Bachelor Thesis

Prediction models for Breast cancer

Name: Barbara Tip
Student number: s1592998
First supervisor: Prof. Dr. S. Siesling
Second supervisor: Dr. H.Koffijberg
External organisation: Evidencio
Supervisor external organisation: T.Hueting

Date: 17-07-2018

UNIVERSITY OF TWENTE.

Abstract

Background: To support clinical decision making in the current medical practice, a growing number of clinical prediction models are being developed. Yet, the application of these models in daily medical practice falls behind. It is uncertain how many prediction models exist to support decision making in breast cancer care and how capable they are at supporting clinical decisions. The objective of this study was to identify and assess clinically relevant prediction models regarding breast cancer care.

Method: A literature search was performed to identify as many as relevant clinical prediction models regarding breast cancer developed between January, 1st 2010 and June 2018. Prediction models concerning breast cancer in men and women were included. Prediction models were reported in an overview using the following variables; country of origin, year of publication, predicted outcome, model features, performance (concordance index), calculation method, presentation, internal/external validated, number of patients in the validation cohort, transparency, and publication journal. Thereafter, critical analysis of the articles and prediction models were performed on quality of reporting. In addition, to search if there is a relationship between the impact of the journals and the (in)completeness of the models. The identified prediction models were assessed on transparency and reproducibility. Reproducible models were translated to user-friendly online calculators. Finally, to identify when models can be used in Dutch breast cancer care, these models were mapped to the Dutch breast cancer pathway.

Results: A total of 91 studies were included, describing the development of 142 prediction models. The most common calculation methods regarding the prediction models were the cox proportional hazard method (75) and the logistic regression method (65). Most of the articles and prediction models were developed using patients from China and the USA. Out of the 142 prediction models, 101 were presented as a nomogram and 26 models as a table. The most common models features (amount of times used) were patient age (65) and ER status (62). Twenty-four (26.4%) of the studies described the prediction models with full transparency. The overall quality of reporting was poor as 111 models did not have all the data available. Models lacking transparent description were usually reported without the model intercept or baseline hazard. There is no relationship found between the impact of the journal where the study was published in and the quality of reporting. Also, the reporting quality of the prediction models did not increase after the official publication of the TRIPOD statement in 2015. Lastly, 24 full transparent models were uploaded on Evidencio and 14 prediction models were uploaded after calculating the missing data. A total of 38 models were uploaded on Evidencio and using the digital breast cancer guideline Oncoguide, the models were mapped to their corresponding decision moment.

Conclusion: Generally, the quality of reporting is poor, most articles did not report all the necessary parameters to reproduce a clinical prediction model. To improve the application of prediction models for breast cancer care, the quality of reporting must be better. All models on Evidencio were assigned to the location in the guideline of which the model may support clinical decision making. Further assessment is necessary on the clinical impact and validity of the models before implementing them in the guideline.

Keywords

Prediction model, prediction, breast cancer, Dutch care path

Table of Contents

Introduction

In the current Dutch medical practice breast cancer is listed as the most common cancer diagnosed in women being 28% of all cancers in women[1]. Among men, breast cancer is diagnosed 100 times per year. Approximately 2.500 women are diagnosed per year with non-invasive breast cancer, and circa 14.500 with invasive breast cancer. On average, the 5-year mortality rate is 22% in breast cancer patients. The relatively low mortality is due to the early detection through population screening. In addition, better staging and providing personalized care play a major role in recent years [2].

Breast cancer can be roughly categorized into invasive cancer or in situ cancer. In situ breast cancer is further classified as ductal or lobular, where invasive or infiltrating cancer can be classified into other types or mixed forms of ductal and lobular[3]. To personalize treatment of breast cancer, information regarding patient characteristics is required, such as the hormone receptor status for estrogen receptor (ER) and the progesterone receptor (PgR)[4]. Also, background information of the patient is required, for example age, menopause, human epidermal growth factor receptor 2 and other co-morbidities[4]. By making a distinction between the before mentioned types of breast cancer, their classified cancer stages, background information of the patient, and the patient's personal preferences, it can be determined which care is desired [5].

According to the guidelines[5], a care process must be a continuing process. The transparency and the quality of care increases by constant monitoring and modification of the care path. This care path is a logistic range of care processes that a patient goes through from the primary care until survival or passing. In addition, the provided care exists of a combination of intra- and extramural care which can largely differ per patient given the unique characteristics [5]. Care paths can be personalised based on individual risks, patient preferences and the social context.

A care path contains a lot of decision and prediction moments. These decision and prediction moments are summarized in the Oncoguide [6]. The Oncoguide consists of different decision trees which have been developed using the "Nationaal borstkanker overleg Nederland" (NABON) guideline. Oncoguide can therefore be applied to all Dutch breast cancer patients. Oncoguide's decision trees can be accessed at www.oncoguide.nl. To support these decisions, clinical prediction models can be used. Clinical prediction models support physicians in tailoring the treatment to the needs of an individual patient.

Clinical prediction models are used to estimate the probability of a certain event with empirically substantiated tools [7]. These tools combine patient specific characteristics with a predictive value to predict the probability that an illness is present, or a certain disease status will occur. For example, the 5-years disease specific survival. A requirement in this situation is that the model is well calibrated and reliable. In other words, that the model has shown a good performance on the (external) validation.

Another important factor of clinical prediction models is decision support [8], including decisions about the need to continue with diagnostic tests and therapies. Such decisions are mostly binary and require the definition of clinical relevant decision thresholds. Therefore, a decision can be supported by a range of choice moments concerning diagnostic, treatment, screening, prognosis and disease prevention. Frequently used prediction models are nomograms and regression models.

Over the past decades there has been an exponential growth of published clinical prediction models regarding breast cancer [9]. Due to the lack of validation and implementation of the models, multiple models predicting the same outcome were developed. This results in confusion among doctors regarding the applicability of the models, especially when the results of the models are conflicting [10]. Consequently, interpretation becomes more difficult and may leads to less attractive use of the model. Recently, de Buy Wenniger concluded that it is about

time to assess existing models on their usefulness instead of developing new prediction models [11]. For example, by comparing and improving them.
To improve the reporting of prediction modelling studies, the "Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD)" statement has been published in 2015[12]. This statement has been endorsed by an international group consisting of doctors, researchers, statisticians, and scientific journalists who all looked at prediction models from their own perspective. This has led to a complete checklist of 22 criteria which have found to be essential in the initiation and validation of a reliable model.

Regarding the validation of prediction models, the TRIPOD statement recommends the use of model discrimination and model calibration as valuable performance measures. Discrimination is the ability of a model to separate individuals who experience the event of interest from those who will not and, is mostly defined using the Concordance statistics (C-index). The C-index value 0.5 and 1.0 correspond to respectively poor and a perfect diagnostic ability[13]. Calibration is mostly assed using a calibration plot which shows the agreement between the predicted and the observed probabilities. The calibration is often described using the calibration slope which ideally is equal to 1, and the calibration intercept (or calibration-in-the large) and is ideally equal to 1[14].

To improve the accessibility and use of prediction models, an online platform has been developed to translate clinical prediction models to user-friendly online calculators. Evidencio enables users to use, create, validate and integrate clinical prediction models. The goal of Evidencio is to connect prediction models from science with the clinical world by making the consequences of choices visible[15].

It is currently unknown which prediction models for breast cancer are available, how well the quality of reporting is, and when these models are clinically applicable in the Dutch breast cancer pathway. The goal of this research is to make a clear overview of the clinical prediction models concerning breast cancer, and increase the amount of usable clinical prediction models in Evidencio and merge the models in the Dutch care path.

Method

Literature search

First, a literature search was performed to identify as much as possible relevant publications concerning breast cancer prediction models in the bibliographic databases Scopus and Google Scholar. To increase the relevance of the findings of this review for current clinical practice we only included papers published from January, 1st 2010 up to June 2018. Search terms for "breast cancer" or "mamma carcinoma" were used in combination with search terms for "prediction model", "nomogram", "prediction" and "validation". The reference list of relevant identified articles were also searched for additional relevant publications, hereby it is possible that prediction models before 2010 are included. Only papers published in English were assessed.

Model selection

All titles and abstracts were screened. When an article described the developing of a new prediction model, the full text of the studies were read. All identified prediction models were reported using the following variables; country of origin, year of publication, predicted outcome, model features, performance (concordance index), calculation method, presentation, internal/external validated, number of patients in the validation cohort, transparency, and publication journal. The overview of identified prediction models was added as supplementary data.

Model evaluation

The articles and models were assessed on quality of reporting. Quality of reporting was based on two terms; complete and incomplete data. Complete data meant that all the necessary data to reproduce the prediction model was reported. Incomplete data meant that a part or all the data necessary to reproduce the prediction model was missing, for example the baseline hazard or the coefficients. Missing data did not include the study group information. Secondly, critical analysis of the articles and prediction models were performed on quality of reporting. In addition, to search if there is a relationship between the impact of the journals and the (in)completeness of the models. The models and journals were compared using the h-index. Journal h-index is one measure of the quality of a journal [16]. The h-index is best used to compare journals within a field. Lastly, the most common missing data was identified.

Translate models to online calculators on Evidencio

Many as possible models of which the underlying statistical formula could be derived were translated to online calculators on www.evidencio.com.

Prediction models can be developed using different analytic methods [17]. The two most common methods are the logistic regression and the cox proportional hazard regression. A logistic regression model is commonly used to predict a binary endpoint and is usually used in diagnostic models. In logistic regression, the predicted probability of the outcome event is calculated using the following formula:

Equation 1 Logistic regression

$$P = \frac{e^{X\beta}}{1 + e^{X\beta}}$$

Where β: intercept in model, X regression coefficient (= log odds ratio). For the transparency of a logistic regression model it is important that the intercept and the regression coefficients or log odds ratios are given [17].

A Cox proportional hazards regression model is used for time-to-time-event outcomes and is used for long-term prognostic outcomes. In a Cox proportional hazards regression, the predicted probability of the outcome event is calculated using the following formula:

Equation 2 Cox proportional hazard regression

$$P(t) = H_0(t)^{\exp(x\beta)}$$

Where H_0 (t): baseline hazard, xβ is the model linear predictor estimated by the summation of the coefficients multiplied by their corresponding variable input. (e.g.$\beta_{var1} \cdot var1 + \beta_{var2} \cdot var2 + \beta_{varX} \cdot varX$). For the transparency of a cox proportional hazard regression model it is important that the baseline hazard and the regression coefficients or the hazard ratios are given [17].

If the models were not described in full, the outcome for an individual could be derived for some models if the formula was presented in for instance a nomogram, table or risk score.

Position in Dutch care path

A care path contains a lot of decision and prediction moments. These decision and prediction moments are summarized in the Oncoguide. Oncoguide offers care providers digital decision support displayed in decision trees based on guidelines. The goal of Oncoguide is to support health care professionals and patients when taking treatment decisions[18].

To assess the clinical applicability, all prediction models in Evidencio were mapped to the Dutch care path for breast cancer patients. The Dutch care path for breast cancer patients consists of multiple decision trees which are summarized in the Oncoguide[6]. Figure 1 is the simplest version of the Dutch care path. By looking at the variable 'prediction' from the Excel overview and combine this with the decision moments in the decision tree it will be possible to implement the prediction models in the care path. This leads to more clarity about which models can be applied on which decision moment.

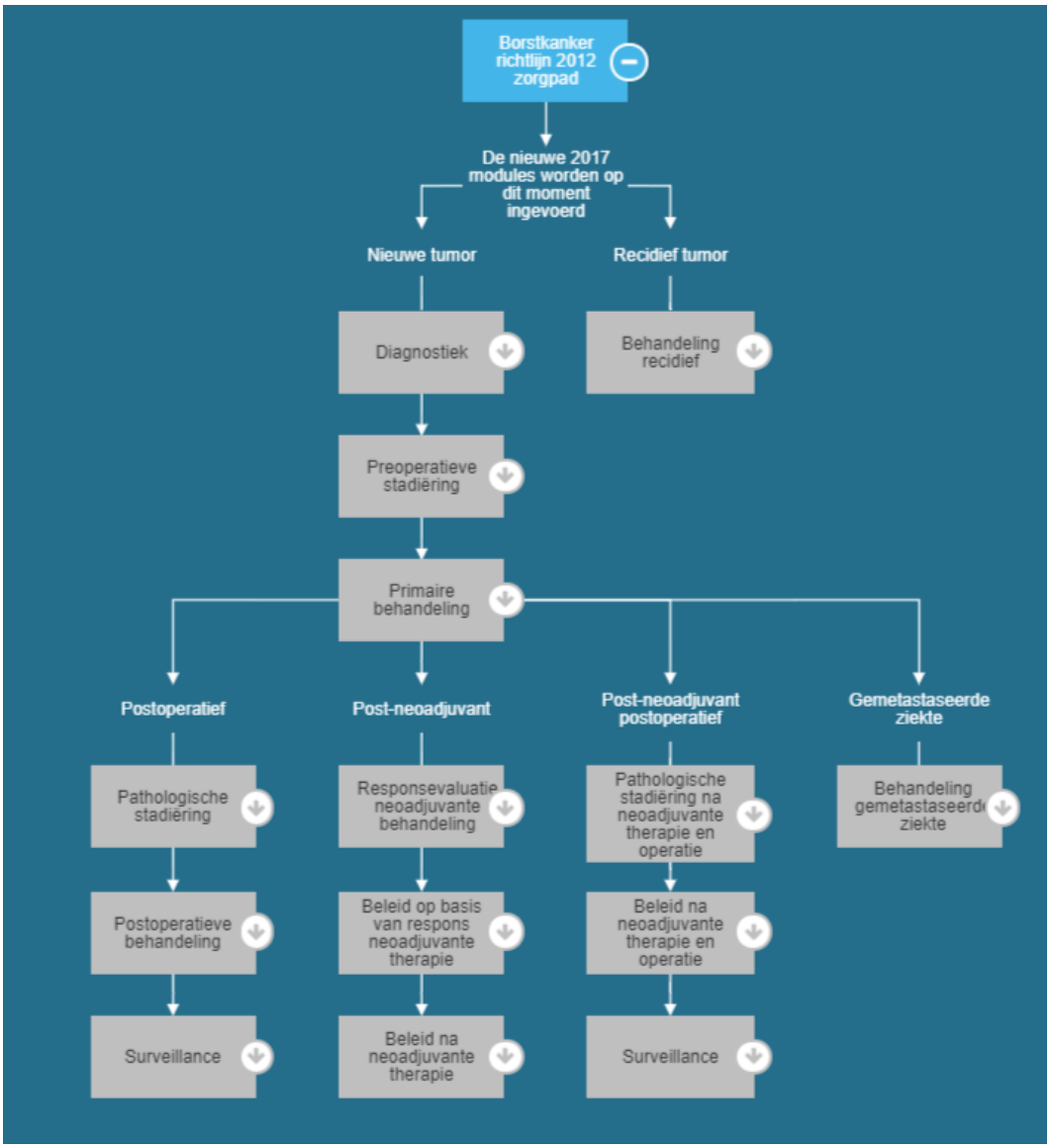

Figure 1Decision tree Dutch care path breast cancer [6]

Results

Literature search

A total of 105 articles were identified, of which 91 were eligible for inclusion in the review. These 91 articles combined described a total of 142 prediction models for breast cancer patients. The 91 development studies are shown in the Excel file. 14 studies were not included in the review due to missing full text, not written in English or if it was a validation study of an already included model. Figure 2 shows the overview of the whole search process.

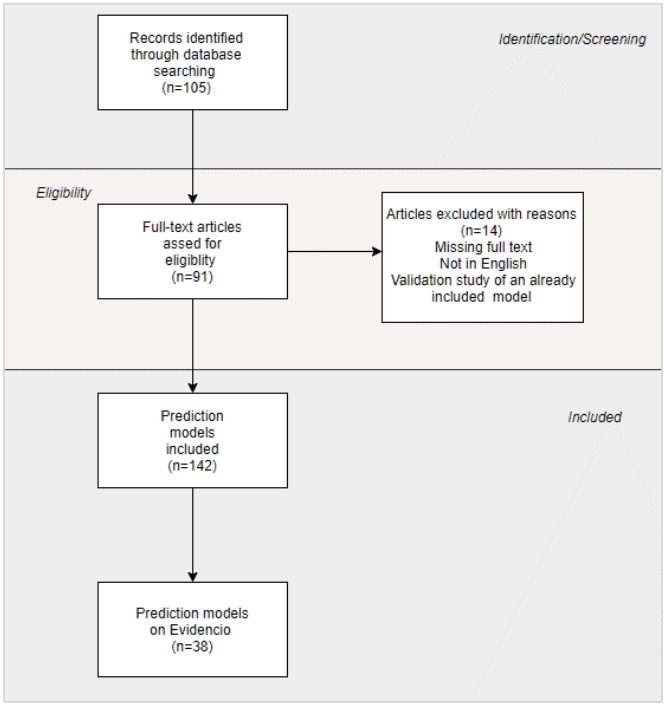

Figure 2 Search process

Model selection

As shown in figure 2, 91 articles were included in the overview. The 91 articles included 142 different prediction models. An example of the overview is shown in Appendix 2. The full overview is available in the attached Excel file.

Model evaluation

Most of the articles and prediction models were developed using patients from China and the USA. Twenty articles (22%) and 39 (28%) prediction models were published in China and 22 (24%) articles and 28 (20%) prediction models in the USA. The whole overview of the publication countries can be found in appendix 3. Out of the 142 prediction models, 101 were presented as a nomogram, 27 models as a table, 6 models as a formula, 3 models as a decision tree, and 3 models as a scorecard. Figure 3 shows the presentations methods per country.

All of the prediction models combined identified more than 100 different features to predict a breast cancer related outcome. The most common models features (amount of times used) were; patient age (65), ER status (62), nodal stage (53), HER2 status (43), tumor grade (41), PgR (40), tumor size (34), tumor stage, lymphovascular space invasion (20), and histology (17).

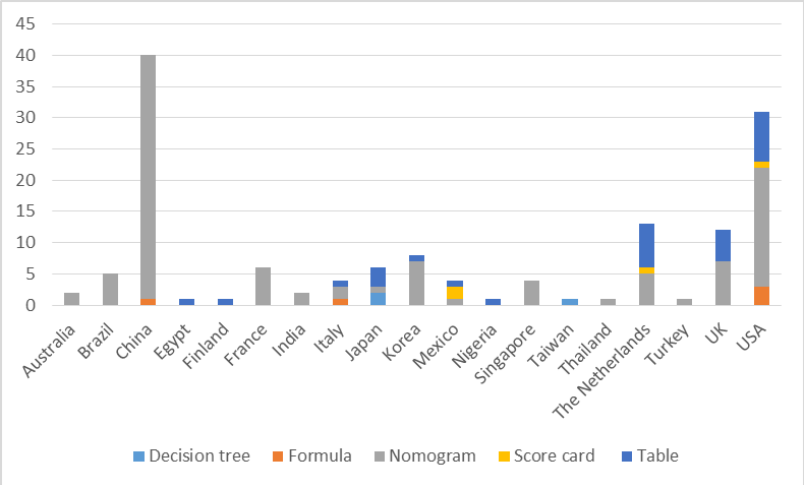

Figure 3 Model presentation methods per publication country

Concerning the calculation method, 55 articles used the logistic regression method and 35 articles used the cox proportional hazard method. The two remaining articles used another calculation method. Regarding the prediction models, 75 used the cox proportional hazard method and 65 the logistic regression method. The remaining two prediction models used other calculation methods. The total number of patients used to develop the prediction model is summarized in table 1. The median number of patients used to develop a prediction model is 724 (IQR: 348, 2044). The smallest patients population contained 64 patients, and the largest population contained 2,392,998 patients.

Table 1 Population group

	N
Min	64
Q1	348
Median	724
Q3	2044
Max	2392998

General

All 91 studies were reported in the excel file and can be found in the supplemental material. Figure 4 shows the division of studies providing complete and incomplete description of the developed models. A total of 67 (73.6%) studies were lacking proper reporting on the development of the prediction model. The remaining 24 (26.4%) studies provided a transparent description of the developed prediction model.

The 91 articles described a total of 142 prediction models for breast cancer patients. Of these 142 models, 31 prediction models contained all the relevant data to reproduce the model. The other 111 did not have all the data available. As can be seen in figure 5.

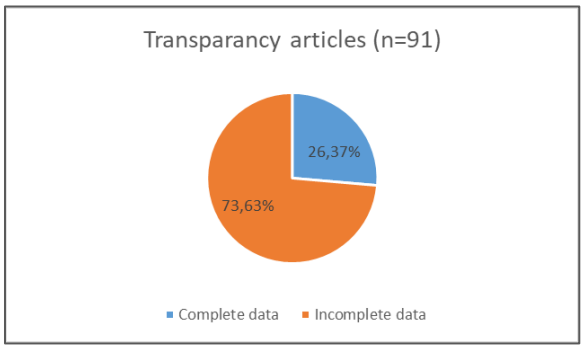

Figure 4 Transparancy articles

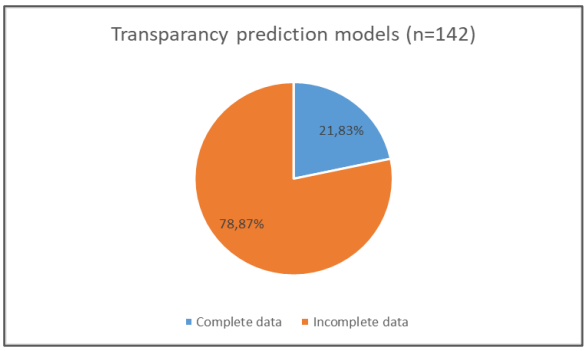

Figure 5 Transparancy prediction models

A total of 142 prediction models were assessed. The underlying statistical formula was fully/clearly presented for 31 models. The formula could be derived for 96 models, but 16 models could not be reproduced. Table 2 shows the exact missing data of the prediction models.

Table 2 Missing data prediction models

Data missing	**Number of prediction models**	**%**
Baseline hazard	50	45,05%
Baseline hazard and nomogram	5	4,50%
Baseline Hazard is missing and the predicted value scale of the nomogram is incorrect	3	2,70%
Baseline Hazard and coefficients	14	12,61%
Coefficients	2	1,80%
Intercept and coefficients	6	5,41%
Intercept and nomogram	7	6,31%
Intercept	24	21,62%
	111	**100%**

Publication year and journals

Figure 6 shows the amount of (in) complete studies per year of publication. From 2010 to 2017, the percentage of incomplete articles increased from 66.67 percent to 92.86 percent.

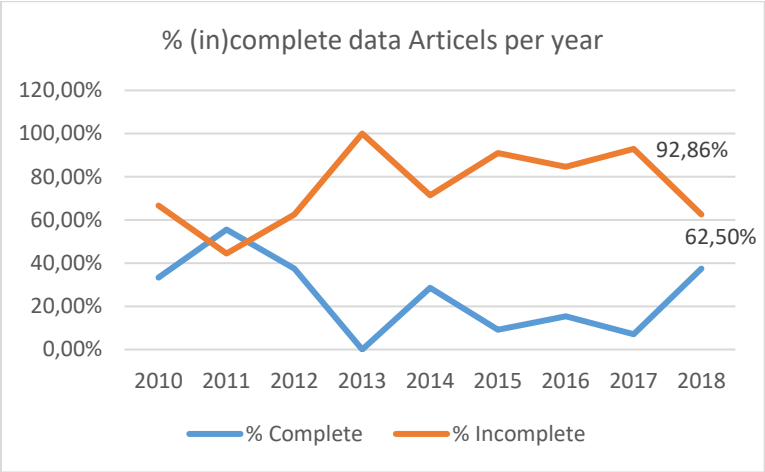

Figure 6 % (not) transparent articles per publication year

Tabel 3 Number of (in)complete articles per publication year

Publication year	2010	2011	2012	2013	2014	2015	2016	2017	2018
Complete	2	5	3	0	2	1	2	1	3
Incomplete	4	4	5	3	5	10	11	13	5
Total	6	9	8	3	7	11	13	14	8

Thereafter, we also investigated whether there will be a relationship between the impact factor (h-index) of the journal in which the study was published and the transparency of the article. Because of the small number of articles per journal it is not reliable to conclude if there is a relationship. Appendix 4 gives an overview of studies and the corresponding journal.

Position in Dutch Care path

After the analyses of the articles and prediction models, 38 models have been translated and updated to Evidencio. These 38 models came from 30 different articles. The 38 prediction models were totally random chosen. Table 4 is an overview of the prediction models on Evidencio with corresponding prediction.

Table 4 Prediction models on Evidencio with corresponding prediction.

PREDICTION MODEL	PREDICTION
[9]	likelihood of achieving axillary pathologic complete response to neoadjuvant chemo(immuno)therapy
[84]	Risk of breast cancer
[83]B PREDICT	10- year overall survival
[83]A PREDICT	5- year overall survival
[82]	Probability of axillary lymph node metastasis in early breast cancer patients with positive axillary ultrasound
[81]	risk of breast cancer in intraductal neoplasms with nipple discharge
[80]B INFLUENCE	4-year loco regional recidief
[80]A INFLUENCE	3- year recurrence risk
[80] E INFLUENCE	2-year locoregional recurrence
[80] D INFLUENCE	5-year locoregional recurrence
[80] C INFLUENCE	1-year locoregional recurrence
[8]	predict individual probability of BC in radiological lessions classified as BI-RADS Category 4
[71]	Likelihood of N2 orN3 stage in clinical T1-2NOMO
[70]	sentinel lymph node metastasis
[7]	Nonsentinel lymph node metastasis
[6] PKUPH	Probability of Non-sentinel lymph node metastasis
[5]B	3-Year Overall Survival
[5]A	1-Year Overall Survival
[4]B	Loco-regional control (7-years)
[4]A	Loco-regional control (5-years)
[35]A	5-year survival
[32]	Probabillity of positive surgical margins following lumpectomy
[31]	Probability of axillary pCR
[30]	Probabilty of pCR
[3]B	2 year relapse-free surival
[3]A	Probability of pathologic complete response (pCR)
[28]A	Axillary lymph node status
[2]	Probability of pathologic complete response (pCR)
[19]	Axillary Response to neoadjuvant chemotherapy in clincically node-positive patients
[17]	Pathological complete remission after preoperative chemotherapy
[16]	Preoperative diagnosis of ALN status in patients with EIBC
[15]	Level 2 lymph node metastasis
[14]	Absolute breast cancer risk
[13]	Breast cancer risk
[12]	Probability of brain metastasis
[11]	breast synchronous metastasis
[10]	ALN pCR probability
[1]	10- year proportion IBR-free

After combining the prediction of the models and the decisions moments in the Dutch care path the following overview is made [Figure 7]. The yellow numbers are the prediction models. It should be noted that the pathway provided in figure 7 is only a general overview of the different steps within the breast cancer care path. There are more decision trees behind every grey box shown in figure 7 that provide more in-depth information on that specific place. The clinical prediction models may therefore be used at different places inside the care pathway even if they were placed at the same position in the current overview. A more in depth analysis should pinpoint exactly when the prediction models can be used optimally.

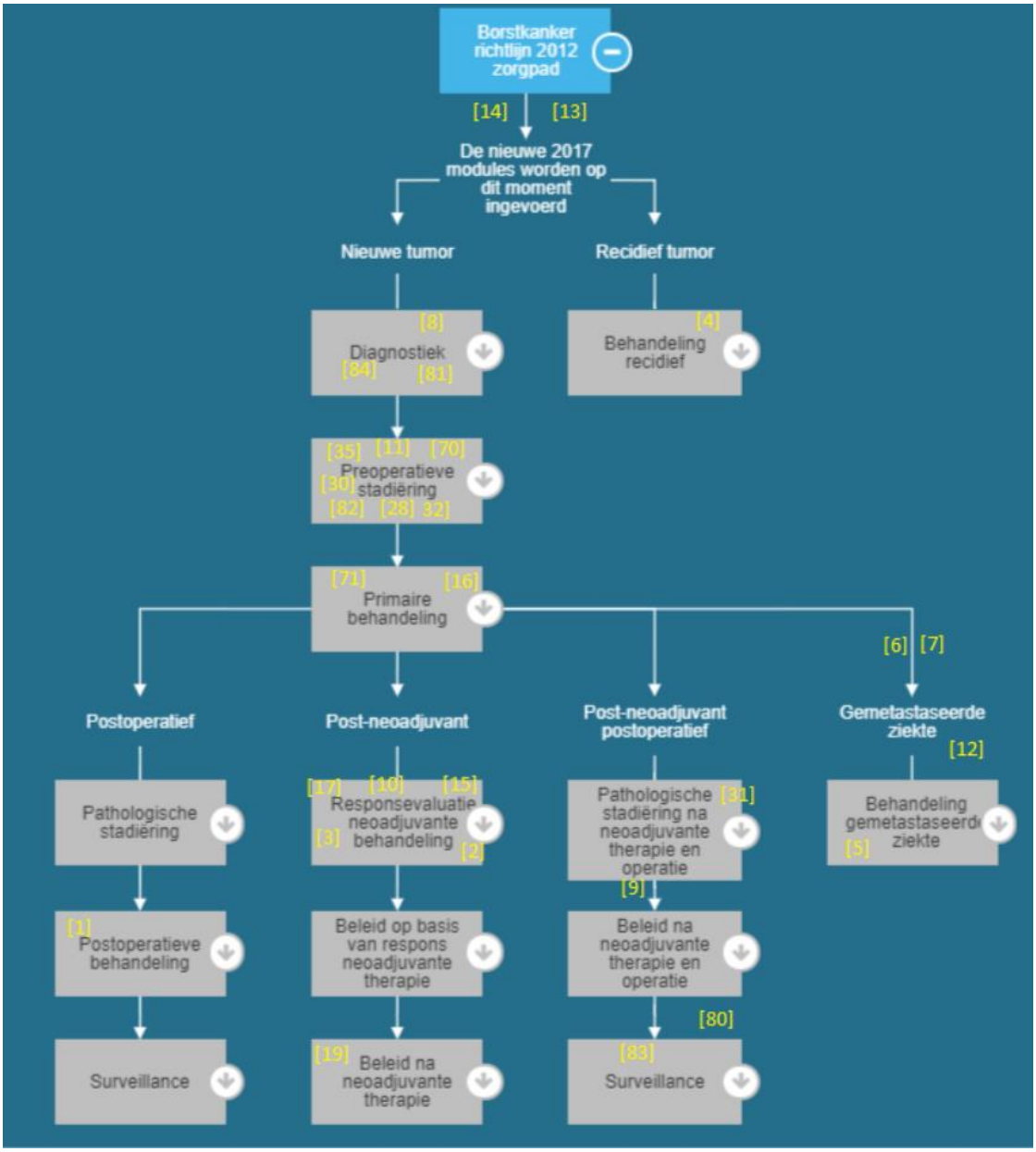

Figure 7 Global overview including the models and the corresponding decision points

Discussion

The objective of this study was to identify as many as possible breast cancer prediction models and to assess the models on transparency, reproducibility and clinical applicability. This is based on four phases. Firstly, identifying as many as relevant clinical prediction models regarding breast cancer in a literature study. Secondly, describe the clinical prediction models in a clear overview. Thirdly, evaluate the prediction models and criticize the quality of reporting of the articles. Lastly, translate as many as possible prediction models to the online platform 'Evidencio' and merge them into the existing Dutch care path.

A total of 142 prediction models regarding breast cancer where identified in 91 development studies. The overall quality of reporting was poor as 111 models did not have all the data available and 31 models where reported transparently, allowing immediate translation to an online calculator. Lastly, 38 models were translated to Evidencio and mapped to their corresponding decision moment using the digital breast cancer guideline Oncoguide.

Concerning the missing data, the first thing that is remarkable is that the baseline hazard was missing for 73 of the 76 cox proportional hazard regression models. Only 3 of the 76 models had all the data available. Six models were missing both the baseline hazard and a nomogram, this means that it was not possible to translate these models to online calculators.

Furthermore, from all the prediction models were 65 logistic regression models. It is remarkable that the intercept was missing for 37 of the 65 logistic regression models. Seven prediction models missed both the intercept and the nomogram, this means that it will not be possible to get access to all the data and upload the prediction model on Evidencio unless the auteur of the article will be contacted. However, 28 prediction models did have all the data available.

Concerning the incomplete articles per publication year, it is remarkable that the percentage of incomplete articles per publication year increased because the TRIPOD statement was published in 2015 and the primary aim of this statement is to improve the transparency and completeness of reporting of prediction modelling studies. It is to be expected that the availability of the TRIPOD statement and checklist caused a decrease in studies lacking a transparent description. However, the number of incomplete articles after 2015 is still very high.

Study limitations

It is likely that a couple of prediction models are missing in this overview, because we looked from 2010 till July 2018 and prediction models were also developed before 2010. Besides, it is also possible that not all available models between 2010 and July 2018 were identified as the literature search was not performed systematically and a limited amount of different search terms was used. Furthermore, there are prediction models that are not specifically developed for breast cancer patients but can be applied to breast cancer patients. Such as a model that predicts toxicity in patients undergoing chemotherapy. We did not contact authors in cases where the reporting was incomplete, as the main focus of this study was to create an overview of reported studies.

Because of the poor reporting of the models, it will be much more difficult to reproduce and validate the models. Hereby, it took a lot of time to translate the prediction models to online calculators on Evidencio. If the reporting of the models was more complete, there would be more models on Evidencio and also more models implemented in the Dutch care pathway.

Concerning the implementation of the prediction models in the Dutch care pathway, it is a global overview and that means that not all decision points are covered. Only the 38 models from Evidencio are merged in the pathway. When all models were implemented you will get a much better, completer and more useful overview.

Future perspectives

Overall, the quality of reporting was poor. A lot of data in the articles were missing and in 16 prediction models, the underling formula could not be derived. For future research it is recommended to contact the authors of the respective studies to obtain the underlying formula of the developed prediction models.

Before implementing the prediction models in the Dutch care path, it is necessary to follow a few steps. It is preferred to validate and compare the identified models on their performance in Dutch patients. Thereafter, the Dutch care path, all the decision trees, must be fully presented so the prediction models can be implemented at the exact decision/prediction moment. After that, the prediction models can be merged into the Dutch care path.

We strongly recommend authors, and peer-reviewers to follow the TRIPOD-statement for reporting newly developed prediction models in the future. It would facilitate a consistent manner of reporting and it guarantees the inclusion of important items needed for interpretation and reproduction of the models.

We recommend clinicians to use the prediction models on Evidencio, because they are user-friendly, validated and they are making the consequences of choices visible.

Conclusion

Generally, the quality of reporting is poor, most articles did not report all the necessary parameters to reproduce a clinical prediction model. To improve the application of prediction models for breast cancer care, the quality of reporting must be better. All models on Evidencio were assigned to the location in the guideline of which the model may support clinical decision making. Further assessment is necessary on the clinical impact and validity of the models before implementing them in the guideline.

Appendix 1: References

[1] Borstkanker in Nederland. *IKNL.* [Online] 2 March 2018. https://iknl.nl/oncologische-zorg/tumorteams/borstkanker.
[2] NHG-standaard borstkanker. *Nederlands huisartsen genootschap.* [Online] 2 March 2018. https://www.nhg.org/standaarden/volledig/nhg-standaard-borstkanker.
[3] *Difference between invasive lobular and invasive ductal carcinoma of the breast:results and therapeutic implications.* Romualdo Borroso-Sousa, Otto Metzger-Filho. 2016, Medical oncology, pp. 261–266.
[4] Treatments for Hormone-Receptor-Positive breast cancer. *Breastcancer.* [Online] 6 May 2018. http://www.breastcancer.org/symptoms/diagnosis/hormone status/treatment hrpos.
[5] Richtlijnen borstkanker. *Oncoline.* [Online] 5 March 2018. https://www.onconline.nl/index.php?pagina=/richtlijnen/item/pagina.php&id=41430&richtlijn_id=1069.
[6] Oncoguide borstkanker richtlijnen 2012. *IKNL.* [Online] 6 March 2018. https://www.oncoguide.nl/#!guidelines/1258/1258.
[7]*Stand van zaken methodologie van onderzoek: ziektekansen voorspellen: rekenen met predictieregels.* Verbeek AJ, Verbeek JF, van Dijk JA, Verbeek AL. 2014, Nederlands tijdschrift Geneeskunde, p. 158.
[8] Samenvatting. *repub.eur.* [Online] 6 march 2018. https://repub.eur.nl/pub/1169/21.
[9] Biomarkers, predictiemodellen en borstdensiteit. *RIVM.* [Online] 7 March 2018. https://www.rivm.nl/Onderwerpen/B/Bevolkingsonderzoek_borstkanker_voor_professional/Actuele_ontwikkelingen/Biomarkers_predictiemodellen_en_borstdensiteit.
[10] *Prognostic models: clinically useful or quickly forgotten?* Wyatt JC, Altman DG. 1995, BMJ.
[11] *Wildgroei van risicomodellen voor hart- en vaatziekten.* L, Maillette de Buy Wenniger. 2016, Nederlands Tijdschrift Geneeskunde
[12] *Transparent reporting of multivariable prediction model for individual prognosis or diagnois (TRIPOD): the TRIPOD statement.* Gary S Collins, Johannes B Reitsma, Douglas G altman and Karel GM Moons. 2015, BMC Medicine, pp. 1-13.
[13] *Decision curve analysis: a novel method for evaluating prediction models.* Elkin, Andrew J. Vickers and Elena B. 2006, Medical decision making, pp. 565-574.
[14] Pepe M, Janes H. Methods for evaluating prediction performance of biomarkers and Tests. [boekaut.] Gail M, Pfeiffer R, Satten G, Cai T, Gandy A Lee Ml. *Risk Assessment and Evaluation of predictions. Lecture notes in statistics.* New York : sn, 2013, p. 2015.
[15] Background. *Evidencio.* [Online] 6 March 2018. https://www.evidencio.com/about.
[16] Journal metric, what is journal h-index? *University of South Australia.* [Online] 8 June 2018. https://guides.library.unisa.edu.au/c.php?g=169983&p=1119055.
[17] *How to develop, validate, and compare clinical prediction models involving radiological parameters: Study design and statistical methods. .* Kyangwa Han, Kijung Song, Byoung Wook Choi. 2016, Korean Journal of Radiology, pp. 339-350.
[18] Oncologische zorg oncoguide. *IKNL* [Online] 1 July 2018. https://www.iknl.nl/oncologische-zorg/oncoguide

Appendix 2: Example overview prediction models

Prediction model	Country	Publication Year	Prediction	Variables	Performances (concordance)	Calculation method	Presentation	Population	Validation	Validation cohort	Evidencio	Possibillity*	Transparancy	Journal
[65]	Brazil/USA	2003	likelihood of additional, non-sentinel lymph node metastases in a patient with a positive SLN	tumor type and nuclear grade, lymphovascular invasion, multifocality of primary tumor, ER/ number of negative SLNs/ number of positive SLNs/ pathologic size in cm/ method of detection of SLN metastases	0.76	Logistic regression	Nomogram	702 Patients	Internal		No	Yes	Intercept and coefficients are missing	Surgical Oncology
[79]	Taiwan	2005	5- year Local regional recurrence of breast carcioma after mastectomy	Age/ Estrogen receptor status/ Lymph node positive/ Lymphovascular invasion/ Adjuvant radiotherapy		Cox regressie	Decision tree	1010 Patients	Internal		No	Yes	Yes	International journal of radiation oncology biology physics
[55]c Neoadjuvant	USA	2005	10-year metastases-free survival	Residual tumor size at surgery/ Number of metastatic nodes at surgery/ histologic grade/ er status/ histologic type	0.71	Cox regressie	Nomogram	337Patients	Internal		No	Yes	Baseline Hazard and coefficients are missing	Journal of Clinical Oncology
[55]b Neoadjuvant	USA	2005	5- year metastases-free survival	Residual tumor size at surgery/ Number of metastatic nodes at surgery/ histologic grade/ er status/ histologic type	0.71	Cox regressie	Nomogram	337 Patients	Internal		No	Yes	Baseline Hazard and coefficients are missing	Journal of Clinical Oncology
[55]a Neoadjuvant	USA	2005	Probability of pathologic complete response (pCR)	ER status/ T (initial, TNM)/ Histologic grade/ age/ Number of courses	0.77	Logistic regression	Nomogram	337 Patients	Internal		No	Yes	Intercept and coefficients are missing	Journal of Clinical Oncology
[73]b	USA	2006	Absolute breast cancer risk for women younger than 50 years	age at birth of firt live child/ number of affected mother or sisters/ number of previous bening breast bipsy examinations/ weight/ mammographic density	0.747	Logistic regression	Formula	1774 Patients			No	Yes	Yes	Journal of National Cancer institute
[73]a	USA	2006	Absolute breast cancer risk for women older than 50 years	age at birth of firt live child/ number of affected mother or sisters/ number of previous bening breast bipsy examinations/ weight/ mammographic density	0.779	Logistic regression	Formula	1774 Patients			No	Yes	Yes	Journal of National Cancer institute
[61]b Neoadjuvant	USA	2006	Probability of breast conservation	ER status/ initial diameter/ histologic grade/ multicentricity/ histologic type	0.71	Logistic regression	Nomogram	1147 Patients	Internal/ External	0.71 and 651 patients from Texas	No	Yes	Yes	American Cancer Society
[61]a Neoadjuvant	USA	2006	Probability of residual tumor size less than 3 cm	ER status/Initial diameter/ histologic grade/ histologic type/ number of type	0.67	Logistic regression	Nomogram	496 Patients from France	Internal/ External	0.79 and 651 patients from Texas	No	Yes	Yes	American Cancer Society
[29]b	UK	2006	Breast cancer within 1 year in postmenopausal women	Age/Hispanic/Race/BMI/Age of birth of first child/ Prior breast procedure/ First-degree family history of breast cancer/ current hormone therapy/ surgical menopause/previous mammographic outcome/Breat density (BI-RADS)	0.624	Logistic regression	Table	2392998 Patients	Internal		No	No	Intercept and Nomogram is missing	Journal of the National cancer institute

Appendix 3: Summary Excel file

Publication Country	Number of articles
Australia	1
Brazil	2
Brazil/USA	1
China	20
Egypt	1
France	1
France/USA	3
Finland	1
India	1
Italy	4
Japan	6
Korea	6
Mexico	1
Nigeria/USA	1
Singapore	1
Taiwan	1
Thailand	1
The Netherlands	10
Turkey	1
UK	4
UK/Canada/USA	1
USA	22
USA/Italy	1
Calculation method	
Cox regression method	35
Logistic regression method	55
Other	2

Publication Country	Number of prediction models
Australia	2
Brazil	4
Brazil/USA	1
China	39
Egypt	1
France	1
France/USA	8
Finland	1
India	2
Italy	4
Japan	6
Korea	8
Mexico	1
Nigeria/USA	1
Singapore	4
Taiwan	1
Thailand	1
The Netherlands	15
Turkey	1
UK	11
UK/Canada/USA	1
USA	28
USA/Italy	1
Calculation method	
Cox regression method	75
Logistic regression method	65
Other	2
Presentation model	
Nomogram	101
Scorecard	2
Table	26
Table and scorecard	2
Formula	5
Formula and table	1
Decision tree	3
Other	2

Appendix 4: Publication journals

Journal	H Index	Number of Articels	Complete	Incomplete	% incomplete/ Number of articels
Journal of Clinical Oncology	483	11	3	8	73%
Journal of National Cancer institute	326	6	3	3	50%
Clinical cancer research	285	7	0	7	100%
Plos One	241	9	1	8	89%
Internation journal of Radiation Oncology Biology Physics	221	2	1	1	50%
International Journal of Cancer	206	1	1	0	0%
British journal of Cancer	204	6	0	6	100%
European Journal of Cancer	184	2	1	1	50%
Cancer epidemiology, biomarkers & prevention	172	1	1	0	0%
Cancer letters	152	1	0	1	100%
Annals of Surgical Oncology	149	1	0	1	100%
Radiotherapy and Oncology	136	1	1	0	0%
Breast cancer Research and Treatment	133	14	6	8	57%
European radiology	131	1	0	1	100%
Medicine	130	1	0	1	100%
Human Pathology	127	1	0	1	100%
Breast cancer Research	126	4	2	2	50%
Scientific reports	122	2	1	1	50%
journal of Clinical Pathology	113	4	0	4	100%
Journal of Cellular and Molecular Medicine	108	1	0	1	100%
BMC Cancer	104	4	1	3	75%
Investigative radiology	98	1	0	1	100%
Journal of Surgical Oncology	97	3	0	3	100%
Neuro-Oncology	94	2	0	2	100%
Supportive care in cancer	92	1	0	1	100%
Journal of Cancer Research and Clinical Oncology	84	2	0	2	100%
Oncotarget	77	11	1	10	91%
American Journal of Clinical Oncology	70	1	0	1	100%
Breast	67	4	0	4	100%
Journal of experimental & Clinical Cancer Research	63	1	0	1	100%
Clinical breast cancer	61	3	1	2	67%
Asian pacific journal of cancer prevention	59	1	1	0	0%
Surgical Oncology	54	7	1	6	86%
Journal of the American Heart Association	49	1	0	1	100%
Tumori journal	48	1	1	0	0%
Breast cancer	44	1	1	0	0%
Journal of clinical laboratory analysis	42	2	0	2	100%
Translational Oncology	39	1	0	1	100%
Oncotarget and Therapy	38	1	0	1	100%

Oncology letters	32	3	0	3	100%
Indian journal of Cancer	29	2	0	2	100%
Cancer medicine	27	5	0	5	100%
Journal of Breast cancer	24	1	0	1	100%
Healthcare Informatics Research	17	1	0	1	100%
American Cancer Society	16	5	2	3	60%
JPRAS Open	3	1	0	1	100%
		142			