

Development and External Validation of Prediction Models for 10-Year Survival of Invasive Breast Cancer. Comparison with PREDICT and CancerMath

Solon Karapanagiotis¹, Paul D.P. Pharoah², Christopher H. Jackson¹, and Paul J. Newcombe¹



Abstract

Purpose: To compare PREDICT and CancerMath, two widely used prognostic models for invasive breast cancer, taking into account their clinical utility. Furthermore, it is unclear whether these models could be improved.

Experimental Design: A dataset of 5,729 women was used for model development. A Bayesian variable selection algorithm was implemented to stochastically search for important interaction terms among the predictors. The derived models were then compared in three independent datasets ($n = 5,534$). We examined calibration, discrimination, and performed decision curve analysis.

Results: CancerMath demonstrated worse calibration performance compared with PREDICT in estrogen receptor (ER)-positive and ER-negative tumors. The decline in discrimination performance was -4.27% (-6.39 to -2.03) and -3.21% (-5.9 to -0.48) for ER-positive and ER-negative tumors, respectively.

Our new models matched the performance of PREDICT in terms of calibration and discrimination, but offered no improvement. Decision curve analysis showed predictions for all models were clinically useful for treatment decisions made at risk thresholds between 5% and 55% for ER-positive tumors and at thresholds of 15% to 60% for ER-negative tumors. Within these threshold ranges, CancerMath provided the lowest clinical utility among all the models.

Conclusions: Survival probabilities from PREDICT offer both improved accuracy and discrimination over CancerMath. Using PREDICT to make treatment decisions offers greater clinical utility than CancerMath over a range of risk thresholds. Our new models performed as well as PREDICT, but no better, suggesting that, in this setting, including further interaction terms offers no predictive benefit. *Clin Cancer Res*; 24(9); 2110–5. ©2018 AACR.

Introduction

Accurate prognostic predictions are crucial in clinical decision making around adjuvant therapy following surgery for early breast cancer. Adjuvant therapies, such as chemotherapy, are associated with serious side-effects and so are typically only in the patient's interest when there is non-negligible risk of mortality. Inaccurate risk predictions can lead to over-prescription of chemotherapy to patients who do not require it, and unnecessarily increased mortality among patients whose risk is underestimated. Over the years, several predictive models have been proposed, that can aid in the decision to treat patients with systemic therapy, based on clinical and pathological factors (1).

One such tool, PREDICT, was proposed in 2010 (www.predict.nhs.uk). It provides predictions of 5- and 10-year breast cancer survival and the benefits of hormonal therapy, chemotherapy and trastuzumab (2). The model uses data on factors such as patient age, tumor size, tumor grade, number of positive nodes, estrogen receptor (ER) status, Ki67 status, mode of detection and HER2 status (3). PREDICT has been validated in many cohorts (4–6). An updated version improving performance has been published recently (7). Another prognostic tool, CancerMath, was proposed in 2009 (<http://cancer.lifemath.net/>). CancerMath models risk predictions as a function of tumor size, number of nodes and various other prognostic factors (8), and is also able to predict the impact of various adjuvant therapies. CancerMath has been validated using data from patients in the United States (9), as well as a South East population (10).

Although both PREDICT and CancerMath are widely used for cancer prognosis, to date, there are limited data on the concordance of these two prediction models. Only, Laas and colleagues (11), compared the overall performance of CancerMath and PREDICT using information from 965 women with ER-positive and HER2-negative early breast cancer from the United States and Canada. Further comparison in larger populations may be necessary. Furthermore, risk models are routinely evaluated using only standard performance metrics of discrimination and calibration. Such metrics give limited information as to the clinical value of a prediction model (12). This is because they

¹MRC Biostatistics Unit, University of Cambridge, Cambridge, United Kingdom.
²Department of Oncology, University of Cambridge, Cambridge, United Kingdom.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (<http://clincancerres.aacrjournals.org/>).

C.H. Jackson and P.J. Newcombe are joint senior authors.

Corresponding Author: Solon Karapanagiotis, MRC Biostatistics Unit, University of Cambridge, Institute of Public Health Forgive Site, Robinson Way, Cambridge, Cambridgeshire, CB2 0SR. Phone: 746-066-5830; E-mail: solon.karapanagiotis@mrc-bsu.cam.ac.uk

doi: 10.1158/1078-0432.CCR-17-3542

©2018 American Association for Cancer Research.

Translational Relevance

Accurate risk prediction in breast cancer patients is crucial for treatment decisions around adjuvant therapy. Two widely used prediction tools are PREDICT and CancerMath. To date, a formal comparison of their predictive performance taking into account their clinical utility has never been undertaken. In this article, using several large and independent validation cohorts, we compared their expected clinical usefulness using a decision-theoretic framework. We demonstrated that using PREDICT to make treatment decisions offers greater clinical utility than CancerMath over a range of risk thresholds, including those at which treatment decisions around chemotherapy are currently made in the UK. Moreover, we explored extensions to the original models using a novel variable selection methodology. We found no meaningful improvement, demonstrating surprising robustness of the PREDICT model in its current form. Overall, our results suggest that PREDICT is an accurate risk prediction tool and provides greater clinical utility compared with CancerMath.

focus purely on the accuracy of the predictions from the model and do not consider the clinical consequences of decisions made using the model (13) such as who should receive adjuvant therapy. For example, a false-positive prediction from the model would result in a patient being treated unnecessarily, whereas a false-negative prediction would result in a patient not getting a treatment that she would benefit from. The relative harms or benefits of these alternative clinical outcomes should be weighted appropriately when evaluating the models. Decision curve analysis (14) is a decision-theoretic method to compare prediction models with respect to their clinical utility. It provides a calculation of net benefit: The expected utility of a decision to treat patients at some threshold, compared to an alternative policy such as treating nobody. "Utility" is defined precisely below, however, conceptually, it balances the number of patients whose lives are saved against the cost of unnecessarily treating patients who would have survived anyway. In practice, adjuvant chemotherapy is recommended as a treatment option from the Cambridge Breast Unit (Addenbrooke's Hospital, Cambridge, UK) if a patient's 10-year predicted survival probability using chemotherapy, minus their predicted probability without, is greater than 5% (15). Hence, it would be of interest to evaluate whether or not using the existing thresholds alongside the current prediction models lead to better clinical utility.

The aim of this study was 3-fold. First, to perform the first direct comparison of CancerMath and PREDICT in a large European validation cohort, second, to explore whether modelling interactions among the risk factors used in PREDICT and CancerMath can lead to more accurate prognostic models, and finally, to exemplify the use of decision curve analysis as a means of comparing prognostic models.

Materials and Methods

Study population

Training data. The training data used for model development were based on data from 5,729 subjects with invasive breast cancer diagnosed in East Anglia, UK between 1999 and 2003

identified by the Eastern Cancer Registration and Information Centre (ECRIC). All analyses used data censored on December 31, 2012. For more details on the source of data and eligibility criteria see ref. 2.

Information obtained included age at diagnosis (years), ER status (positive or negative), tumor grade (I, II, III), tumor size (mm), number of positive lymph nodes, presentation (screening vs. clinical), and type of adjuvant therapy (chemotherapy, endocrine therapy, or both). Separate models were derived for ER-positive and ER-negative breast cancer. The outcome modeled was 10-year breast cancer-specific mortality (BCSM). BCSM is defined as time to death from breast cancer.

Validation data. The validation data consisted of three cohorts, the Nottingham/Tenovus Breast Cancer Study (1,944 subjects; ref. 16), the Breast Cancer Outcome Study of Mutation Carriers (981 subjects; ref. 17), and the Prospective Study of Outcomes in Sporadic and Hereditary Breast Cancer (2,609 subjects; ref. 18). The same predictors as in the training set were available. The outcome measure was again 10-year BCSM status. If BCSM status was missing, the patient's data were excluded from the analysis. CancerMath was validated on a subset of this validation dataset (3,431 ER-positive and 1,418 ER-negative subjects) whose tumor size and/or number of positive nodes did not exceed the values against which it was originally validated, 50 mm and 10 nodes, respectively.

Development of novel prognostic models. We developed novel prognostic models by building upon the existing PREDICT and CancerMath frameworks to incorporate interaction terms between the covariates. To this end, we used a Bayesian variable selection algorithm, Reversible Jump Markov Chain Monte Carlo (RJMCMC) to stochastically search for important interaction terms (19). We present two implementations of the algorithm. The first using a Weibull model for the baseline survival time (denoted Weibull RJMCMC hereafter), and the second with a logistic regression model for the binary outcome of 10-year survival (denoted Logistic RJMCMC hereafter). For both models, all predictors were transformed as in the latest version of PREDICT (7). The RJMCMC method searched over models with all combinations of pairwise interactions between the predictors listed, while including the main effects of all predictors in both models. For further details on the RJMCMC set-up see Supplementary Material.

Both PREDICT and CancerMath are web-based interfaces, implemented in Javascript. The Javascript code, was downloaded in April 2017, and translated into R (20) and subsequently cross-checked with a random subset of 20 patients, to verify the accuracy of the R translation. Some of the predictors used in PREDICT and CancerMath were not available in our data sets, such as the chemotherapy generation. We considered that all the patients received second generation chemotherapy, as the data were older than 10 years. All other missing predictors were treated as unknown.

Comparison of models in terms of predictive performance and decision curve analysis. We assessed the predictive performance of the models on the validation data by examining measures of calibration, discrimination, and clinical net benefit. 95% confidence intervals (CI) were obtained by bootstrapping using 1,000 re-samples of the validation set.

Calibration. Calibration was assessed graphically, by plotting the predicted 10-year mortality risks (*x* axis) against the observed probabilities (*y* axis), calculated using smoothing techniques (21). Ideally, if predicted and observed 10-year survival probabilities agree over the whole range of probabilities, the plots show a 45-degree line.

Discrimination. Discrimination is the ability of the risk score to differentiate between patients who do and do not experience the event during the study period. Model discrimination was assessed by calculating the area under the receiver-operator characteristic curve (AUC). The AUC is the probability that the predicted mortality of a randomly selected patient who died is higher than that of a randomly selected patient who survived; the higher the AUC, the better the model is at identifying patients with a worse survival.

Decision curve analysis. We used decision curve analysis to describe and compare the clinical consequences of the models (14). The net benefit of a model is the difference between the proportion of true positives (TP) and the proportion of false positives (FP) weighted by the threshold *R* (expressed on the odds scale) at which an individual is classified as "high risk":

$$\text{Net benefit} = \frac{TP}{n} - \frac{FP}{n} \left[\frac{R}{1-R} \right].$$

Here, *n* is the total sample size and *R* is the "high risk designation" threshold at which treatment is prescribed. The risk threshold (*R*) used by clinicians to prescribe treatment implies a relative value for either receiving treatment if death was likely or avoiding treatment if death was not likely. For example, if the clinician views unnecessary chemotherapy in nine women as an acceptable cost for correctly treating one high risk woman, this is equivalent to treating all women with $\leq 10\%$ mortality risk. The decision curve is constructed by varying *R* and plotting net benefit on the *y* axis against alternative values of *R* on the *x* axis. We calculated the net benefit of each prediction model and two reference strategies: treat none or treat all. The model with the highest net benefit at a particular risk threshold *R* enables us to treat as many high-risk people as possible while avoiding harm from unnecessarily treating people at low risk.

In practice, treatment for breast cancer is given if it is expected to reduce the predicted risk by some desired magnitude, or more. For instance, clinicians in the Cambridge Breast Unit (Addenbrooke's Hospital, Cambridge, UK) use the absolute 10-year survival benefit from chemotherapy to guide decision making for adjuvant chemotherapy as follows: <3% no chemotherapy; 3%–5% chemotherapy discussed as a possible option; >5% chemotherapy recommended (15). Assuming that chemotherapy reduces the 10-year risk of death by 22% (22), a target risk reduction between 3% and 5%, corresponds to risks between 14% and 23% on the decision curves. The R software, version 3.3.3 (20) was used for data analysis.

Results

The demographic, pathologic and treatment characteristics for the training and validation data are presented in Table 1. Estimated coefficients and the probabilities of inclusion in the model for each interaction term are given in the Supplementary Material

Table 1. Characteristics of women in the training and validation data sets

	Training data	Validation data
Total number of women	5,729	5,534
Breast cancer deaths, <i>n</i> (%)	994 (17)	1,079 (20)
Missing breast cancer deaths, <i>n</i> (%)	0 (0)	118 (2)
Other deaths, <i>n</i> (%)	680 (12)	208 (4)
Missing other deaths, <i>n</i> (%)	0 (0)	118 (2)
Median age at diagnosis (IQR), years	58 (17)	40 (13)
Median tumor size (IQR), mm	19 (13)	20 (15)
Missing tumor size, <i>n</i> (%)	0 (0)	157 (3)
Median number of nodes (IQR)	0 (2)	0 (2)
Grade, <i>n</i> (%)		
I	987 (18)	702 (13)
II	2,932 (51)	1,918 (35)
III	1,810 (32)	2,914 (53)
ER status, <i>n</i> (%)		
Positive	4,711 (82)	3,879 (70)
Negative	1,018 (18)	1,655 (30)
Adjuvant therapy, <i>n</i> (%)		
None	735 (13)	1,310 (24)
Chemo	909 (16)	1,562 (28)
Hormone	2,981 (52)	1,020 (18)
Both	1,104 (19)	1,642 (30)

(Supplementary Tables S2–S5 and Supplementary Fig. S1). Subsequent performance characteristics are presented only for the validation data.

Calibration

Figure 1 shows the calibration plots of the four prognostic models. Individual plots with 95% CIs can be found in Supplementary Material (Supplementary Fig. S2). CancerMath systematically underestimated the 10-year death probability, especially for the ER-negative tumors. The other models showed good calibration except for high-risk individuals, whose risk was consistently overestimated by all models. Compared to PREDICT, CancerMath demonstrated worse calibration performance.

Discrimination

Table 2 represents the percentage change in discriminative performance (PC_{AUC}) of the three models with respect to PREDICT. The PC_{AUC} was calculated as follows: $PC_{AUC} = (AUC_{\text{model}} - AUC_{\text{PREDICT}}) / AUC_{\text{PREDICT}} \times 100$. This allowed us to directly compare the between-model performance. The Weibull RJMCMC and Logistic RJMCMC models had superior discrimination to PREDICT for ER-positive subjects. Logistic RJMCMC also performed better than PREDICT in ER-negative tumors. However, none of these improvements were statistically significant. Again, CancerMath performed notably worse than PREDICT.

Decision curve analysis

Figure 2 demonstrates the decision curves for the four models. For risk thresholds ≤ 0.05 for ER-positive tumors and ≤ 0.15 for ER-negative tumors, no model provided benefit compared with a policy of treating all patients. For threshold probabilities around 0.55 or higher for ER-positive tumors and ≥ 0.60 for ER-negative tumors, the models also gave lower net benefit than a policy of treating no patients. Within the threshold range 0.05–0.55 for ER-positive tumors, and 0.15–0.60 for ER-negative tumors, CancerMath provided the lowest net benefit for both ER-positive and ER-negative tumors. All the other curves were very similar over these ranges, indicating that the net-benefit from using any of

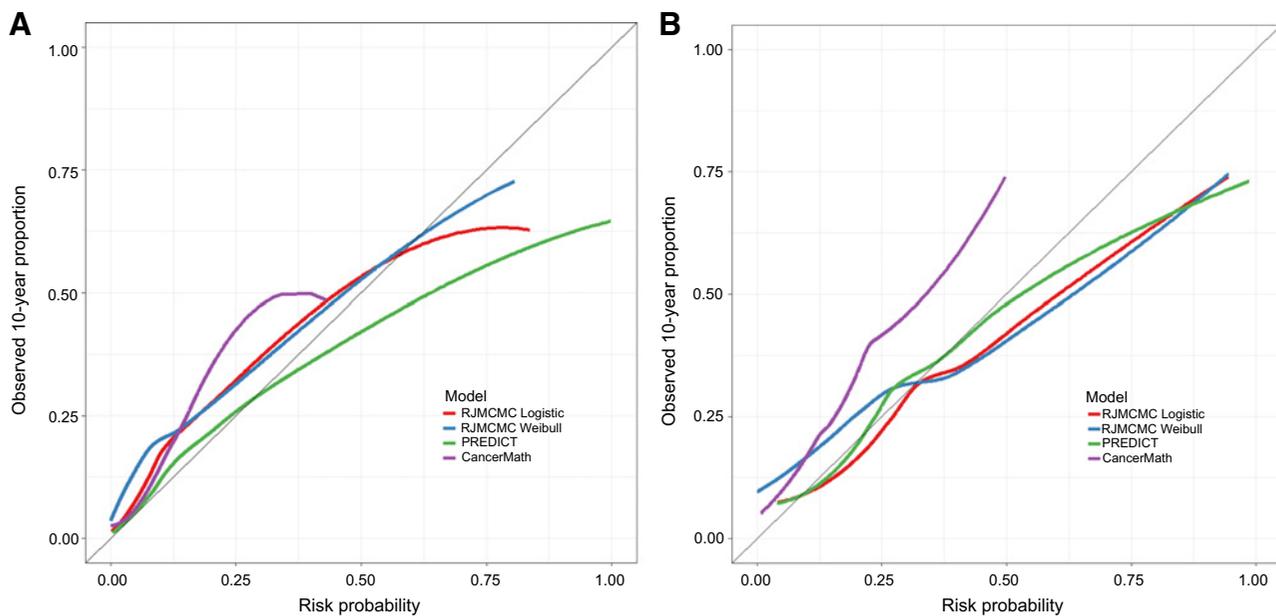


Figure 1.

Calibration plots of the four models in the validation set for (A) ER-positive, and (B) ER-negative tumors. Ten-year mortality risks based on the four prognostic models are plotted against the observed mortality rate. The 45-degree line represents the perfect calibration slope.

these models to prescribe chemotherapy, compared with treating nobody, could be as high as 0.1, which is of equivalent value to treating an additional 10% of patients who would have died without treatment.

As already mentioned, risks between 14% and 23% correspond to absolute risk reductions from chemotherapy treatment, in the range 3%–5%. In this range adjuvant chemotherapy is discussed as a treatment option (15). Risk >23% correspond to absolute risk reductions of 5% or more where chemotherapy is firmly recommended. We see that in both thresholds ranges PREDICT outperformed CancerMath. Individual decision curve plots with 95% CIs can be found in Supplementary Material (Supplementary Fig. S3).

Discussion

Accurate prognostic tools are critical to personalized breast cancer treatment, allowing clinicians and patients to make more informed decisions about adjuvant therapy. Using several large European validation cohorts, we have performed the first detailed comparison of the two most widely used breast cancer prognostic models: PREDICT and CancerMath. We found that survival probabilities from PREDICT offer both improved accuracy and discrimination over CancerMath. Using decision curve analysis,

we also demonstrate that the use of PREDICT to make treatment decisions offers greater clinical utility than CancerMath over a range of risk thresholds, including those at which adjuvant chemotherapy is currently recommended by clinicians in the UK. We also explored novel prognostic models, extending PREDICT to include interaction terms among the predictors. Our new models performed as well as PREDICT, but no better, suggesting that, in this setting, the incorporation of interaction terms offers no predictive benefit.

Although all models demonstrated good overall prediction of 10-year mortality for both ER-positive and -negative tumors, they consistently overestimated risk in women with the worse prognosis. This highlights the shortcoming of current prediction models, showing significant discrepancies especially in high risk populations. Laas and colleagues (11), arrived at the same conclusion but using a smaller sample of ER-positive HER2-negative breast cancer subjects. This could have consequences for the individual patient if their survival probability for breast cancer is underestimated.

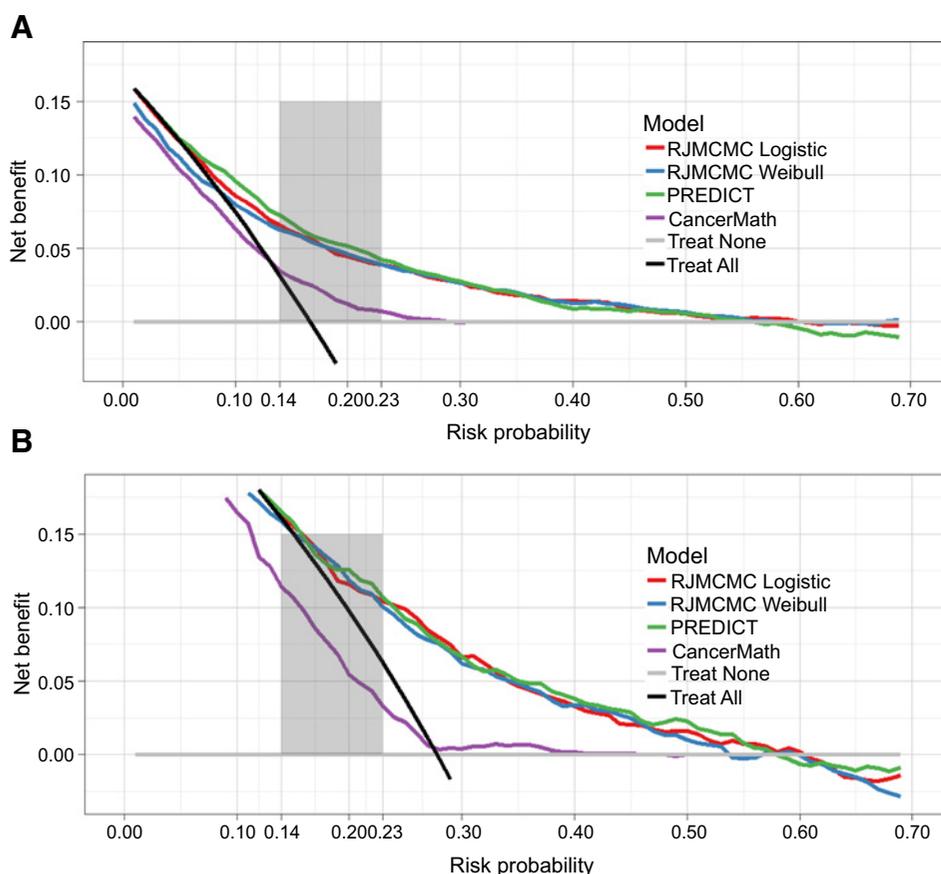
We further applied decision curve analysis because this offers important information beyond the standard performance metrics of discrimination and calibration (14, 23). To our knowledge, this is the first study to take into account the clinical utility of treatment decisions in breast cancer risk models. The main clinical utility of breast cancer prognostic models is to facilitate the decision of whether a patient will benefit from systemic treatment or not. Treatment is often prescribed at a risk threshold, which implies the clinician's or patient's relative value of unnecessary treatment versus worthwhile use. From our results, both our models and PREDICT would provide the greatest clinical utility when recommending treatment to patients with risks falling in the range of 0.05 and 0.55 for ER-positive tumors and between 0.15 and 0.60 for ER-negative

Table 2. Percentage change in AUC (PC_{AUC}) between PREDICT and each of the other models

PREDICT vs.	ER positive PC_{AUC} (%; 95% CI)	ER negative PC_{AUC} (%; 95% CI)
CancerMath	-4.27 (-6.39 to -2.03)	-3.21 (-5.9 to -0.48)
Logistic RJMCMC	0.04 (-1.04-1.12)	0.19 (-0.19-0.57)
Weibull RJMCMC	0.06 (-0.96-1.11)	-0.40 (-0.83-0.03)

NOTE: The 95% confidence intervals (CI) are based on 1,000 bootstrap replicates.

Karapanagiotis et al.

**Figure 2.**

Decision curve analysis for 10-year mortality for each model for (A) ER-positive, and (B) ER-negative tumors. The y axis measures net benefit. The 0.14 to 0.23 shaded area on the x axis corresponds to 3%–5% absolute risk of death reduction with and without chemotherapy assuming a relative risk reduction of 22% if treated. These are the risk ranges where chemotherapy is discussed as a treatment option. A model is of clinical value if it has the highest net benefit compared with the simple strategies of treating all patients (black line) or no patients (horizontal gray line) across the full range of threshold probabilities at which a patient would choose to be treated. The units on the y axis may be interpreted as the benefit associated with one patient who would die without treatment and who receives therapy.

tumors. It has been suggested that when the absolute 10-year survival benefit from chemotherapy is 3% to 5%, chemotherapy is discussed as a treatment option (15). We demonstrated that in this range PREDICT outperformed CancerMath. Note that these calculations assumed a 22% risk reduction when treated with second-generation chemotherapy (22). However, it is important to emphasize that there are no studies to date that have evaluated what would be an acceptable range of threshold probabilities that would help with adjuvant therapy recommendation.

Several potential limitations of our study should be considered. First, our comparison focused on web-based models, which are readily available for use in daily practice and incorporate only a modest set of risk factors which would be measured as a matter of course. We also acknowledge that other models may provide more accurate predictions, for example by incorporating more biomarker information, such as genetics or other "omics." Second, some data were missing from the validation cohorts, such as the chemotherapy generation, which were required in the PREDICT and CancerMath models. As an approximation, we assumed that all patients received second-generation chemotherapy since the data were older than 10 years. Because this same approximation was used in the fitting and validation of all models, we think it is unlikely to add much bias to their relative performance. We also note that any patient with missing information (predictors or outcome) was excluded from the training and validation data. Overall this was a small number of patients (<5%) so is unlikely to have biased our results.

In summary, we have compared two widely used prediction models for survival in patients with early-stage breast cancer, PREDICT and CancerMath. The former outperformed the latter in discriminatory accuracy, calibration, and clinical utility. We also developed novel prognostic models that performed as well as PREDICT, indicating that the incorporation of interaction terms offers only marginal predictive benefit. We have shown that PREDICT offers robust predictive performance, equivalent to our more complex modeling strategy, and for the first time quantified its clinical utility in identifying high-risk women that would benefit from adjuvant treatment. We did not elicit the relative benefits and harms of treatment directly from patients or clinicians, but inferred them from the risk threshold currently used to prescribe chemotherapy. Further research will be required to more precisely determine relative utilities of treatment versus no treatment, which could lead to new risk thresholds to focus decision curve analysis on.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: S. Karapanagiotis, P.D.P. Pharoah, P.J. Newcombe
 Development of methodology: S. Karapanagiotis, C.H. Jackson, P.J. Newcombe
 Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): S. Karapanagiotis, P.J. Newcombe
 Writing, review, and/or revision of the manuscript: S. Karapanagiotis, P.D.P. Pharoah, C.H. Jackson, P.J. Newcombe

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): P.D.P. Pharoah

Acknowledgments

This work was supported by the Medical Research Council (www.mrc.ac.uk; MC_UU_00002/9 to S. Karapanagiotis and P.J. Newcombe, MC_UU_00002/11; to C.H. Jackson) and Cancer Research UK (C490/A16561; to P.D.P. Pharoah). P.D.P. Pharoah, P.J. Newcombe, and S. Karapanagiotis were also supported by the National Institute for Health Research Biomedical Research

Centre at the University of Cambridge. We thank the anonymous reviewers for their constructive comments, which helped improve the article.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received November 27, 2017; revised January 23, 2018; accepted February 11, 2018; published first February 14, 2018.

References

- Engelhardt EG, Garvelink MM, de Haes JHCJM, van der Hoeven JJ, Smets EM, Pieterse AH, et al. 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* 2013;32:238–50.
- 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.
- 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.
- 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.
- 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.
- Wong H-S, 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 (Baltimore)* 2015;94:e593.
- 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.
- Chen LL, Nolan ME, Silverstein MJ, Mihm MC Jr, 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.
- Michaelson JS, Chen LL, Bush D, Fong A, Smith B, Younger J. Improved web-based calculators for predicting breast carcinoma outcomes. *Breast Cancer Res Treat* 2011;128:827–35.
- Miao H, Hartman M, Verkooijen HM, Taib NA, Wong HS, Subramaniam S, et al. Validation of the CancerMath prognostic tool for breast cancer in Southeast Asia. *BMC Cancer* 2016;16:820.
- Laas E, Mallon P, Delomenie M, Gardeux V, Pierga JY, Cottu P, et al. Are we able to predict survival in ER-positive HER2-negative breast cancer? A comparison of web-based models. *Br J Cancer* 2015;112:912–7.
- Vickers AJ, Cronin AM. Traditional statistical methods for evaluating prediction models are uninformative as to clinical value: towards a decision analytic framework. *Sem Oncol* 2010;37:31–8.
- Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ* 2016;352:i6.
- Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak* 2006;26:565–74.
- Down SK, Lucas O, Benson JR, Wishart GC. Effect of PREDICT on chemotherapy/trastuzumab recommendations in HER2-positive patients with early-stage breast cancer. *Oncol Lett* 2014;8:2757–61.
- 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.
- 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.
- Copson E, Eccles B, Maishman T, Gerty S, Stanton L, Cutress RI, et al. Prospective observational study of breast cancer treatment outcomes for UK women aged 18–40 years at diagnosis: The POSH study. *J Natl Cancer Inst* 2013;105:978–88.
- Newcombe P, Raza Ali H, Blows F, Provenzano E, Pharoah PD, Caldas C, et al. Weibull regression with Bayesian variable selection to identify prognostic tumour markers of breast cancer survival. *Stat Methods Med Res* 2017;26:414–36.
- R Core Team. R: A Language and Environment for Statistical Computing. R Found Stat Comput 2017;{ISBN} 3–900051-07-0. Available from: <http://www.R-project.org>.
- Austin PC, Steyerberg EW. Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. *Stat Med* 2014;33:517–35.
- Group EBCTC, others. 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.
- Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* 2014;35:1925–31.

Clinical Cancer Research

Development and External Validation of Prediction Models for 10-Year Survival of Invasive Breast Cancer. Comparison with PREDICT and CancerMath

Solon Karapanagiotis, Paul D.P. Pharoah, Christopher H. Jackson, et al.

Clin Cancer Res 2018;24:2110-2115. Published OnlineFirst February 14, 2018.

Updated version Access the most recent version of this article at:
doi:[10.1158/1078-0432.CCR-17-3542](https://doi.org/10.1158/1078-0432.CCR-17-3542)

Supplementary Material Access the most recent supplemental material at:
<http://clincancerres.aacrjournals.org/content/suppl/2018/04/06/1078-0432.CCR-17-3542.DC2>

Cited articles This article cites 22 articles, 1 of which you can access for free at:
<http://clincancerres.aacrjournals.org/content/24/9/2110.full#ref-list-1>

E-mail alerts [Sign up to receive free email-alerts](#) related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, use this link
<http://clincancerres.aacrjournals.org/content/24/9/2110>.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.