98.3) (0.96–1.00)	− 7 (− 3.0)	0.65	208 (91.0)	205 (89.5) (84.8–93.2)	3 (1.5)	0.81
Chemotherapy	None	5363 (66.8)	5100 (95.1)	5156 (96.1) (95.6–96.6)	− 56 (− 1.0)	0.43	4886 (91.1)	4947 (92.2) (91.5–92.9)	− 61 (− 1.1)	0.38
	Generation 1	1158 (14.4)	1069 (92.3)	1061 (91.6) (89.9–93.2)	8 (0.7)	0.81	989 (85.4)	999 (86.3) (84.2–88.2)	− 10 (− 0.9)	0.75
	Generation 2	673 (8.4)	639 (95.0)	640 (95.1) (93.2–96.6)	− 1 (− 0.1)	1.00	608 (90.4)	599 (89.0) (86.4–91.3)	9 (1.4)	0.71
	Generation 3	151 (1.9)	142 (94.1)	139 (92.1) (86.5–95.8)	3 (2.0)	0.79	134 (88.7)	129 (85.4) (78.8–90.6)	5 (3.3)	0.67
	Generation unknown	687 (8.6)	643 (93.6)	638 (92.9) (90.7–94.7)	5 (0.7)	0.84	603 (87.8)	594 (86.5) (83.7–88.9)	9 (1.3)	0.71

Table 2 (continued)

Predictors	Categories	*N* (%)	5-year BCSS				10-year BCSS			
			Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value	Predicted (%)	Observed (%) (95% CI)	Difference (%)	*p* value
Adjuvant systemic therapy	None	3712 (46.2)	3552 (95.7)	3603 (97.1) (96.5–97.6)	−51 (−1.4)	0.40	3422 (92.2)	3490 (94.0) (93.2–94.8)	−68 (−1.8)	0.25
	Both	1817 (22.6)	1719 (94.6)	1750 (96.3) (95.3–97.1)	−31 (−1.7)	0.45	1632 (89.8)	1628 (89.6) (88.1–91.0)	4 (0.2)	0.92
	Only hormonal therapy	1651 (20.6)	1545 (93.6)	1553 (94.1) (92.8–95.2)	−8 (−0.5)	0.84	1464 (88.7)	1457 (88.3) (86.6–89.8)	7 (0.4)	0.84
	Only chemotherapy	852 (10.6)	774 (90.8)	728 (85.5) (82.9–87.7)	46 (5.3)	0.10	705 (82.7)	693 (81.3) (78.6–83.9)	12 (1.4)	0.66
Targeted therapy	No	7609 (94.7)	7203 (94.7)	7236 (95.1) (94.6–95.6)	−33 (−0.4)	0.70	6868 (90.3)	6892 (90.6) (89.9–91.2)	−24 (−0.3)	0.78
	Yes	423 (5.3)	388 (91.6)	398 (94.1) (91.8–96.3)	−10 (−2.4)	0.61	356 (84.1)	376 (88.9) (85.9–91.9)	−20 (−4.7)	0.29

N total number, *BCSS* breast cancer-specific survival, *CI* confidence interval, *ER* oestrogen receptor, *PR* progesterone receptor, *HER2* human epidermal growth factor receptor 2, *n.a* not applicable due to low number of patients

half of the women did not receive adjuvant systemic treatment (Table 1).

Calibration: overall survival

In the entire population, predicted 5-year OS was 87.0% compared to an observed OS of 89.2%. The 2.2% difference was statistically significant ($p = 0.032$), but not outside the range of 3% (Table 1, Fig. 1). CancerMath significantly underestimated 5-year OS in the following subgroups: patients with a tumour diameter 1.01–2.00 cm, no positive lymph nodes, grade II tumours, a positive ER and PR status, patients not receiving adjuvant systemic therapy and patients treated with targeted therapy (Table 1).

Predicted 10-year OS was 74.1%, compared to an observed OS of 76.0%. The difference of 1.9% was borderline significant ($p = 0.05$) (Table 1, Fig. 1). CancerMath significantly underestimated 10-year OS in women with at least one of the following characteristics: between 60 and 69 years, a tumour diameter of 0.51–1.00 cm, no positive lymph nodes, grade II tumours, and in case no adjuvant systemic therapy was administered. Besides, the model overestimated 10-year OS in women with 4–9 positive lymph nodes (Table 1).

Calibration: breast cancer-specific survival

Overall, predicted 5-year BCSS was 94.5% compared to an observed BCSS of 95.0%, indicating an underestimation of 0.5%, which was not statistically significant ($p = 0.62$). Predicted 10-year BCSS was 89.9% compared to an observed BCSS of 90.5%, indicating an underestimation of 0.6%, which was also not statistically significant ($p = 0.58$) (Table 2, Fig. 2). CancerMath estimated both 5- and 10-year BCSS accurate in all predefined subgroups (Table 2).

Discrimination: overall survival

In the entire population, the AUC for 5-year OS was 0.77 (95% CI 0.75–0.79) (Table 3, Fig. 3), varying from 0.37 to 0.87 in all subgroups. The discriminatory accuracy for 5-year OS was lowest in women with a tumour diameter < 0.11 cm and women who received adjuvant chemotherapy without endocrine therapy. (AUCs 0.37 and 0.59, respectively) (Table 3).

The AUC for 10-year OS was 0.77 (95% CI 0.76–0.78) in the entire population (Table 3, Fig. 3), ranging from 0.58 to 0.86 in the subgroups. The lowest AUC (0.58) was found in women treated with adjuvant chemotherapy without endocrine therapy (Table 3).

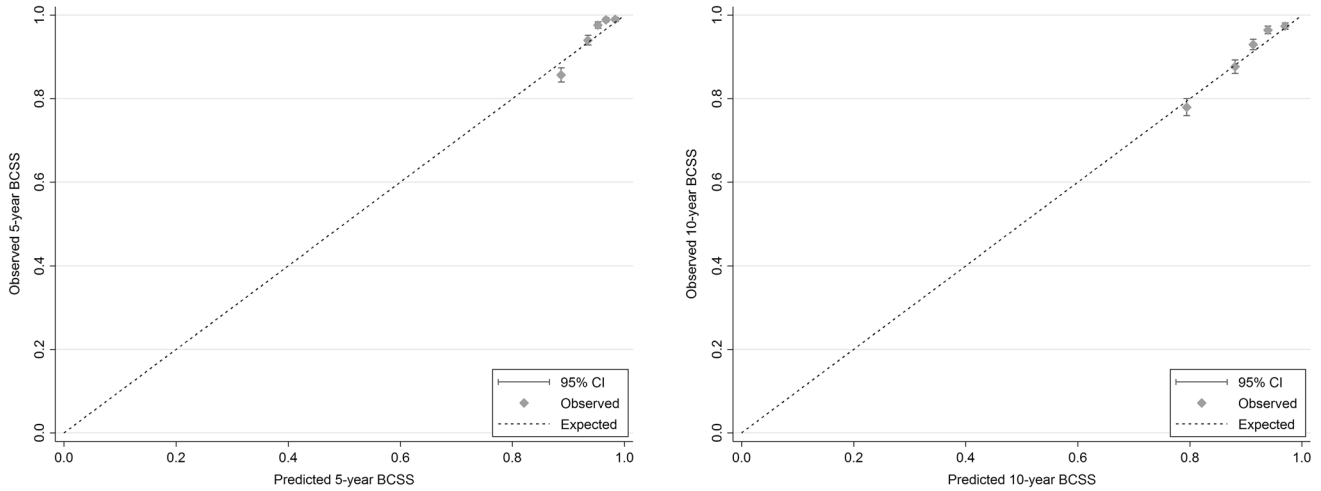

Fig. 2 Observed and predicted 5-year (left) and 10-year (right) breast cancer-specific survival in Dutch non-metastatic breast cancer patients diagnosed in 2005

Discrimination: breast cancer-specific survival

Overall, the AUC for 5-year BCSS was 0.78 (95% CI 0.76–0.80) (Table 4, Fig. 4), ranging from 0.61 to 0.99 in the subgroups, with the lowest AUC found in patients treated with adjuvant chemotherapy without endocrine therapy (Table 4).

In the entire population, the AUC for 10-year BCSS was 0.73 (95% CI 0.72–0.75) (Table 4, Fig. 4), ranging from 0.60 to 0.86 in the subgroups, with the lowest AUC found in patients with 4–9 positive lymph nodes (Table 4).

Discussion

This study showed that CancerMath accurately predicts OS in most subgroups, and accurately predicts BCSS in all subgroups of the Dutch validation population. OS should be interpreted with caution in several subgroups, but it predicted accurately in the subgroups in which PREDICT [9] was less accurate (ER-negative disease, patients > 79 years, T3 tumours and patients treated with both endocrine therapy and chemotherapy). For BCSS, CancerMath can safely be used in the entire population. CancerMath predicted OS less accurate in grade II tumours, patients without positive lymph nodes, a tumour diameter of 1.01–2.00 cm, a positive hormonal receptor status and patients aged 60–69 years. The fact that CancerMath predicts BCSS well in the entire population, but that OS is not accurately estimated (mostly underestimated) in several subgroups, may partly be explained by the difference in life expectancy, which is higher in the Netherlands than in the United States [25]. Another possible reason for the underestimation in hormonal receptor positive disease may be the different definition of positivity. In the Netherlands, hormonal receptor positivity is defined as ≥ 10% positive nuclei, while in the US a cut-off of 1% is used. This may have resulted in a more favourable group of hormonal receptor positive disease in the Dutch validation population, and consequently an underestimation of OS. In our study, over 80% of the patients had ER-positive breast cancer, while not even half of the patients received adjuvant endocrine therapy. It may feel counterintuitive that CancerMath underestimates 5-year OS in ER-positive disease even when there is less endocrine therapy given than expected. Further analysis of our data shows that this specific patient group mainly consists of older patients with favourable tumour characteristics, which may in part explain this counterintuitive result. Discriminatory accuracy of the model was lowest in patients treated with chemotherapy without endocrine treatment, which may partly be a reflection of triple-negative disease, which showed an overestimation of BCSS (although not significant). Notably, it is difficult to interpret predicted treatment effects in patients treated with adjuvant systemic therapy, as patients included in the validation already received treatment, independent of the model's prediction.

In the United Kingdom, a significant overestimation of 10-year BCSS was observed [18], contrary to our study, which may possibly be explained by the different populations studied, or by the different definition of BCSS. In Southeast Asia, the model underestimated 5-year OS within almost all subgroups [26]. However, in the subgroups in which we observed an underestimation of 5-year OS, the model predicted accurately for the Southeast Asian population. These discrepancies may possibly be explained by the fact that Asian patients are more likely to have a poorer breast cancer prognosis as compared to western breast cancer patients [27].

Table 3 Discriminatory performance of CancerMath of 5- and 10-year overall survival by subgroups of predictors in CancerMath

Predictors	Categories	*N* (%)	5-year OS AUC (95% CI)[a]	10-year OS AUC (95% CI)[a]
Entire cohort		8032 (100.0)	0.77 (0.75–0.79)	0.77 (0.76–0.78)
Age (years)	< 40	420 (5.2)	0.74 (0.65–0.83)	0.66 (0.59–0.73)
	40–49	1526 (19.0)	0.75 (0.69–0.80)	0.69 (0.65–0.73)
	50–59	2229 (27.8)	0.74 (0.69–0.78)	0.70 (0.67–0.73)
	60–69	1929 (24.0)	0.70 (0.65–0.74)	0.65 (0.62–0.68)
	70–79	1312 (16.3)	0.71 (0.67–0.75)	0.65 (0.62–0.68)
	> 79	616 (7.7)	0.67 (0.62–0.71)	0.65 (0.60–0.70)
Tumour diameter (cm)	< 0.1	39 (0.5)	0.37 (n.a.)	0.82 (0.62–1.00)
	0.11–0.5	256 (3.2)	0.77 (0.64–0.91)	0.70 (0.59–0.82)
	0.51–1.00	1211 (15.1)	0.70 (0.63–0.77)	0.71 (0.67–0.75)
	1.01–2.00	3678 (45.8)	0.74 (0.71–0.77)	0.75 (0.73–0.77)
	2.01–3.00	2088 (26.0)	0.74 (0.71–0.77)	0.76 (0.74–0.78)
	3.01–4.00	585 (7.3)	0.75 (0.70–0.80)	0.78 (0.74–0.82)
	4.01–5.00	175 (2.2)	0.75 (0.66–0.83)	0.78 (0.72–0.85)
Number of positive nodes	0	5291 (65.9)	0.75 (0.73–0.78)	0.76 (0.74–0.77)
	1–3	2180 (27.1)	0.78 (0.75–0.81)	0.78 (0.76–0.81)
	4–9	532 (6.6)	0.72 (0.66–0.77)	0.72 (0.68–0.77)
	9–10	29 (0.4)	0.87 (0.74–1.00)	0.82 (0.65–1.00)
ER status	Unknown	116 (1.4)	0.77 (0.58–0.96)	0.86 (0.77–0.95)
	Positive	6535 (81.4)	0.79 (0.77–0.81)	0.77 (0.76–0.79)
	Negative	1381 (17.2)	0.68 (0.64–0.72)	0.72 (0.69–0.75)
PR status	Unknown	507 (6.3)	0.75 (0.68–0.83)	0.78 (0.73–0.83)
	Positive	5109 (63.6)	0.79 (0.77–0.81)	0.78 (0.76–0.80)
	Negative	2416 (30.1)	0.72 (0.69–0.75)	0.73 (0.71–0.75)
HER2 status	Unknown	1413 (17.6)	0.81 (0.77–0.84)	0.81 (0.79–0.84)
	Positive	1000 (12.5)	0.79 (0.74–0.83)	0.75 (0.71–0.78)
	Negative	5619 (70.0)	0.76 (0.73–0.78)	0.76 (0.75–0.78)
HR/HER2 status	Unknown	1417 (18.3)	0.80 (0.77–0.84)	0.82 (0.79–0.84)
	HR+/HER2−	4827 (60.1)	0.78 (0.75–0.80)	0.77 (0.75–0.78)
	HR+/HER2+	646 (8.0)	0.81 (0.75–0.87)	0.76 (0.71–0.80)
	HER2 positive	341 (4.3)	0.71 (0.63–0.79)	0.70 (0.63–0.77)
	Triple negative	747 (9.3)	0.65 (0.60–0.70)	0.72 (0.67–0.76)
Histological type	Ductal	6526 (81.3)	0.77 (0.76–0.79)	0.77 (0.75–0.78)
	Lobular	790 (9.8)	0.77 (0.72–0.83)	0.80 (0.76–0.83)
	Ductal and lobular	289 (3.6)	0.72 (0.60–0.85)	0.74 (0.66–0.81)
	Other	427 (5.3)	0.73 (0.66–0.81)	0.78 (0.73–0.84)
Grade	Unknown	317 (3.9)	0.77 (0.68–0.85)	0.76 (0.70–0.82)
	I	1854 (23.1)	0.80 (0.75–0.85)	0.79 (0.77–0.82)
	II	3488 (43.4)	0.79 (0.77–0.82)	0.78 (0.76–0.80)
	III	2373 (29.5)	0.71 (0.68–0.74)	0.72 (0.70–0.74)
Hormonal therapy	None	4564 (56.8)	0.76 (0.73–0.78)	0.76 (0.74–0.78)
	Tamoxifen	853 (10.6)	0.67 (0.60–0.74)	0.75 (0.71–0.78)
	Aromatase inhibitor	1724 (21.5)	0.79 (0.76–0.82)	0.79 (0.76–0.81)
	Tamoxifen to aromatase inhibitor	659 (8.2)	0.71 (0.59–0.83)	0.71 (0.65–0.78)
	Ovarian ablation	3 (0.04)	n.a.	n.a.
	Ovarian ablation and tamoxifen	229 (2.9)	0.70 (0.47–0.94)	0.73 (0.62–0.83)

Table 3 (continued)

Predictors	Categories	*N* (%)	5-year OS AUC (95% CI)[a]	10-year OS AUC (95% CI)[a]
Chemotherapy	None	5363 (66.8)	0.81 (0.79–0.82)	0.80 (0.78–0.81)
	Generation 1	1158 (14.4)	0.66 (0.61–0.71)	0.65 (0.61–0.69)
	Generation 2	673 (8.4)	0.69 (0.60–0.78)	0.67 (0.61–0.72)
	Generation 3	151 (1.9)	0.70 (0.60–0.81)	0.63 (0.53–0.74)
	Generation unknown	687 (8.6)	0.72 (0.65–0.79)	0.72 (0.67–0.77)
Adjuvant systemic therapy	None	3712 (46.2)	0.81 (0.78–0.83)	0.80 (0.78–0.83)
	Both	1817 (22.6)	0.67 (0.61–0.73)	0.67 (0.61–0.72)
	Only hormonal therapy	1651 (20.6)	0.77 (0.74–0.80)	0.77 (0.74–0.80)
	Only chemotherapy	852 (10.6)	0.59 (0.54–0.64)	0.58 (0.53–0.63)
Targeted therapy	No	7909 (94.7)	0.77 (0.75–0.79)	0.77 (0.76–0.78)
	Yes	423 (5.3)	0.75 (0.67–0.94)	0.65 (0.58–0.73)

N total number, *OS* overall survival, *AUC* area under the curve, *CI* confidence interval, *ER* oestrogen receptor, *PR* progesterone receptor, *HER*2 human epidermal growth factor receptor 2, *n.a* not applicable due to low number of patients

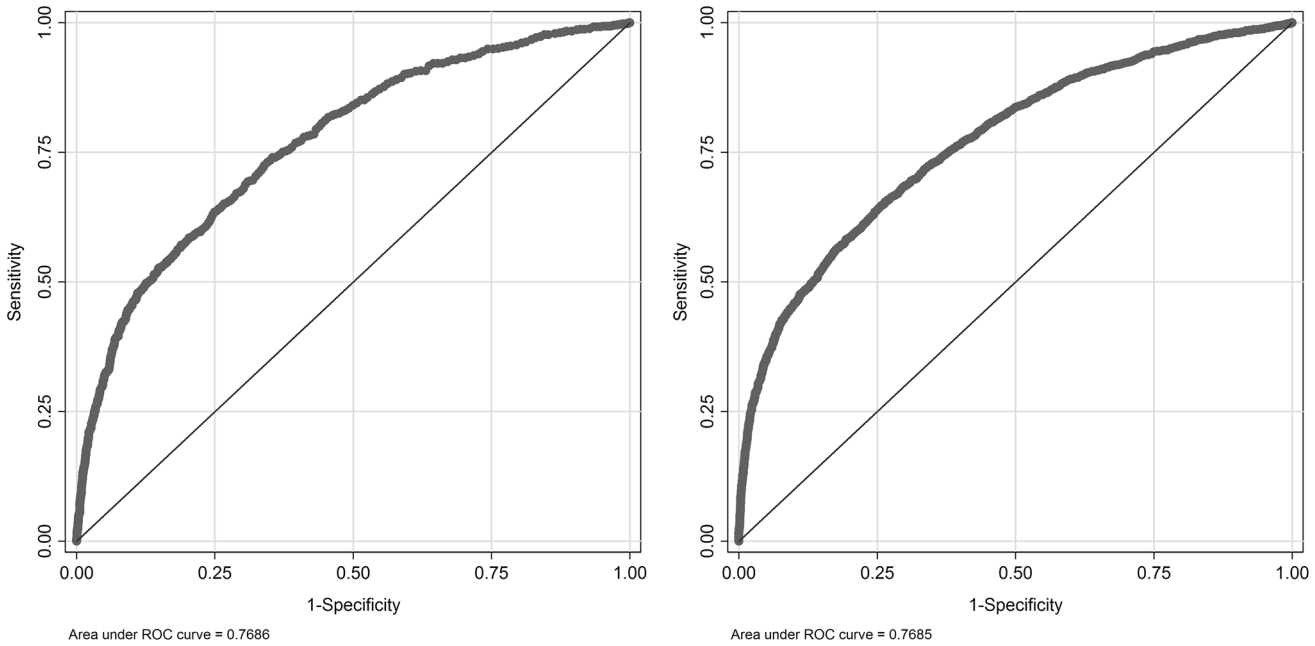

Fig. 3 Discriminatory accuracy of 5-year (left) and 10-year (right) overall survival for in Dutch non-metastatic breast cancer patients diagnosed in 2005. *ROC* receiver operating characteristic curve

Currently, a frequently used online prediction model is PREDICT [28], which has been validated in several countries [9–18]. It was shown that PREDICT has a lower predictive accuracy in the youngest and oldest Dutch women, in larger tumours and in patients treated with both endocrine therapy and chemotherapy [15–17]. Our study showed that CancerMath accurately estimated both OS and BCSS in all of these subgroups.

A statistically significant difference between predicted and observed outcomes of > 3% was considered clinically relevant. This was based on the Dutch guideline in which it is stated that a 10-year survival benefit of 3-5% is an indication for adjuvant systemic therapy [10]. Importantly, a 3% difference may not be clinically relevant for someone with a predicted benefit of 12% and an observed benefit of 15%, as adjuvant systemic therapy will be administered in both cases. However, in case these percentages are 1% and 4%,

Table 4 Discriminatory performance of CancerMath of 5- and 10-year breast cancer-specific survival by subgroups of predictors in CancerMath

Predictors	Categories	*N* (%)	5-year BCSS AUC (95% CI)[a]	10-year BCSS AUC (95% CI)[a]
Entire cohort		8032 (100.0)	0.78 (0.76–0.80)	0.73 (0.72–0.75)
Age (years)	< 40	420 (5.2)	0.75 (0.66–0.84)	0.67 (0.59–0.74)
	40–49	1526 (19.0)	0.76 (0.70–0.82)	0.72 (0.68–0.76)
	50–59	2229 (27.8)	0.80 (0.75–0.84)	0.76 (0.72–0.79)
	60–69	1929 (24.0)	0.79 (0.74–0.83)	0.75 (0.71–0.78)
	70–79	1312 (16.3)	0.79 (0.75–0.84)	0.75 (0.71–0.79)
	> 79	616 (7.7)	0.72 (0.65–0.79)	0.69 (0.63–0.76)
Tumour diameter (cm)	< 0.1	39 (0.5)	n.a.	n.a.
	0.11–0.5	256 (3.2)	0.99 (0.99–1.00)	0.81 (0.64–0.98)
	0.51–1.00	1211 (15.1)	0.75 (0.62–0.89)	0.70 (0.62–0.78)
	1.01–2.00	3678 (45.8)	0.74 (0.69–0.79)	0.71 (0.67–0.74)
	2.01–3.00	2088 (26.0)	0.71 (0.68–0.75)	0.66 (0.63–0.70)
	3.01–4.00	585 (7.3)	0.67 (0.60–0.73)	0.63 (0.57–0.69)
	4.01–5.00	175 (2.2)	0.65 (0.52–0.79)	0.65 (0.55–0.75)
Number of positive nodes	0	5291 (65.9)	0.75 (0.71–0.79)	0.69 (0.67–0.72)
	1–3	2180 (27.1)	0.76 (0.73–0.80)	0.71 (0.68–0.74)
	4–9	532 (6.6)	0.64 (0.57–0.71)	0.60 (0.54–0.65)
	9–10	29 (0.4)	0.80 (0.63–0.98)	0.64 (0.42–0.87)
ER status	Unknown	116 (1.4)	0.90 (n.a.)	0.86 (0.73–1.00)
	Positive	6535 (81.4)	0.76 (0.73–0.79)	0.73 (0.71–0.75)
	Negative	1381 (17.2)	0.65 (0.61–0.69)	0.65 (0.61–0.68)
PR status	Unknown	507 (6.3)	0.74 (0.62–0.86)	0.71 (0.64–0.78)
	Positive	5109 (63.6)	0.77 (0.74–0.81)	0.73 (0.70–0.76)
	Negative	2416 (30.1)	0.71 (0.68–0.74)	0.68 (0.66–0.71)
HER2 status	Unknown	1413 (17.6)	0.79 (0.74–0.85)	0.74 (0.69–0.79)
	Positive	1000 (12.5)	0.77 (0.72–0.83)	0.68 (0.64–0.73)
	Negative	5619 (70.0)	0.78 (0.75–0.81)	0.74 (0.72–0.76)
HR/HER2 status	Unknown	1417 (18.3)	0.80 (0.75–0.86)	0.75 (0.70–0.79)
	HR+/HER2−	4827 (60.1)	0.77 (0.73–0.80)	0.74 (0.71–0.76)
	HR+/HER2+	646 (8.0)	0.77 (0.68–0.86)	0.66 (0.60–0.72)
	HER2 positive	341 (4.3)	0.68 (0.59–0.77)	0.66 (0.58–0.73)
	Triple negative	747 (9.3)	0.63 (0.58–0.68)	0.63 (0.58–0.68)
Histological type	Ductal	6526 (81.3)	0.79 (0.76–0.81)	0.74 (0.72–0.76)
	Lobular	790 (9.8)	0.72 (0.64–0.81)	0.71 (0.65–0.77)
	Ductal and lobular	289 (3.6)	0.84 (0.71–0.98)	0.73 (0.62–0.83)
	Other	427 (5.3)	0.71 (0.58–0.84)	0.69 (0.58–0.80)
Grade	Unknown	317 (3.9)	0.86 (0.80–0.93)	0.76 (0.67–0.84)
	I	1854 (23.1)	0.68 (0.55–0.81)	0.67 (0.59–0.74)
	II	3488 (43.4)	0.77 (0.72–0.81)	0.70 (0.67–0.74)
	III	2373 (29.5)	0.66 (0.63–0.69)	0.63 (0.60–0.66)
Hormonal therapy	None	4564 (56.8)	0.82 (0.79–0.84)	0.76 (0.74–0.79)
	Tamoxifen	853 (10.6)	0.73 (0.60–0.74)	0.70 (0.65–0.75)
	Aromatase inhibitor	1724 (21.5)	0.73 (0.68–0.78)	0.70 (0.66–0.74)
	Tamoxifen to aromatase inhibitor	659 (8.2)	0.75 (0.64–0.86)	0.62 (0.54–0.70)
	Ovarian ablation	3 (0.04)	n.a.	n.a.
	Ovarian ablation and tamoxifen	229 (2.9)	0.69 (0.43–0.94)	0.73 (0.62–0.83)

Table 4 (continued)

Predictors	Categories	*N* (%)	5-year BCSS AUC (95% CI)[a]	10-year BCSS AUC (95% CI)[a]
Chemotherapy	None	5363 (66.8)	0.81 (0.78–0.84)	0.75 (0.73–0.77)
	Generation 1	1158 (14.4)	0.67 (0.61–0.72)	0.65 (0.61–0.69)
	Generation 2	673 (8.4)	0.71 (0.62–0.80)	0.68 (0.61–0.74)
	Generation 3	151 (1.9)	0.76 (0.64–0.87)	0.70 (0.59–0.81)
	Generation unknown	687 (8.6)	0.72 (0.65–0.79)	0.73 (0.68–0.78)
Adjuvant systemic therapy	None	3712 (46.2)	0.73 (0.70–0.77)	0.71 (0.68–0.75)
	Both	1817 (22.6)	0.69 (0.63–0.74)	0.69 (0.63–0.74)
	Only hormonal therapy	1651 (20.6)	0.65 (0.61–0.69)	0.61 (0.57–0.64)
	Only chemotherapy	852 (10.6)	0.61 (0.56–0.66)	0.61 (0.56–0.66)
Targeted therapy	No	7909 (94.7)	0.67 (0.64–0.70)	0.61 (0.60–0.63)
	Yes	423 (5.3)	0.76 (0.67–0.85)	0.67 (0.59–0.76)

N total number, *BCSS* breast cancer-specific survival, *AUC* area under the curve, *CI* confidence interval, *ER* oestrogen receptor, *PR* progesterone receptor, *HER*2 human epidermal growth factor receptor 2, *n.a.* not applicable due to low number of patients

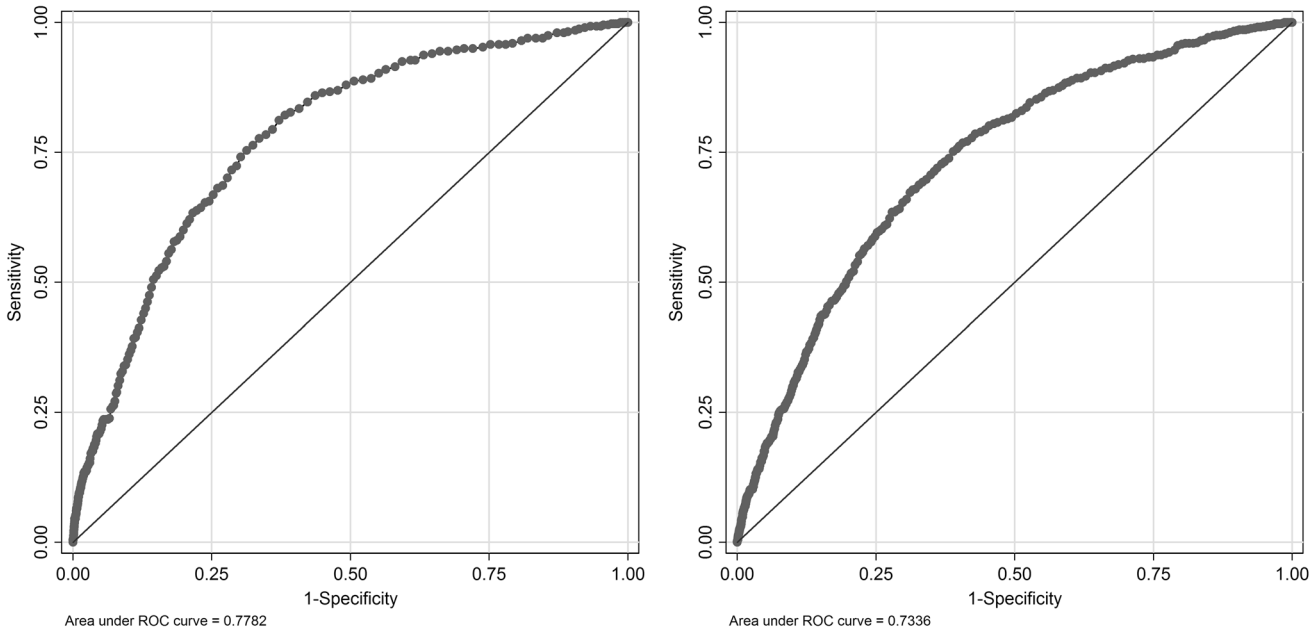

Fig. 4 Discriminatory accuracy of 5-year (left) and 10-year (right) breast cancer-specific survival for in Dutch non-metastatic breast cancer patients diagnosed in 2005. *ROC* receiver operating characteristic curve

the indication for adjuvant systemic therapy will change. It is therefore crucial to interpret the over- or underestimations for individual patients in light of the initial predicted survival without adjuvant systemic therapy.

Strengths and limitations

A strength of this study is its population-based character, increasing the reliability and generalisability of the results. A limitation of this study is that HER2 status determination and administration of trastuzumab was introduced in 2005. Consequently, not every patient with HER2-positive disease may have been treated with trastuzumab while they would have been treated accordingly today. This could

consequently have led to less adequate predictions for these patients. Furthermore, we did not have access to death certificates and could not determine whether patients died due to breast cancer. We now classified deceased patients as breast cancer-related when they experienced a distant metastasis (85% of all deceased patients experienced a distant metastasis). As this is a proxy, outcomes should be interpreted with care. Lastly, for several subgroups the confidence intervals around the AUCs were wide, making it difficult to draw definite conclusions. It is therefore recommended for future validation studies to ensure sufficient numbers of patients with specific characteristics.

Implications for clinical practice

CancerMath predicts OS accurate in the groups where PREDICT was less accurate [15–17], meaning that CancerMath is a reliable complement to PREDICT in clinical practice, or may even be favoured over PREDICT as it also estimates breast cancer-specific survival accurately. A possible advantage of CancerMath may be that this model provides the absolute difference in life expectancy in years, which may be preferred by some patients. However, a limitation of the model is the lack of predictions of the use of trastuzumab. As we now have two prediction models which still have their limitations but both predict accurate in the entire Dutch population, and either PREDICT or CancerMath may be favoured in some subgroups, we should focus on further improving these models. Taking into account that more and more patients are nowadays treated with neoadjuvant instead of adjuvant chemotherapy, and that neither PREDICT nor CancerMath was developed on this patient group, these models should be updated for patients treated with neoadjuvant systemic therapy. Furthermore, it would be of great relevance to ask patients which type or model they favour in terms of the presentation of the outcomes, or if they miss any information such as long-term side effects. Based on patient experiences and more information on for example breast cancer subtypes and outcomes of genomic tests, we should be able to integrate this knowledge in an updated prediction model that is valid for all specific patient groups.

Conclusions

CancerMath accurately predicts OS in most subgroups, and accurately predicts BCSS in all subgroups of the Dutch validation population. It may function as a valuable complement to PREDICT in patient groups in which PREDICT performed less accurate. As CancerMath also provides the absolute gain in life expectancy in years, it may even be favoured over PREDICT by some patients, but this should be subject of further research. The shift to more neoadjuvant systemic treatment, the reduction in axillary dissections, the increasing use of genomic tests and the growing knowledge on breast cancer subtypes poses a challenge in updating existing prediction models.

Author contributions LH and MCM analysed and interpreted the data in this study, and wrote the manuscript. SS, LH and MCM designed the study. TH helped with the statistical analysis. All authors helped interpreting the results, helped writing the manuscript and approved the final manuscript.

Funding The authors received no specific funding for this work.

Data availability The datasets generated during and/or analysed during the current study are not publicly available due to the strict privacy regulation of the Netherlands Cancer Registry but are available in aggregated form from the corresponding author on reasonable request.

Compliance with ethical standards

Conflict of interest Dr. Gabe Sonke receives institutional research support from AstraZeneca, Merck, Novartis, and Roche. All other authors declare that they have no conflicts of interest.

Informed consent This study was approved by the privacy committee (CvT) of the Netherlands Cancer Registry, and complies with the current laws in the Netherlands.

Research involving human participants and/or animals All procedures performed in studies involving human participants were in accordance with the ethical standards of the privacy committee (CvT) of the Netherlands Cancer Registry and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

1. Early Breast Cancer Trialists' Collaborative Group (2008) Adjuvant chemotherapy in oestrogen-receptor-poor breast cancer: patient-level meta-analysis of randomised trials. Lancet 371(9606):29–40
2. 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(9472):1687–1717
3. Clarke M (2006) Meta-analyses of adjuvant therapies for women with early breast cancer: the Early Breast Cancer Trialists' Collaborative Group overview. Ann Oncol 17(suppl_10):x59–x62
4. Bertos NR, Park M (2011) Breast cancer—one term, many entities? J Clin Invest 121(10):3789–3796
5. Higgins MJ, Baselga J (2011) Targeted therapies for breast cancer. J Clin Invest 121(10):3797–3803
6. Prat A, Perou CM (2011) Deconstructing the molecular portraits of breast cancer. Mol Oncol 5(1):5–23
7. Laas E, Mallon P, Delomenie M, Gardeux V, Pierga J, Cottu P, Lerebours F, Stevens D, Rouzier R, Reyal F (2015) Are we able to predict survival in ER-positive HER2-negative breast cancer? A comparison of web-based models. Br J Cancer 112(5):912

8. Engelhardt EG, Garvelink MM, de Haes JC, 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
9. dos Reis FJC, Wishart GC, Dicks EM, Greenberg D, Rashbass J, Schmidt MK, van den Broek AJ, Ellis IO, Green A, Rakha E (2017) An updated PREDICT breast cancer prognostication and treatment benefit prediction model with independent validation. Breast Cancer Res 19(1):58
10. Wishart G, Bajdik C, Azzato E, Dicks E, Greenberg D, Rashbass J, Caldas C, Pharoah P (2011) A population-based validation of the prognostic model PREDICT for early breast cancer. Eur J Surg Oncol 37(5):411–417
11. Wishart G, Bajdik C, Dicks E, Provenzano E, Schmidt M, Sherman M, Greenberg D, Green A, Gelmon K, Kosma V (2012) PREDICT Plus: development and validation of a prognostic model for early breast cancer that includes HER2. Br J Cancer 107(5):800
12. Wishart GC, Rakha E, Green A, Ellis I, Ali HR, Provenzano E, Blows FM, Caldas C, Pharoah PD (2014) Inclusion of KI67 significantly improves performance of the PREDICT prognostication and prediction model for early breast cancer. BMC Cancer 14(1):908
13. Maishman T, Copson E, Stanton L, Gerty S, Dicks E, Durcan L, Wishart G, Pharoah P, Eccles D, Group PS (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
14. Wong H-S, Subramaniam S, Alias Z, Taib NA, Ho G-F, Ng C-H, Yip C-H, Verkooijen HM, Hartman M, Bhoo-Pathy N (2015) The predictive accuracy of PREDICT: a personalized decision-making tool for Southeast Asian women with breast cancer. Medicine 94(8):e593. https://doi.org/10.1097/MD.0000000000000593
15. De Glas N, Bastiaannet E, Engels C, De Craen A, Putter H, Van De Velde C, Hurria A, Liefers G, Portielje J (2016) Validity of the online PREDICT tool in older patients with breast cancer: a population-based study. Br J Cancer 114(4):395
16. van Maaren M, van Steenbeek C, Pharoah P, Witteveen A, Sonke GS, Strobbe L, Poortmans P, Siesling S (2017) Validation of the online prediction tool PREDICT v. 2.0 in the Dutch breast cancer population. Eur J Cancer 86:364–372
17. Engelhardt EG, van den Broek AJ, Linn SC, Wishart GC, Emiel JT, van de Velde AO, Smit VT, Voogd AC, Siesling S, Brinkhuis M (2017) Accuracy of the online prognostication tools PREDICT and Adjuvant! for early-stage breast cancer patients younger than 50 years. Eur J Cancer 78:37–44
18. Karapanagiotis S, Pharoah PD, Jackson CH, Newcombe PJ (2018) Development and external validation of prediction models for 10-year survival of invasive breast cancer comparison with PREDICT and CancerMath. Clin Cancer Res 24(9):2110–2115
19. Chen LL, Nolan ME, Silverstein MJ, Mihm MC, Sober AJ, Tanabe KK, Smith BL, Younger J, Michaelson JS (2009) The impact of primary tumor size, lymph node status, and other prognostic factors on the risk of cancer death. Cancer 115(21):5071–5083
20. Bickell NA, Weidmann J, Fei K, Lin JJ, Leventhal H (2009) Underuse of breast cancer adjuvant treatment: patient knowledge, beliefs, and medical mistrust. J Clin Oncol 27(31):5160–5167. https://doi.org/10.1200/JCO.2009.22.9773
21. Belkora JK, Hutton DW, Moore DH, Siminoff LA (2011) Does use of the adjuvant! model influence use of adjuvant therapy through better risk communication? J Natl Compr Canc Netw 9(7):707–712
22. Fritz AG (2000) International classification of diseases for oncology: ICD-O. World Health Organization, Geneva
23. Sobin LH, Gospodarowicz MK, Wittekind C (2011) TNM classification of malignant tumours. Wiley, New York
24. CancerMath Therapy Calculator: http://www.lifemath.net/cancer/breastcancer/therapy/index.php. Accessed 22 Jun 2018
25. World Health Organization (2018) Life expectancy and healthy life expectancy, data by country. http://apps.who.int/gho/data/node.main.688?lang=en
26. Miao H, Hartman M, Verkooijen HM, Taib NA, Wong H-S, Subramaniam S, Yip C-H, Tan E-Y, Chan P, Lee S-C (2016) Validation of the CancerMath prognostic tool for breast cancer in Southeast Asia. BMC Cancer 16(1):820
27. Telli ML, Chang ET, Kurian AW, Keegan TH, McClure LA, Lichtensztajn D, Ford JM, Gomez SL (2011) Asian ethnicity and breast cancer subtypes: a study from the California Cancer Registry. Breast Cancer Res Treat 127(2):471–478
28. 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. https://doi.org/10.1186/bcr2464