European Journal of Cancer 195 (2023) 113401

Contents lists available at ScienceDirect

European Journal of Cancer

journal homepage: www.ejcancer.com

External validation and clinical utility assessment of PREDICT breast cancer prognostic model in young, systemic treatment-naïve women with node-negative breast cancer

Yuwei Wang [a], Annegien Broeks [b], Daniele Giardiello [a,c], Michael Hauptmann [d], Katarzyna Jóźwiak [d], Esther A. Koop [e], Mark Opdam [a], Sabine Siesling [f,g], Gabe S. Sonke [h], Nikolas Stathonikos [i], Natalie D. ter Hoeve [i], Elsken van der Wall [j], Carolien H.M. van Deurzen [k], Paul J. van Diest [i], Adri C. Voogd [l], Willem Vreuls [m], Sabine C. Linn [a,h,i], Gwen M.H. E. Dackus [a,i,*,1], Marjanka K. Schmidt [a,n,*,1]

[a] Department of Molecular Pathology, the Netherlands Cancer Institute, Amsterdam, the Netherlands
[b] Core Facility Molecular Pathology and Biobanking, the Netherlands Cancer Institute, Amsterdam, the Netherlands
[c] Eurac Research, Institute of Biomedicine, Epidemiology and Biostatistics, Bolzano, Italy
[d] Institute of Biostatistics and Registry Research, Brandenburg Medical School Theodor Fontane, Neuruppin, Germany
[e] Department of Pathology, Gelre Ziekenhuizen, Apeldoorn, the Netherlands
[f] Department of Research and Development, Netherlands Comprehensive Cancer Organization, Utrecht, the Netherlands
[g] Department of Health Technology and Services Research, Technical Medical Centre, University of Twente, Enschede, the Netherlands
[h] Department of Medical Oncology, the Netherlands Cancer Institute, Amsterdam, the Netherlands
[i] Department of Pathology, University Medical Center Utrecht, Utrecht, the Netherlands
[j] Division of Internal Medicine and Dermatology, University Medical Center Utrecht, Utrecht, the Netherlands
[k] Department of Pathology, Erasmus MC Cancer Institute, Rotterdam, the Netherlands
[l] Department of Epidemiology, Maastricht University, Maastricht, the Netherlands
[m] Department of Pathology, Canisius Wilhelmina Ziekenhuis, Nijmegen, the Netherlands
[n] Department of Clinical Genetics, Leiden University Medical Center, Leiden, the Netherlands

ARTICLE INFO

Keywords:
PREDICT Breast
Breast cancer
Prognosis prediction
External validation
Clinical utility

ABSTRACT

Background: The validity of the PREDICT breast cancer prognostic model is unclear for young patients without adjuvant systemic treatment. This study aimed to validate PREDICT and assess its clinical utility in young women with node-negative breast cancer who did not receive systemic treatment.

Methods: We selected all women from the Netherlands Cancer Registry who were diagnosed with node-negative breast cancer under age 40 between 1989 and 2000, a period when adjuvant systemic treatment was not standard practice for women with node-negative disease. We evaluated the calibration and discrimination of PREDICT using the observed/expected (O/E) mortality ratio, and the area under the receiver operating characteristic curve (AUC), respectively. Additionally, we compared the potential clinical utility of PREDICT for selectively administering chemotherapy to the chemotherapy-to-all strategy using decision curve analysis at predefined thresholds.

Results: A total of 2264 women with a median age at diagnosis of 36 years were included. Of them, 71.2% had estrogen receptor (ER)-positive tumors and 44.0% had grade 3 tumors. Median tumor size was 16 mm. PREDICT v2.2 underestimated 10-year all-cause mortality by 33% in all women (O/E ratio:1.33, 95%CI:1.22–1.43). Model discrimination was moderate overall ($AUC_{10\text{-year}}$:0.65, 95%CI:0.62–0.68), and poor for women with ER-negative tumors ($AUC_{10\text{-year}}$:0.56, 95%CI:0.51–0.62). Compared to the chemotherapy-to-all strategy, PREDICT only showed a slightly higher net benefit in women with ER-positive tumors, but not in women with ER-negative tumors.

* Correspondence to: Division of Molecular Pathology, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands.
E-mail addresses: g.dackus@nki.nl (G.M.H.E. Dackus), mk.schmidt@nki.nl (M.K. Schmidt).
[1] Shared senior and corresponding authors.

https://doi.org/10.1016/j.ejca.2023.113401
Received 24 August 2023; Received in revised form 19 October 2023; Accepted 19 October 2023
Available online 30 October 2023

Conclusions: PREDICT yields unreliable predictions for young women with node-negative breast cancer. Further model updates are needed before PREDICT can be routinely used in this patient subset.

1. Introduction

The prognosis of young women with early-stage breast cancer in Europe has significantly improved since the 1990s, mostly due to increased usage of adjuvant systemic treatment [1]. This has sparked interest in chemotherapy de-escalation for certain patient groups, for example young breast cancer patients. Prognostic models like PREDICT [2] and CancerMath [3] can aid in the treatment decision-making process by predicting both the mortality after surgery and the benefit of adjuvant systemic treatment. Oncologists can use these predictions in the shared decision-making process to potentially de-escalate chemotherapy safely.

PREDICT (www.predict.nhs.uk), one of the most widely used prognostic models for breast cancer patients, has undergone several updates [2,4–8]. Previous versions of PREDICT have been validated in various populations, including in women from Western Europe [9–11], North America [5], and Asia [12]. However, studies have shown that PREDICT underestimated breast cancer-specific mortality of young women [9,11, 12]. To address this issue, PREDICT was updated to include age at diagnosis (version 2.2, v2.2) [2]. The latest version 2.3 (v2.3, not yet available online) also included progesterone receptor (PR) status as an additional predictor [8].

PREDICT v2.2 and v2.3 underwent extensive external validation, with the results summarized in Table 1 [2,13–20]. These studies included women of different ages and eras. However, there are concerns about the accuracy of PREDICT in predicting mortality after surgery for young women, because the majority of young women included in these validation studies received systemic therapy. Accurate prediction of mortality after surgery is crucial when PREDICT-aided treatment decision-making is used. Therefore, this study aimed to assess the predictive performance and clinical utility of PREDICT v2.2 and v2.3 in a population-based, systemic treatment-naïve cohort of young women with node-negative breast cancer. Our study population was minimized with indication bias as all women were diagnosed in an era when node-negativity was considered to associate with favorable outcomes.

2. Materials and methods

2.1. Study population

We selected women from the population-based PARADIGM cohort for the external validation. Patient selection and data collection of the PARADIGM cohort have been reported previously [23]. In brief, all women diagnosed under age 40 years with node-negative invasive breast cancer, between 1989 and 2000, were selected from the Netherlands Cancer Registry. In this period, women with node-negative disease were considered low-risk. As a result, they only underwent locoregional treatment, and did not receive hormone therapy, trastuzumab, or chemotherapy. Women with previous malignancies or bilateral breast cancer at diagnosis were excluded. For the current study, we excluded women diagnosed under age 25 years (Fig. 1), since PREDICT cannot provide predictions for them. Vital status was obtained through linkage with the municipality population register. Cause of death, however, was unknown. In total we included 2264 women, 27 of whom were lost to follow-up within 10 years after diagnosis.

2.2. Statistical analysis

2.2.1. Predictors of the PREDICT algorithms and missing data

The PREDICT model consists of two separate algorithms which use different predictors for women with ER-positive and ER-negative tumors. The predicted all-cause mortality for women with ER-positive (ER expression $\geq$ 1%) and ER-negative (ER expression $<$ 1%) tumors were therefore computed separately. breast cancer detection mode was assumed to be 'clinically detected' as all included women were diagnosed outside the age range covered by the Dutch breast cancer screening program [24]. Furthermore, the screening program for women with a genetic predisposition were not yet implemented before 2000 [25]. Information on Ki67 status was missing in all women and therefore set to 'unknown', as was allowed by PREDICT. Complete information on ER status, tumor size and grade is required for PREDICT, therefore, multiple imputation by chained equations was performed. The imputation procedure is outlined in the Supplementary Methods.

2.2.2. Predictive performance and clinical utility

Predictive performance was evaluated by calibration, the agreement between observed and predicted mortality, and discrimination, the ability to differentiate between women who did or did not die during follow-up. Calibration was calculated using the ratio of observed to expected all-cause mortality (O/E ratio). A value $>$ 1 indicates that the observed mortality is larger than the predicted mortality, implying the model underestimates mortality, while a value $<$ 1 suggests that the model overestimates mortality. For the calibration plot, we divided women into quintiles based on their predicted 10-year all-cause mortality. The observed 10-year all-cause mortality, calculated using the Kaplan-Meier method, was then plotted against the average predicted mortality for each quintile. Discriminative ability was assessed using the area under the receiver operating characteristic curve (AUC). In this study, we defined a moderate discrimination as an AUC of 0.60 or higher. Details on calculating 95% confidence intervals (CI) are given in the Supplementary Methods.

We evaluated the predictive performance of PREDICT v2.2 and v2.3 in all women, as well as in subgroups based on ER status and immunohistochemical breast cancer subtype defined by hormone receptor (HR) and human epidermal growth factor receptor two (HER2) status. Immunohistochemical breast cancer subtypes include HR-positive/HER2-negative (ER-positive and/or PR-positive and HER2-negative), HR-positive/HER2-positive (ER-positive and/or PR-positive and HER2-positive), HR-negative/HER2-positive (ER-negative, PR-negative and HER2-positive), and triple-negative (ER-negative, PR-negative and HER2-negative). We also determined the predictive performance of PREDICT in women with triple-negative tumors with and without germline *BRCA1* or *BRCA2* (*BRCA1/2*) mutations (see Supplementary Methods).

The potential clinical utility of PREDICT v2.2 and v2.3 in aiding adjuvant chemotherapy decision-making was calculated using decision curve analysis [26]. Clinicians from the TRANSBIG consortium recommended the administration of adjuvant chemotherapy for women whose absolute 10-year mortality was $\geq$ 12% for those with ER-positive tumors, and $\geq$ 8% for those with ER-negative tumors [27]. Therefore, this study classified women as high-risk if their predicted 10-year all-cause mortality was $\geq$ 12% for those with ER-positive tumors, or $\geq$ 8% for those with ER-negative tumors. All other women were considered low-risk. Furthermore, the Dutch guideline recommends adjuvant chemotherapy when the treatment could yield an absolute10-year survival benefit of at least 3–5% [28]. This equates to an absolute mortality risk ranging from approximately 10–15%, given that adjuvant chemotherapy can give a relative reduction in the risk of dying from breast cancer of around 40–60% [28,29]. Thus, sensitivity analysis was conducted using mortality thresholds ranging from 10% to 15%.

Table 1
Summary of the external validation studies on PREDICT v2.2 and v2.3.

Study author	External validation cohort	Young patients proportion and treatment	Assessment of clinical utility	Assessment methods and results[e]
Candido dos Reis (2017) [2][a]	BCOS • Hospital-based cohort • Country: the Netherlands • N = 981 • Year of diagnosis: 1990–2000 • Follow-up: until 2013 • Age at diagnosis: < 50 years • Systemic therapy usage: 67%[b]	• Proportion of women < 40 years: 29%[b] • Systemic therapy usage: 59% [11][b]	No	• Calibration for 10-year mortality of PREDICT v2.2 ■ All women: the number of breast cancer-specific deaths was overestimated by 7% (*p-value = 0.25*). ■ All women: the number of non-breast cancer deaths was underestimated by 9% (*p-value = 0.66*). ■ All women: the number of all-cause deaths was overestimated by 6% (*p-value = 0.34*). • Discrimination for 10-year mortality of PREDICT v2.2 ■ AUC: 0.741 (ER-positive); 0.632 (ER-negative)
	NTBCS • Population-based cohort • Country: the United Kingdom • N = 1944 • Year of diagnosis: 1989–1998 • Follow-up: until October 2012 • Age at diagnosis: 20–79 years • Systemic therapy usage: not reported	• Proportion of women < 40 years: 8% • Systemic therapy usage: not reported	No	• Calibration for 10-year mortality of PREDICT v2.2 ■ All women: the number of breast cancer-specific deaths was overestimated by 2% (*p-value = 0.74*). ■ All women: the number of non-breast cancer deaths was underestimated by 19% (*p-value = 0.039*). ■ All women: the number of all-cause deaths was underestimated by 4% (*p-value = 0.36*). • Discrimination for 10-year mortality of PREDICT v2.2 ■ AUC: 0.790 (ER-positive); 0.680 (ER-negative)
	POSH: • Population-based cohort • Country: the United Kingdom • N = 2609 • Year of diagnosis: 2000–2008 • Follow-up: until December 2014 • Age at diagnosis: ≤ 40 years	• Proportion of women ≤ 40 years: 100% • Systemic therapy usage: 98% [13][c]	No	• Calibration for 10-year mortality of PREDICT v2.2 ■ All women: the number of breast cancer-specific deaths was overestimated by 9% (*p-value = 0.018*). ■ All women: the number of non-breast cancer deaths was overestimated by 57% (*p-value < 0.001*). ■ All women: the number of all-cause deaths was overestimated by 12% (*p-value < 0.001*). • Discrimination for 10-year mortality of PREDICT v2.2 ■ AUC: 0.746 (ER-positive); 0.715 (ER-negative)
Karapanagiotis (2018) [15]	BCOS, NTBCS and POSH	Not applicable	Yes	Using a risk threshold between 14% and 23%, the net benefit of PREDICT v2.2 was superior compared to CancerMath.
van Maaren (2017) [14]	• Population-based cohort • Country: the Netherlands • N = 8834 • Year of diagnosis: 2005 • Follow-up: median follow-up time was 10.4 years • Systemic therapy usage: 38.2%	• Proportion of women < 40 years: 5.6% • Systemic therapy usage: 88% [21][d]	No	• Calibration for 10-year all-cause mortality of PREDICT v2.2 ■ All women: the number of deaths was overestimated by 3.65% (*p-value = 0.072*). ■ ER-positive: the number of deaths was overestimated by 0.32% (*p-value = 0.892*). ■ ER-negative: the number of deaths was overestimated by 13.37% (*p-value < 0.001*). ■ Women < 40 years: the number of deaths was overestimated by 10.52% (*p-value = 0.289*). • Discrimination for 10-year all-cause mortality of PREDICT v2.2 ■ AUC: 0.78 (ER-positive); 0.76 (ER-negative)
Gray (2018) 22	• Population-based cohort • Country: the United Kingdom • N = 45,789 • Year of diagnosis: 2001–2015 • Follow-up: until February 2017 • Age at diagnosis: all age groups • Systemic therapy: 89% of the women were treated	• Proportion of women < 49 years: 18.8% • Systemic therapy usage: not reported	No	• Calibration for 10-year all-cause mortality of PREDICT v2.2 ■ All women: the number of deaths was underestimated by 1.96% (*p-value = 0.151*). ■ ER-positive: the number of deaths was underestimated by 4.29% (*p-value = 0.005*). ■ ER-negative: the number of deaths was overestimated by 5.62% (*p-value = 0.053*). ■ Women < 35 years: the number of deaths was overestimated by 11.77% (*p-value = 0.368*). ■ Women aged 35–49 years: the number of deaths was overestimated by 14.44% (*p-value < 0.001*). • Discrimination for 10-year all-cause mortality of PREDICT v2.2 ■ AUC: 0.767 (ER-positive); 0.761 (ER-negative)
Aguirre (2019) [17]	• Population-based cohort • Country: Spain • N = 535 • Year of diagnosis: 2000–2008 • Follow-up: until December 2013 • Age at diagnosis: median was 59 years (ranged between 50 and 71 years) • Systemic therapy usage: 96%	• Proportion of women < 40 years: 0%	No	• Calibration for 5-year all-cause mortality of PREDICT v2.2 ■ ER-positive: the observed mortality was underestimated by 21.21% (*p-value = 0.425*). ■ ER-negative: the observed mortality was overestimated by 31.93% (*p-value = 0.180*). • Discrimination for 5-year all-cause mortality of PREDICT v2.2 ■ C-index: 0.768 (ER-positive); 0.697 (ER-negative)
Zaguirre (2020) [18]	• Hospital-based cohort • Country: Japan • N = 636 • Year of diagnosis: 2001–2013	• Proportion of women < 40 years: 5.7%	No	• Calibration for 10-year all-cause mortality of PREDICT v2.2 ■ All women: the mortality was overestimated by 17.14% (*p-value = 0.106*).

(*continued on next page*)

Table 1 (*continued*)

Study author	External validation cohort	Young patients proportion and treatment	Assessment of clinical utility	Assessment methods and results[e]
	• Follow-up: until October 2019 • Age at diagnosis: median was 57 years • Systemic therapy usage: 89.3%	• Systemic therapy usage: not reported		■ ER-positive: the mortality was overestimated by 18.75% (*p-value = 0.150)*. ■ ER-negative: the mortality was overestimated by 19.58% (*p-value = 0.270)*. ■ Women < 40 years: the mortality was underestimated by 13.33% (*p-value = 0.845)*. • Discrimination for 10-year all-cause mortality of PREDICT v2.2: AUC = 0.707
Grootes (2022) [8]	• Population-based cohort • Country: New Zealand • N = 11,365 • Year of diagnosis: 2000–2014 • Follow-up: until December 2014 • Mean age at diagnosis: 57.1 years • Chemotherapy usage: 35% • Hormone therapy usage: 62%	• Proportion of women < 40 years: not reported	Yes	• Calibration for 15-year breast cancer-specific mortality of PREDICT v2.2 and v2.3 ■ All women: the number of deaths was overestimated by 17.62% (*p-value < 0.001;* v2.2*)* and 18.33% *(p-value <0.001;* v2.3*)*. ■ ER-positive: the number of deaths was overestimated by 11.18% (*p-value = 0.005;* v2.2*)* and 12.04% *(p-value = 0.003;* v2.3*)*. ■ ER-negative: the number of deaths was overestimated by 25.05% (*p-value < 0.001;* v2.2*)* and 25.76% *(p-value < 0.001;* v2.3*)*. • Discrimination for up to 15-year breast cancer-specific mortality of PREDICT v2.2 and v2.3 ■ v2.2 C-index: 0.898 (ER-positive); 0.807 (ER-negative) ■ v2.3 C-index: 0.902 (ER-positive); 0.809 (ER-negative) • Clinical utility when using PREDICT v2.2 and PREDICT v2.3 ■ 2.4% women changed from a lower risk category to a higher risk category when using PREDICT v2.3.
Agostinetto (2022) [19]	• Data from the clinical trial ALTTO • Country: 44 countries across the world • N = 2794 • Year of diagnosis: 2007–2011 • Median follow-up: 6 years • Mean age at diagnosis: 57.1 years • Chemotherapy usage: 100% • Hormone therapy usage: 53% • Anti-HER2 therapy: 100%	• Proportion of women ≤ 40: 17.7% • Systemic therapy usage: 100%	No	• Calibration for 5-year all-cause mortality of PREDICT v2.2 ■ All women: the mortality was overestimated by 55.75% (*p-value < 0.001)*. ■ HR-positive: the mortality was overestimated by 39.42% (*p-value < 0.001)*. ■ HR-negative: the mortality was overestimated by 65.40% (*p-value < 0.001)*. ■ Women ≤ 40 years: the mortality was overestimated by 54.58% (*p-value < 0.001)*. • Discrimination for 5-year all-cause mortality of PREDICT v2.2 ■ All women: AUC = 0.7375 ■ HR-positive: AUC = 0.7681 ■ HR-negative: AUC = 0.7187 ■ Women ≤ 40 years: AUC = 0.7609
Muranen (2023) [20]	CIMBA • Studies include women from genetic clinics • Country: multiple countries from Europe, Oceania, and North America • Germline *BRCA1* mutation carriers: N = 2892 • Germline *BRCA2* mutation carriers: N = 1813 • Year of diagnosis: since 1990 • Systemic therapy usage: not reported systemically, but 90% of *BRCA1* carriers with ER-negative tumors received chemotherapy	• Proportion of women < 40 years: not reported	No	• Calibration for 10-year breast cancer-specific mortality of PREDICT v2.3[f] ■ ER-positive *BRCA1*: the mortality was overestimated by 23.53% *(p-value = 0.010)*. ■ ER-positive *BRCA2*: the mortality was underestimated by 5.56% *(p-value = 0.385)*. ■ ER-negative *BRCA1*: the mortality was overestimated by 30.43% *(p-value < 0.001)*. ■ ER-negative *BRCA2*: the mortality was overestimated by 40.00% *(p-value < 0.001)*. • Discrimination for 10-year breast cancer-specific mortality ■ Gönen & Heller unbiased concordance: 0.565 (ER-positive *BRCA1*); 0.604 (ER-positive *BRCA2*); 0.651 (ER-negative *BRCA1*); 0.554 (ER-positive *BRCA2*)
	BCAC • Most studies are population-based, with a few family-based studies • Country: multiple countries from Europe, Oceania, and North America • Germline *BRCA1* mutation carriers: N = 316 • Germline *BRCA2* mutation carriers: N = 432 • Systemic therapy usage: not reported systemically	• Proportion of women < 40 years: not reported	No	• Calibration for 10-year breast cancer-specific mortality of PREDICT v2.3[f] ■ ER-positive *BRCA2*: the mortality was underestimated by 10.00% *(p-value = 0.423)*. ■ ER-negative *BRCA1*: the mortality was overestimated by 25.00% *(p-value = 0.054)*. • Discrimination for 10-year breast cancer-specific mortality ■ Gönen & Heller unbiased concordance: 0.653 (ER-positive *BRCA2*); 0.651 (ER-negative *BRCA1*)

Abbreviations: ER, estrogen receptor; CI, confidence interval; AUC, area under the receiver operating characteristic curve; HR, hormone receptor.

[a] This study contains the results of refitting PREDICT (to generate PREDICT v2.2), and external validating PREDICT in three independent cohorts.

[b] Data was extracted from a previous study using the BCOS cohort.

[c] Data was extracted from a previous study using the POSH cohort.

[d] Data was extracted from a previous study which reported the percentage of systemic therapy usage in the Netherlands.

[e] Only the results of all women and/or women with different estrogen-receptor status, and women diagnosed at a young age (if available) were shown. Calibration was reflected by the relative difference between the predicted and observed mortality rate or number of deaths; if the original studies reported relative differences, we used the values reported from the original studies (with the corresponding *p*-values if reported), otherwise we approximated the relative differences by dividing the absolute differences between the predicted and observed mortality rates or number of deaths by the predicted values. The corresponding *p*-values of the relative differences were calculated using test statistic defined as *(observed number of deaths – predicted number of deaths)*^{2} */ predicted number of deaths*, and assuming that it followed a Chi-square distribution. For the studies that did not report the numbers of predicted and observed deaths, we multiplied the total number of patients used for calibration with the predicted and observed mortality rates or (1 – survival rates) to approximate the values.

[f] Since the study did not provide the exact number of patients used for calibration, we approximated it using the number of carriers, percentages of ER-positive and ER-negative tumors, and patients without metastasis at diagnosis. We then used this approximated total number of patients, predicted mortality, and observed mortality to calculate the predicted and observed number of deaths.

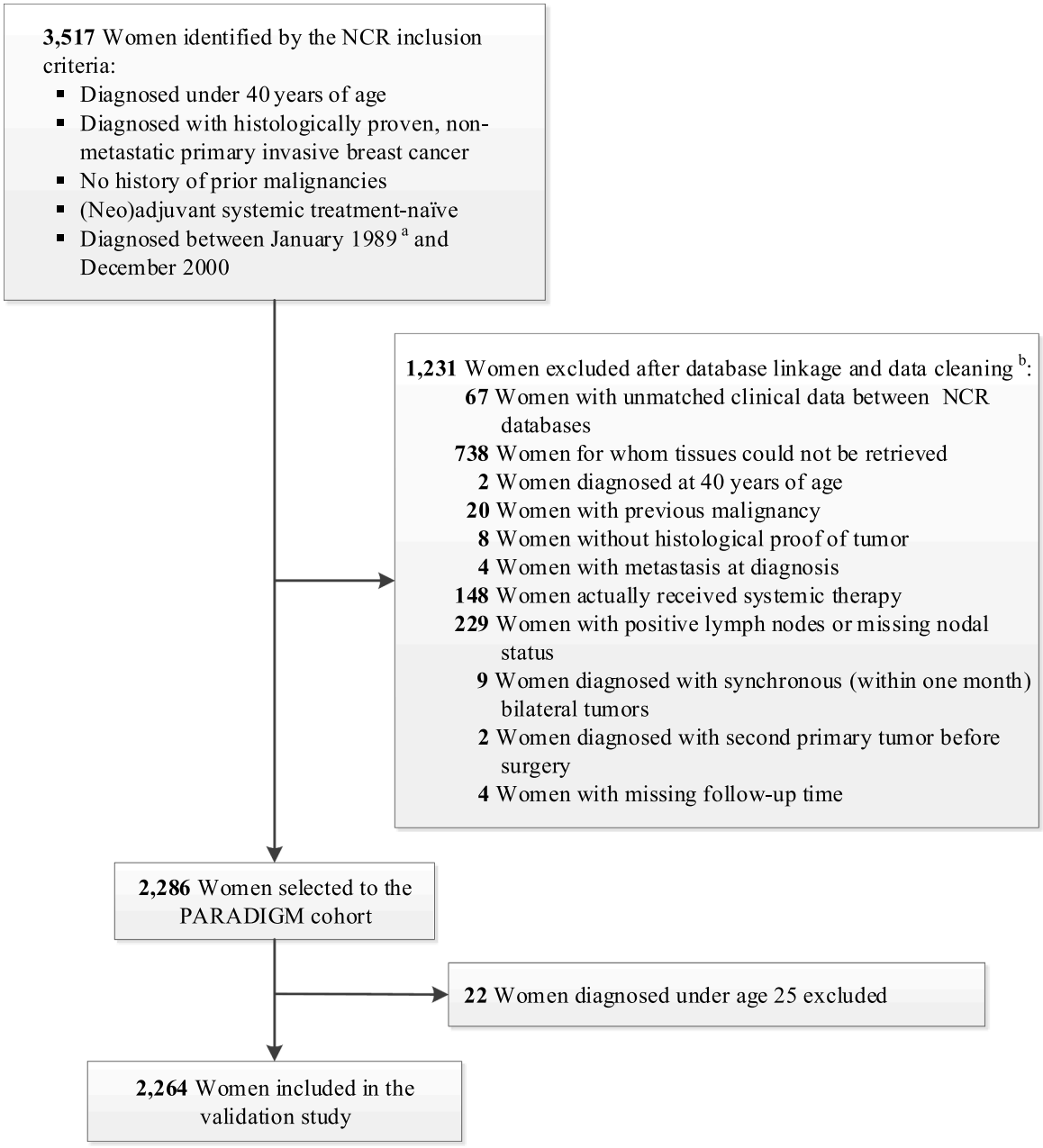

Fig. 1. Flow chart of the women included in the validation study. Abbreviation: NCR, Netherlands Cancer Registry. [a] The Netherlands Cancer Registry provides nationwide registry since 1989. [b] The exclusion steps are in subsequent order.

Furthermore, to account for different perspectives regarding the relative risk reduction from the treatment, we lowered the risk thresholds to 5% in the sensitivity analysis. Clinical utility was represented by net benefit, calculated as the correctly predicted high-risk women (true-positive) minus the weighted falsely predicted high-risk women (false-positive). The weight was assigned under the assumption that the harm of unnecessary chemotherapy in low-risk women weighs less than the benefit of chemotherapy in high-risk women [26]. The true-negative, false--negative, sensitivity, and specificity were also computed. The net benefit of PREDICT was compared to a treatment strategy that considered all

women as high-risk, thus recommending chemotherapy to all. Detailed information on the calculation of clinical utility is in the Supplementary Methods.

All analyses were performed using R version 4.1.1 [30].

3. Results

3.1. Baseline characteristics

For the 2264 women included in this study, median follow-up was 20.6 years and median age at diagnosis was 36 years. Most women had HR-positive/HER2-negative tumors (60.4%), followed by triple-negative (22.9%), HR-positive/HER2-positive (12.2%), and HR-negative/HER2-positive (4.6%) tumors, respectively. Patient characteristics stratified by ER status are shown in Table 2.

3.2. Predictive performance

PREDICT v2.2 significantly underestimated 10-year all-cause mortality by 33% in all women (1.33, 95%CI:1.22–1.43). The underestimation was 45% in women with ER-positive tumors and 13% in women with ER-negative tumors. Calibration in each immunohistochemical breast cancer subtype is shown in Table 3. For women with triple-negative tumors and germline *BRCA1/2* mutations (26%, 98/380),

Table 2
Baseline characteristics of the validation cohort.

		All women (N = 2264)	ER status	
			Negative (N = 559)	Positive (N = 1379)
Median age, years (Q1-Q3)		36.0 (33.0, 38.0)	35.0 (32.5, 38.0)	36.0 (34.0, 38.0)
ER status, No. (%)	Negative	559 (28.8%)	559 (100%)	0 (0%)
	Positive	1379 (71.2%)	0 (0%)	1379 (100%)
	Missing	326	0	0
PR status, No. (%)	Negative	710 (36.7%)	531 (95.5%)	178 (12.9%)
	Positive	1225 (63.3%)	25 (4.5%)	1199 (87.1%)
	Missing	329	3	2
HER2 status, No. (%)	Negative	1614 (83.3%)	465 (83.6%)	1145 (83.2%)
	Positive	324 (16.7%)	91 (16.4%)	231 (16.8%)
	Missing	326	3	3
Median tumor size, mm (Q1-Q3)		16.0 (12.0, 22.0)	20.0 (15.0, 25.0)	15.0 (12.0, 20.0)
Missing tumor size, No.		324	67	165
Tumor grade, No. (%)	Grade 1	367 (17.6%)	8 (1.4%)	337 (24.5%)
	Grade 2	800 (38.4%)	97 (17.4%)	653 (47.4%)
	Grade 3	917 (44.0%)	454 (81.2%)	387 (28.1%)
	Missing	180	0	2
Immuno-histochemical breast cancer subtype	HR-positive/ HER2-negative	1167 (60.4%)	22 (4.0%)[a]	1145 (83.2%)
	HR-positive/HER2-positive	235 (12.2%)	3 (0.5%)[a]	231 (16.8%)
	HR-negative/HER2-positive	88 (4.6%)	88 (15.9%)	0 (0%)
	Triple-negative	442 (22.9%)	442 (79.6%)	0 (0%)
	Germline *BRCA1/2*-mutated[b]	98 (25.8%)	Not applicable	
	Germline *BRCA1/2*-non mutated[b]	282 (74.2%)		
	Missing mutation status[b]	62		
	Missing	332	4	3
Surgery, No. (%)	Lumpectomy	1410 (62.3%)	352 (63.0%)	881 (63.9%)
	Mastectomy	811 (35.8%)	196 (35.1%)	474 (34.4%)
	Surgery not specified	43 (1.9%)	11 (2.0%)	24 (1.7%)
Radiotherapy, No. (%)	No	754 (33.3%)	179 (32.0%)	442 (32.1%)
	Yes	1510 (66.7%)	380 (68.0%)	937 (67.9%)

Abbreviations: Q1, Quartile 1; Q3, Quartile 3; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor two; HR, hormone receptor.

[a] Women with estrogen receptor-negative tumors who were classified as hormone receptor-positive was due to progesterone receptor-positive tumors.

[b] Germline *BRCA1/2* mutation status was only available for 380 women with triple-negative breast cancer.

Table 3
10-year O/E ratio and AUC of PREDICT v2.2 and v2.3 based on multiple-imputed data.

	No. women[a]	No. events[a]	PREDICT v2.2		PREDICT v2.3	
			O/E ratio (95% CI)	AUC (95% CI)	O/E ratio (95% CI)	AUC (95% CI)
All women	2264	492	1.33 (1.22–1.43)	0.65 (0.62–0.68)	1.32 (1.22–1.43)	0.65 (0.62–0.68)
ER status						
ER-positive	1614	330	1.45 (1.31–1.59)	0.69 (0.65–0.72)	1.45 (1.31–1.60)	0.69 (0.66–0.72)
ER-negative	650	162	1.13 (0.97–1.29)	0.56 (0.51–0.62)	1.11 (0.96–1.27)	0.56 (0.51–0.62)
Immunohistochemical breast cancer subtype						
HR-positive/HER2-negative	1376	244	1.37 (1.21–1.53)	0.68 (0.64–0.72)	1.39 (1.22–1.55)	0.69 (0.65–0.72)
HR-positive /HER2-positive	274	94	1.65 (1.37–1.94)	0.57 (0.50–0.64)	1.64 (1.36–1.92)	0.57 (0.50–0.64)
HR-negative/HER2-positive	106	29	1.08 (0.73–1.43)	0.55 (0.42–0.68)	1.05 (0.71–1.39)	0.55 (0.42–0.68)
Triple-negative	506	124	1.15 (0.96–1.33)	0.55 (0.48–0.61)	1.12 (0.94–1.30)	0.55 (0.48–0.61)
those with a germline *BRCA1/2* mutation[b]	98	31	1.46 (1.03–1.88)	0.51 (0.38–0.64)	1.42 (1.01–1.84)	0.51 (0.38–0.64)
those without a germline *BRCA1/2*-mutation[b]	282	66	1.09 (0.86–1.32)	0.56 (0.48–0.64)	1.06 (0.84–1.29)	0.56 (0.48–0.64)

Abbreviations: ER, estrogen receptor; HR, hormone receptor; HER2, human epidermal growth factor receptor two; O/E ratio, the ratio of observed and expected all-cause mortality; CI, confidence interval; AUC, area under the receiver operating characteristic curve.

[a] The number of women in each subgroup was the median number over 50 imputed datasets. The total numbers of ER status and immunohistochemical breast cancer subtypes might be slightly higher or lower than 2264 due to rounding errors.

[b] Germline *BRCA1/2* mutation status was only available for 380 women with triple-negative breast cancer.

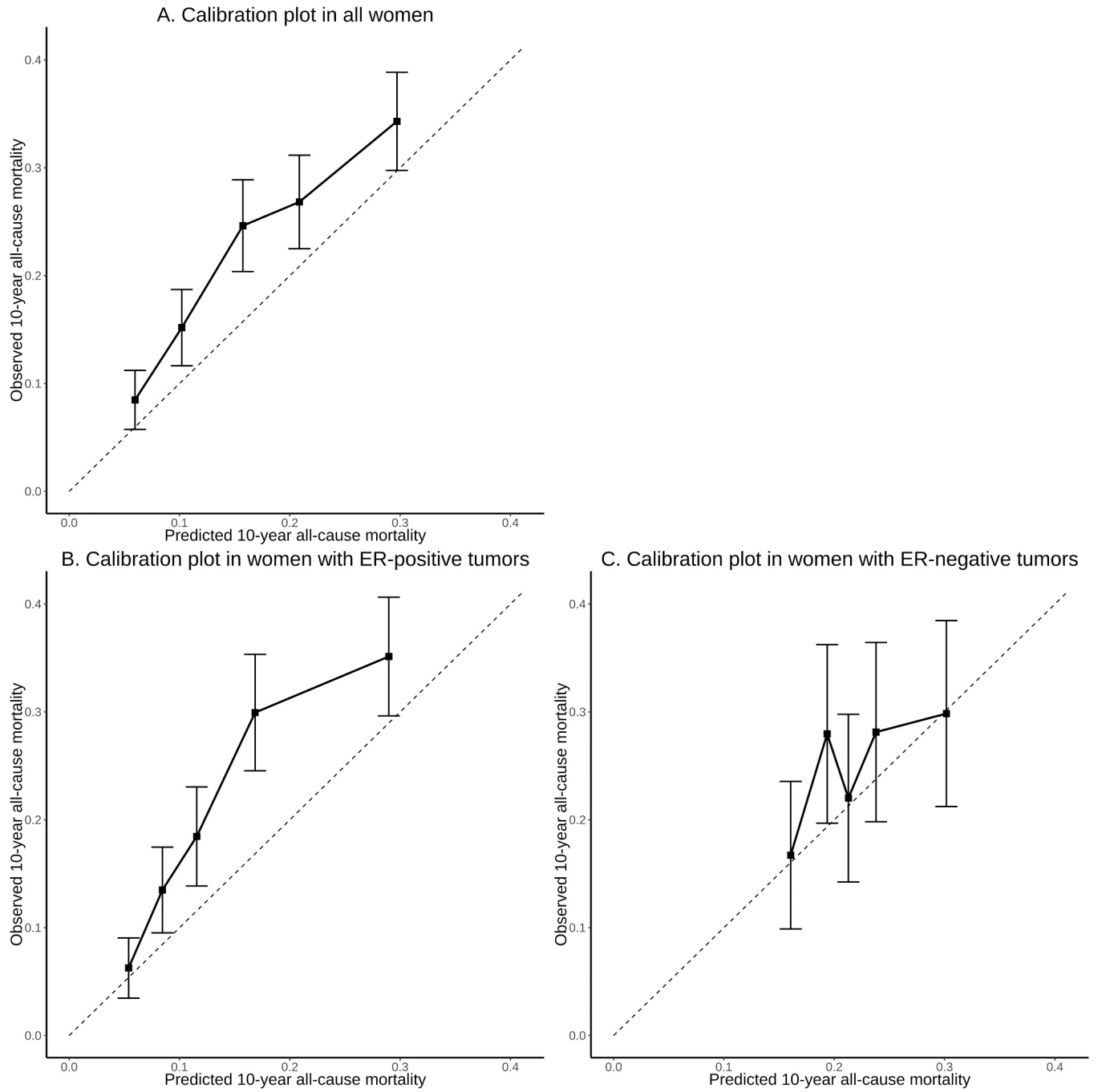

Fig. 2. Calibration plots of PREDICT v2.2 for 10-year all-cause mortality in all women (A), women with ER-positive tumors (B) and women with ER-negative tumors (C). Abbreviations: ER = estrogen receptor.

PREDICT v2.2 underestimated 10-year all-cause mortality by 46%, whereas in women without the mutation only 9%, which is deemed acceptable (Table 3). The calibration plot showed that the underestimation of mortality was less pronounced for women with ER-positive tumors who had a low predicted mortality, but became more apparent as the predicted mortality increased (Fig. 2). Overall, the model had moderate discrimination in this patient population ($AUC_{10\text{-year}}$:0.65, 95%CI:0.62–0.68). For women with ER-positive tumors, PREDICT v2.2 had a 69% chance to correctly differentiate between those who died within 10 years and those who did not ($AUC_{10\text{-year}}$:0.69, 95% CI:0.65–0.72), compared to a 56% chance in women with ER-negative tumors ($AUC_{10\text{-year}}$:0.56, 95%CI: 0.51–0.62). The predictive performance of PREDICT v2.3 was similar to that of PREDICT v2.2 (Table 3). The predictive performance of PREDICT v2.2 and v2.3 at 5 and 15 years is shown in Supplementary Tables 1–2. Results based on imputed data aligned with results using only cases with complete information on ER, PR, HER2 status, tumor size, and grade (Supplementary Tables 3–4).

3.3. Clinical utility

The sensitivity, specificity, and net benefit of PREDICT v2.2 and v2.3, as well as the chemotherapy-to-all strategy are shown in Table 4. For women with ER-positive tumors, both versions of PREDICT showed a sensitivity of about 70% and a specificity of about 60%. For women with ER-negative tumors, the sensitivity of PREDICT v2.2 and v2.3 was 100%, meaning that the model identified all high-risk women. However,

Table 4
Classification table of PREDICT v2.2 and v2.3, and the chemotherapy-to-all strategy.

		Per 1000 women with ER-positive tumors, at 12% threshold			Per 1000 women with ER-negative tumors, at 8% threshold		
		Observed high-risk	*Observed low-risk*	*Net benefit*[a]	*Observed high-risk*	*Observed low-risk*	*Net benefit*[a]
PREDICT v2.2	***Predicted high-risk***	145 (TP)	324 (FP)	101	249 (TP)	744 (FP)	184
	Predicted low-risk	62 (FN)	470 (TN)	-	0 (FN)	6 (TN)	-
	Sensitivity	70.0%		-	100.0%		-
	Specificity	59.2%		-	0.8%		-
PREDICT v2.3	***Predicted high-risk***	144 (TP)	317 (FP)	101	249 (TP)	744 (FP)	184
	Predicted low-risk	62 (FN)	477 (TN)	-	0 (FN)	7 (TN)	-
	Sensitivity	69.9%		-	100.0%		-
	Specificity	60.1%		-	0.9%		-
Chemotherapy-to-all	***Predicted high-risk***	207 (TP)	793 (FP)	98	249 (TP)	751 (FP)	184
	Predicted low-risk	0 (FN)	0 (TN)	-	0 (FN)	0 (TN)	-
	Sensitivity	100.0%		-	100.0%		-
	Specificity	0%		-	0%		-

Abbreviations: ER, estrogen receptor; TP, true-positive; FN, false-negative; FP, false-positive; TN, true-negative.
[a] Net benefit was calculated using true-positive and false-positive only.

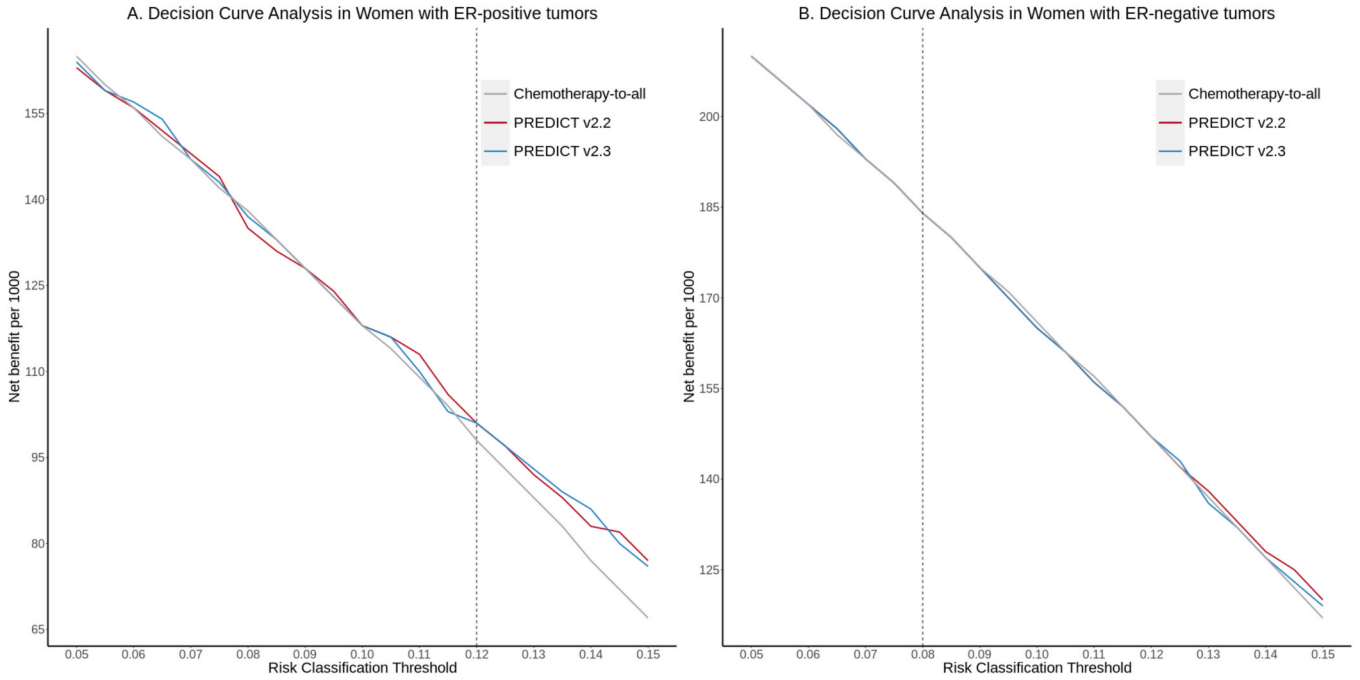

Fig. 3. Net benefit of PREDICT v2.2, PREDICT v2.3 and the chemotherapy-to-all strategy in women with ER-positive tumors (A), and in women with ER-negative tumors (B). Note that the y-axes do not start from zero, and the scale of y-axes are different in panel A and in panel B. Black dashed vertical lines are located at the predefined mortality thresholds for chemotherapy treatment decisions, i.e. 0.12 for women with ER-positive tumors, and 0.08 for women with ER-negative tumors. Abbreviations: ER = estrogen receptor.

with a specificity < 1%, the true-negative rate was extremely low (Table 4). For every 1000 women with ER-positive tumors, PREDICT v2.2 correctly identified 145 women as high-risk and falsely identified 324 women as high-risk when using the 12% mortality threshold. After weighing the benefit of correctly treating 145 high-risk women with chemotherapy to the harm of treating 324 low-risk women with unnecessary chemotherapy, the net benefit of using PREDICT v2.2 was 101 true high-risk women per 1000 women (Table 4). This net benefit was slightly higher than that of the chemotherapy-to-all strategy (98 true high-risk women per 1000 women). For women with ER-negative tumors, PREDICT v2.2 showed the same net benefit as the chemotherapy-to-all strategy at the 8% mortality threshold, as both had a net benefit of 184 true high-risk women per 1000 women. Results of the sensitivity analysis using mortality thresholds from 5% to 15% are shown in Fig. 3 and Supplementary Table 5. For women with ER-positive tumors, PREDICT v2.2 consistently showed an equal or slightly higher net benefit than the chemotherapy-to-all strategy when the mortality threshold was between 9% and 15%. For women with ER-negative tumors, PREDICT v2.2 showed equal or slightly lower net benefit than the treatment-to-all strategy when the mortality threshold was between 5% and 12%. Results of PREDICTv2.3 were similar to those of PREDICT v2.2 (Table 4, Fig. 3 and Supplementary Table 5).

4. Discussion

Our study is the first investigation of the performance of PREDICT v2.2 and v2.3 in young women with node-negative breast cancer who did not receive (neo)adjuvant systemic treatment. Our results show that PREDICT v2.2 and v2.3 significantly underestimate all-cause mortality in this population, particularly in women with ER-positive tumors. The model discrimination is moderate for the total cohort, but poor in women with ER-negative tumors. We showed a slightly higher net benefit of using PREDICT v2.2 and v2.3 to aid chemotherapy decision-making compared to the chemotherapy-to-all strategy for women with

ER-positive tumors, but not for women with ER-negative tumors.

According to the Dutch guideline, most young women with node-negative breast cancers are treated with chemotherapy, especially those with ER-negative tumors (Supplementary Table 6) [28]. Therefore, the chemotherapy-to-all strategy largely reflects current clinical practice. Compared to administering chemotherapy to all women, using PREDICT-aided decisions in women with ER-positive tumors resulted in giving less unnecessary chemotherapy, as was shown by fewer false-positive predictions. However, it also missed many high-risk women (false-negative). Treatment guidelines already recommend avoiding unnecessary chemotherapy in low-risk women based on age, tumor size, tumor grade, ER status, PR status, HER2 status and genomic signatures [28,31,32]. Therefore, the benefit of using PREDICT over guideline-aided chemotherapy decision-making may be limited for young women. The net benefit results show that for women with ER-negative tumors, PREDICT cannot help to de-escalate chemotherapy.

The miscalibration of PREDICT in our study population is unlikely to be explained by the difference in time at diagnosis between the PARADIGM cohort (1989–2000, the Netherlands) and the PREDICT derivation cohort (1999–2003, the UK [4]), as all tumors in the PARADIGM cohort were revised according to current diagnostic standard [23]. A more likely explanation for the miscalibration is inaccurate estimation of treatment benefits in young women. PREDICT extracted the supposed treatment effect from external studies [33,34] and assumed the same relative risk reduction from those treatments across women of all age groups [2]. If the actual treatment effect in young women is higher than predicted, the survival before adjuvant systemic treatment would be overestimated because part of the treatment effect would be attributed to the predicted survival. Indeed a previously published meta-analysis found that young women received significantly higher reduction in breast cancer-specific mortality from anthracycline-based chemotherapy than their older counterparts [29], though a subsequent meta-analysis showed a largely age-independent effect [34]. This discrepancy might originate from the usage of different anthracycline-based regimens in these studies [29,34], indicating the need for a regimen-specific benefit prediction. In addition, both chemotherapy and endocrine treatment reduce the incidence of contralateral breast cancer [35–37], especially in young women [33]. This reduced incidence may, to some extent, translate to a reduced all-cause mortality. Another possible explanation is inaccurate estimation of the age effect, as the number of young women in the PREDICT derivation cohort is very small. This is also reflected by the warning on the website of PREDICT that the model may be less accurate for women diagnosed with ER-positive tumors under 30 years old. Therefore, PREDICT should recalibrate the age effect in young women with a larger sample size.

The poor discriminative ability in young women with node-negative, ER-negative tumors is mainly due to the homogeneous distribution of age, tumor grade, and nodal status (for all women in our cohort) in these women. To improve model discrimination in young women with ER-negative breast cancer, better predictors are required. Recent studies have highlighted the prognostic value of stromal tumor infiltrating lymphocytes (sTILs) in women of all ages with triple-negative and HER2-enriched breast cancers [38–41]. Future updates of PREDICT might therefore benefit from incorporating, amongst others, sTILs as a predictor.

Several factors could influence the generalizability of our results to young women diagnosed in the present era. First of all, more advanced pathological methods might have reclassified a few women in our study population from node-negative to node-positive [42,43]. However, we expect the impact of this potential upstaging to be small, as the prognostic value of occult metastases is only modest at best in systemic treatment-naïve patients [42,44–46]. Second, young women who are diagnosed with breast cancers nowadays are referred to genetic counseling [47], and those who carry germline *BRCA1/2* mutations usually receive annual screening or prophylactic surgeries [48]. These measures may lead to a different case-mix in the germline *BRCA1/2* mutation carriers and our results may therefore not apply to contemporary carriers. For instance, Muranen et al. showed that PREDICT overestimated breast cancer-specific mortality in young, chemotherapy-treated ER-negative breast cancer patients with germline *BRCA1* mutations, indicating a favorable chemotherapy response among these women [20]. In order to ensure the validity of PREDICT in the current mutation carriers, the model should be recalibrated with more current cohorts.

This study has several strengths and limitations. First of all, the key strength is that our results are probably not biased by treatment indication as no systemic treatment was recommended to these women at the time of their diagnosis. Second, we assessed the potential clinical utility of PREDICT regarding chemotherapy decision-making using a comprehensive method. The calculation of clinical utility focused exclusively on chemotherapy and did not include hormone therapy or trastuzumab because in current practice, women with ER-positive tumors typically receive hormone therapy, and those with HER2-positive tumors receive trastuzumab. Third, cause of death was unknown in women from our cohort, although most deaths were expected to be breast cancer-related given the young age at diagnosis. Lastly, Ki67 status was unknown, which is relevant for the risk prediction of women with ER-positive tumors. Despite these limitations, our study provides unique and unbiased results to show the validity and utility of PREDICT in young, systemic treatment-naïve, node-negative breast cancer patients.

5. Conclusions

PREDICT yields unreliable predictions and shows limited clinical utility in young, node-negative breast cancer patients. Clinicians should await future updates to PREDICT which incorporate more recent cohorts and novel powerful predictors to improve model performance before using the model to aid chemotherapy decision-making for this patient population.

Funding

The Netherlands Cancer Institute was supported by an institutional grant of the Dutch Cancer Society and of the Dutch Ministry of Health, Welfare and Sport. This study was supported by grants from the Dutch Cancer Society (KWF, grant number 11655/2018-1 to Marjanka K. Schmidt), the Netherlands Organization for Health Research and Development (project number 836021019, to Sabine C. Linn), A Sister's Hope (to Sabine C. Linn), and De Vrienden van UMC Utrecht (to Sabine C. Linn). The funders of this study had no role in the design and conduct of the study; collection, analysis, and interpretation of the data; writing of the manuscript; and the decision to submit the manuscript for publication.

Ethics statement

The study was approved by the Institutional Review Board of the Netherlands Cancer Institute (IRB code: CFMPB554). All retrospective medical data/biospecimen studies in the Netherlands are executed pursuant to Dutch legislation, international standards and a self-regulatory Code of Conduct (https://www.coreon.org/gedragscode-gezondheidsonderzoek/). Prior to 25 May 2018, national legislation on data protection applied, as well as the International Guideline on Good Clinical Practice. From 25 May 2018, hospitals in the Netherlands also have to adhere to the General Data Protection Regulation. Within this framework, patients are informed and have the opportunity to object or actively consent to the (continued) use of their personal data and biospecimens in research. Hence, the procedures comply both with (inter-) national legislative and ethical standards.

CRediT authorship contribution statement

Yuwei Wang: Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Annegien Broeks:** Methodology, Resources, Writing – review & editing. **Daniele Giardiello:** Methodology, Writing – review & editing. **Michael Hauptmann:** Methodology, Writing – review & editing. **Katarzyna Jóźwiak:** Methodology, Writing – review & editing. **Esther A. Koop:** Methodology, Resources, Writing – review & editing. **Mark Opdam:** Methodology, Resources, Writing – review & editing. **Sabine Siesling:** Methodology, Resources, Writing – review & editing. **Gabe S. Sonke:** Methodology, Resources, Writing – review & editing. **Nikolas Stathonikos:** Methodology, Resources, Writing – review & editing. **Natalie D. ter Hoeve:** Methodology, Resources, Writing – review & editing. **Elsken van der Wall:** Methodology, Resources, Writing – review & editing. **Carolien H. M. van Deurzen:** Methodology, Resources, Writing – review & editing. **Paul J. van Diest:** Methodology, Resources, Writing – review & editing. **Adri C. Voogd:** Methodology, Resources, Writing – review & editing. **Willem Vreuls:** Methodology, Resources, Writing – review & editing. **Sabine C. Linn:** Conceptualization, Funding acquisition, Methodology, Resources, Writing – review & editing. **Gwen M.H.E. Dackus:** Funding acquisition, Data curation, Methodology, Resources, Writing – review & editing. **Marjanka K. Schmidt:** Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing – review & editing.

Declaration of Competing Interest

Sabine C. Linn has been an advisory board member for AstraZeneca, Cergentis, IBM, Novartis, Pfizer, Roche and Sanofi, and has received unrestricted institutional research support or unrestricted educational funding from Agendia, Amgen, AstraZeneca, Bayer, Daiichi Sankyo, Eurocept Pharmaceuticals, Genentech, Immunomedics (now Gilead), Merck, Roche, Sanofi and TESARO (now GSK), and has a pending patent application for a *BRCA*-like ovarian cancer classifier. Paul J. van Diest has a pending patent application for DDX3 as a biomarker for cancer and its related methods. Gabe Sonke has received institutional research support from Agendia, AstraZeneca, Merck, Novartis, Roche and Seagen. Other authors claim no conflict of interest.

Data availability

The clinical data in this study are available from the Netherlands Cancer Registry, hosted by the Netherlands Comprehensive Cancer Organization; however, restrictions apply to the availability of these data, which were used under license for the current study. Other data generated (germline *BRCA1/2* mutations) are available from the authors upon reasonable request and with permission from the Netherlands Comprehensive Cancer Organization. This R script of this paper can be found on https://github.com/SchmidtGroupNKI.

Acknowledgments

We thank Prof. Paul Pharoah for providing the R script of PREDICT v2.2 and v2.3, and Daniel J. van der Meer for summarizing results using data from the Netherlands Cancer Registry.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejca.2023.113401.

References

[1] Autier P, Boniol M, La Vecchia C, Vatten L, Gavin A, Hery C, et al. Disparities in breast cancer mortality trends between 30 European countries: retrospective trend analysis of WHO mortality database. BMJ 2010;341:c3620. https://doi.org/10.1136/bmj.c3620.

[2] Candido Dos Reis FJ, Wishart GC, Dicks EM, Greenberg D, Rashbass J, Schmidt MK, et al. An updated PREDICT breast cancer prognostication and treatment benefit prediction model with independent validation. Breast Cancer Res 2017;19:58. https://doi.org/10.1186/s13058-017-0852-3.

[3] Chen LL, Nolan ME, Silverstein MJ, Mihm Jr MC, Sober AJ, Tanabe KK, et al. The impact of primary tumor size, lymph node status, and other prognostic factors on the risk of cancer death. Cancer 2009;115:5071–83. https://doi.org/10.1002/cncr.24565.

[4] Wishart GC, Azzato EM, Greenberg DC, Rashbass J, Kearins O, Lawrence G, et al. PREDICT: a new UK prognostic model that predicts survival following surgery for invasive breast cancer. Breast Cancer Res 2010;12:R1. https://doi.org/10.1186/bcr2464.

[5] Wishart GC, Bajdik CD, Azzato EM, Dicks E, Greenberg DC, Rashbass J, et al. A population-based validation of the prognostic model PREDICT for early breast cancer. Eur J Surg Oncol 2011;37:411–7. https://doi.org/10.1016/j.ejso.2011.02.001.

[6] Wishart GC, Bajdik CD, Dicks E, Provenzano E, Schmidt MK, Sherman M, et al. PREDICT Plus: development and validation of a prognostic model for early breast cancer that includes HER2. Br J Cancer 2012;107:800–7. https://doi.org/10.1038/bjc.2012.338.

[7] Wishart GC, Rakha E, Green A, Ellis I, Ali HR, Provenzano E, et al. Inclusion of KI67 significantly improves performance of the PREDICT prognostication and prediction model for early breast cancer. BMC Cancer 2014;14:908. https://doi.org/10.1186/1471-2407-14-908.

[8] Grootes I, Keeman R, Blows FM, Milne RL, Giles GG, Swerdlow AJ, et al. Incorporating progesterone receptor expression into the PREDICT breast prognostic model. Eur J Cancer 2022;173:178–93. https://doi.org/10.1016/j.ejca.2022.06.011.

[9] Maishman T, Copson E, Stanton L, Gerty S, Dicks E, Durcan L, et al. An evaluation of the prognostic model PREDICT using the POSH cohort of women aged 40 years at breast cancer diagnosis. Br J Cancer 2015;112:983–91. https://doi.org/10.1038/bjc.2015.57.

[10] de Glas NA, Bastiaannet E, Engels CC, de Craen AJ, Putter H, van de Velde CJ, et al. Validity of the online PREDICT tool in older patients with breast cancer: a population-based study. Br J Cancer 2016;114:395–400. https://doi.org/10.1038/bjc.2015.466.

[11] Engelhardt EG, van den Broek AJ, Linn SC, Wishart GC, Rutgers EJT, van de Velde AO, et al. Accuracy of the online prognostication tools PREDICT and Adjuvant! for early-stage breast cancer patients younger than 50 years. Eur J Cancer 2017;78:37–44. https://doi.org/10.1016/j.ejca.2017.03.015.

[12] Wong HS, Subramaniam S, Alias Z, Taib NA, Ho GF, Ng CH, et al. The predictive accuracy of PREDICT: a personalized decision-making tool for Southeast Asian women with breast cancer. Medicine 2015;94:e593. https://doi.org/10.1097/MD.0000000000000593.

[13] Copson ER, Maishman TC, Tapper WJ, Cutress RI, Greville-Heygate S, Altman DG, et al. Germline BRCA mutation and outcome in young-onset breast cancer (POSH): a prospective cohort study. Lancet Oncol 2018;19:169–80. https://doi.org/10.1016/S1470-2045(17)30891-4.

[14] van Maaren MC, van Steenbeek CD, Pharoah PDP, Witteveen A, Sonke GS, Strobbe LJA, et al. Validation of the online prediction tool PREDICT v. 2.0 in the Dutch breast cancer population. Eur J Cancer 2017;(86):364–72. https://doi.org/10.1016/j.ejca.2017.09.031.

[15] Karapanagiotis S, Pharoah PDP, Jackson CH, Newcombe PJ. Development and external validation of prediction models for 10-year survival of invasive breast cancer. Comparison with PREDICT and CancerMath. Clin Cancer Res 2018;24:2110–5. https://doi.org/10.1158/1078-0432.CCR-17-3542.

[16] Gray E, Marti J, Brewster DH, Wyatt JC, Hall PS, Group SA. Independent validation of the PREDICT breast cancer prognosis prediction tool in 45,789 patients using Scottish Cancer Registry data. Br J Cancer 2018;119:808–14. https://doi.org/10.1038/s41416-018-0256-x.

[17] Aguirre U, Garcia-Gutierrez S, Romero A, Domingo L, Castells X, Sala M, et al. External validation of the PREDICT tool in Spanish women with breast cancer participating in population-based screening programmes. J Eval Clin Pract 2019;25:873–80. https://doi.org/10.1111/jep.13084.

[18] Zaguirre K, Kai M, Kubo M, Yamada M, Kurata K, Kawaji H, et al. Validity of the prognostication tool PREDICT version 2.2 in Japanese breast cancer patients. Cancer Med 2021;10:1605–13. https://doi.org/10.1002/cam4.3713.

[19] Agostinetto E, Ameye L, Martel S, Aftimos P, Ponde N, Maurer C, et al. PREDICT underestimates survival of patients with HER2-positive early-stage breast cancer. NPJ Breast Cancer 2022;8(87). https://doi.org/10.1038/s41523-022-00452-8.

[20] Muranen TA, Morra A, Khan S, Barnes DR, Bolla MK, Dennis J, et al. PREDICT validity for prognosis of breast cancer patients with pathogenic BRCA1/2 variants. NPJ Breast Cancer 2023;9:37. https://doi.org/10.1038/s41523-023-00546-x.

[21] van der Meer DJ, Kramer I, van Maaren MC, van Diest PJ, S CL, Maduro JH, et al. Comprehensive trends in incidence, treatment, survival and mortality of first primary invasive breast cancer stratified by age, stage and receptor subtype in the Netherlands between 1989 and 2017. Int J Cancer 2021;148:2289–303. https://doi.org/10.1002/ijc.33417.

[23] Dackus GM, Ter Hoeve ND, Opdam M, Vreuls W, Varga Z, Koop E, et al. Long-term prognosis of young breast cancer patients (</=40 years) who did not receive adjuvant systemic treatment: protocol for the PARADIGM initiative cohort study. BMJ Open 2017;7:e017842. https://doi.org/10.1136/bmjopen-2017-017842.

[24] Verbeek AL, Broeders MJ. National Evaluation Team for Breast Cancer S, National E, Training Centre for Breast Cancer S. Evaluation of The Netherlands breast cancer screening programme. Ann Oncol 2003;14:1203–5. https://doi.org/10.1093/annonc/mdg324.

[25] NABON/NIV. Richtlijn Screening en Diagnostiek van het Mammacarcinoom; 2007. ⟨https://www.nabon.nl/oude-richtlijnen⟩ [Accessed 12 May 2023].
[26] Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. BMC Med Inf Decis Mak 2008;8:53. https://doi.org/10.1186/1472-6947-8-53.
[27] Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst 2006;98:1183–92. https://doi.org/10.1093/jnci/djj329.
[28] NABON/NIV. Federatie Medisch Specialisten - Richtlijnen database_Borstkanker adjuvante systemische behandeling; 2020. ⟨https://richtlijnendatabase.nl/richtlijn/borstkanker/adjuvante_systemische_therapie.html⟩ [Accessed 23 August 2023].
[29] Early Breast Cancer Trialists' Collaborative G. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687–717. https://doi.org/10.1016/S0140-6736(05)66544-0.
[30] R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2022.
[31] Cardoso F, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rubio IT, et al. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2019;30:1674. https://doi.org/10.1093/annonc/mdz189.
[32] Burstein HJ, Curigliano G, Thurlimann B, Weber WP, Poortmans P, Regan MM, et al. Customizing local and systemic therapies for women with early breast cancer: the St. Gallen International Consensus Guidelines for treatment of early breast cancer 2021. Ann Oncol 2021;32:1216–35. https://doi.org/10.1016/j.annonc.2021.06.023.
[33] Early Breast Cancer Trialists' Collaborative G, Davies C, Godwin J, Gray R, Clarke M, Cutter D, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet 2011;378:771–84. https://doi.org/10.1016/S0140-6736(11)60993-8.
[34] Early Breast Cancer Trialists' Collaborative G, Peto R, Davies C, Godwin J, Gray R, Pan HC, et al. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. Lancet 2012;379:432–44. https://doi.org/10.1016/S0140-6736(11)61625-5.
[35] Kramer I, Schaapveld M, Oldenburg HSA, Sonke GS, McCool D, van Leeuwen FE, et al. The Influence of adjuvant systemic regimens on contralateral breast cancer risk and receptor subtype. J Natl Cancer Inst 2019;111:709–18. https://doi.org/10.1093/jnci/djz010.
[36] Schaapveld M, Visser O, Louwman WJ, Willemse PH, de Vries EG, van der Graaf WT, et al. The impact of adjuvant therapy on contralateral breast cancer risk and the prognostic significance of contralateral breast cancer: a population based study in the Netherlands. Breast Cancer Res Treat 2008;110:189–97. https://doi.org/10.1007/s10549-007-9709-2.
[37] Bertelsen L, Bernstein L, Olsen JH, Mellemkjaer L, Haile RW, Lynch CF, et al. Effect of systemic adjuvant treatment on risk for contralateral breast cancer in the Women's Environment, Cancer and Radiation Epidemiology Study. J Natl Cancer Inst 2008;100:32–40. https://doi.org/10.1093/jnci/djm267.
[38] Loi S, Drubay D, Adams S, Pruneri G, Francis PA, Lacroix-Triki M, et al. Tumor-infiltrating lymphocytes and prognosis: a pooled individual patient analysis of early-stage triple-negative breast cancers. J Clin Oncol 2019;37:559–69. https://doi.org/10.1200/JCO.18.01010.
[39] Park JH, Jonas SF, Bataillon G, Criscitiello C, Salgado R, Loi S, et al. Prognostic value of tumor-infiltrating lymphocytes in patients with early-stage triple-negative breast cancers (TNBC) who did not receive adjuvant chemotherapy. Ann Oncol 2019;30:1941–9. https://doi.org/10.1093/annonc/mdz395.
[40] Kim RS, Song N, Gavin PG, Salgado R, Bandos H, Kos Z, et al. Stromal tumor-infiltrating lymphocytes in NRG oncology/NSABP B-31 adjuvant trial for early-stage HER2-positive breast cancer. J Natl Cancer Inst 2019;111:867–71. https://doi.org/10.1093/jnci/djz032.
[41] de Jong VMT, Wang Y, Ter Hoeve ND, Opdam M, Stathonikos N, Jozwiak K, et al. Prognostic value of stromal tumor-infiltrating lymphocytes in young, node-negative, triple-negative breast cancer patients who did not receive (neo)adjuvant systemic therapy. J Clin Oncol 2022;40(21):2361–74. https://doi.org/10.1200/JCO.21.01536.
[42] Wu Y, Mittendorf EA, Kelten C, Tucker SL, Wei W, Middleton LP, et al. Occult axillary lymph node metastases do not have prognostic significance in early stage breast cancer. Cancer 2012;118:1507–14. https://doi.org/10.1002/cncr.26458.
[43] Weaver DL, Krag DN, Ashikaga T, Harlow SP, O'Connell M. Pathologic analysis of sentinel and nonsentinel lymph nodes in breast carcinoma: a multicenter study. Cancer 2000;88:1099–107.
[44] de Boer M, van Deurzen CH, van Dijck JA, Borm GF, van Diest PJ, Adang EM, et al. Micrometastases or isolated tumor cells and the outcome of breast cancer. N Engl J Med 2009;361:653–63. https://doi.org/10.1056/NEJMoa0904832.
[45] Giuliano AE, Hawes D, Ballman KV, Whitworth PW, Blumencranz PW, Reintgen DS, et al. Association of occult metastases in sentinel lymph nodes and bone marrow with survival among women with early-stage invasive breast cancer. JAMA 2011;306:385–93. https://doi.org/10.1001/jama.2011.1034.
[46] Tan LK, Giri D, Hummer AJ, Panageas KS, Brogi E, Norton L, et al. Occult axillary node metastases in breast cancer are prognostically significant: results in 368 node-negative patients with 20-year follow-up. J Clin Oncol 2008;26:1803–9. https://doi.org/10.1200/JCO.2007.12.6425.
[47] NABON/NIV. Federatie Medisch Specialisten - Richtlijnen database_Screening buiten het BVO; 2020. ⟨https://richtlijnendatabase.nl/uploaded/docs/Mammacarcinoom/Flowchart_2.pdf?u=1RIVkG⟩ [Accessed 23 August 2023].
[48] Sessa C, Balmana J, Bober SL, Cardoso MJ, Colombo N, Curigliano G, et al. Risk reduction and screening of cancer in hereditary breast-ovarian cancer syndromes: ESMO Clinical Practice Guideline. Ann Oncol 2023;34:33–47. https://doi.org/10.1016/j.annonc.2022.10.004.