



Original Investigation | Oncology

Risk Prediction Models for Cardiotoxicity of Chemotherapy Among Patients With Breast Cancer

A Systematic Review

Elisé G. Kaboré, MD; Conor Macdonald, PhD; Ahmed Kaboré, PhD; Romain Didier, MD; Patrick Arveux, MD, PhD; Nicolas Meda, MD, PhD; Marie-Christine Boutron-Ruault, MD, PhD; Charles Guenancia, MD, PhD

Abstract

IMPORTANCE Cardiotoxicity is a serious adverse effect that can occur in women undergoing treatment for breast cancer. Identifying patients who will develop cardiotoxicity remains challenging.

OBJECTIVE To identify, describe, and evaluate all prognostic models developed to predict cardiotoxicity following treatment in women with breast cancer.

EVIDENCE REVIEW This systematic review searched the Medline, Embase, and Cochrane databases up to September 22, 2021, to include studies developing or validating a prediction model for cardiotoxicity in women with breast cancer. The Prediction Model Risk of Bias Assessment Tool (PROBAST) was used to assess both the risk of bias and the applicability of the prediction modeling studies. Transparency reporting was assessed with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) tool.

FINDINGS After screening 590 publications, we identified 7 prognostic model studies for this review. Six were model development studies and 1 was an external validation study. Outcomes included occurrence of cardiac dysfunction (echocardiographic parameters), heart failure, and composite clinical outcomes. Model discrimination, measured by the area under receiver operating curves or C statistic, ranged from 0.70 (95% IC, 0.62-0.77) to 0.87 (95% IC, 0.77-0.96). The most common predictors identified in final prediction models included age, baseline left ventricular ejection fraction, hypertension, and diabetes. Four of the developed models were deemed to be at high risk of bias due to analysis concerns, particularly for sample size, handling of missing data, and not presenting appropriate performance statistics. None of the included studies examined the clinical utility of the developed model. All studies met more than 80% of the items in TRIPOD checklist.

CONCLUSIONS AND RELEVANCE In this systematic review of the 6 predictive models identified, only 1 had undergone external validation. Most of the studies were assessed as being at high overall risk of bias. Application of the reporting guidelines may help future research and improve the reproducibility and applicability of prediction models for cardiotoxicity following breast cancer treatment.

JAMA Network Open. 2023;6(2):e230569. doi:10.1001/jamanetworkopen.2023.0569

Key Points

Question What is the evidence for use of clinical prediction models to predict cardiotoxicity related to chemotherapy treatment in women with breast cancer?

Findings This systematic review identifies 7 studies and 6 outcome prediction models for cardiotoxicity in different populations of patients with breast cancer. Most prognostic models report good to excellent discrimination but fall short in addressing bias due to methodological weaknesses, particularly for sample size, lack of overfitting consideration, handling of missing data, and insufficient model performance assessment.

Meaning These results suggest that there is insufficient evidence to provide personalized risk prediction of cardiotoxicity related to breast cancer treatment in clinical practice; thus, researchers should adhere to best practices in the process of developing and validating predictive models.

+ [Invited Commentary](#)

+ [Supplemental content](#)

Author affiliations and article information are listed at the end of this article.

Open Access. This is an open access article distributed under the terms of the CC-BY License.

Introduction

Breast cancer is now the world's most commonly diagnosed cancer, and it is responsible for 1 in 6 cancer deaths among women.¹ Research has yielded several developments in breast cancer diagnosis and treatment that improve morbidity and mortality outcomes. Breast cancer survival 5 years after diagnosis now exceeds 80% in most high-income countries.² Currently, there are more than 7.7 million breast cancer survivors around the world; this reflects improvements in diagnostics and treatment strategies including radiation therapy, chemotherapy, and biologic agents.^{3,4}

Unfortunately, these therapeutics can cause potentially life-threatening adverse effects. Myocardial dysfunction and heart failure (HF), frequently described as cardiotoxicity, are serious adverse effects that women undergoing breast cancer treatment can experience.⁵ In the literature, the incidence of cardiotoxicity was reported to be as high as 34% when anthracycline and targeted agents are combined.⁶ Moreover, breast cancer survivors have a significantly increased risk of death due to cardiovascular disease (CVD).⁷

Given that cardiotoxicity is a significant factor in the prognosis for patients with breast cancer, many studies have aimed to identify predictors of cardiotoxicity. Age, obesity, hypertension, diabetes, and previous anthracycline treatment have been identified as major predictors of cardiac toxicity.⁸⁻¹¹ Several scores that can assist health care professionals and patients with breast cancer in assessing prognostic indicators and other decisions made in primary care have been developed in this setting. With the increasing number of breast cancer survivors and adverse cardiovascular events, it appears important to systematically assess the available prediction models and their ability to predict cardiotoxicity. The purpose of this study was to identify, describe, and appraise all prognostic models developed to predict cardiotoxicity related to chemotherapy treatment in women with breast cancer.

Methods

A systematic review was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) reporting guideline. The study protocol was registered with the Prospero International Prospective Register of Systematic Reviews (CRD42021290441). An ethics statement was not required because this study was based exclusively on published research.

Inclusion Criteria for Study Selection

The PICOTS (population, intervention model, comparator, outcome, timing and settings) strategy was adopted for the search.¹² We included all developed prognostic models and their corresponding external validation studies (intervention model), involving women diagnosed with breast cancer at any stage (population) who subsequently had cardiotoxicity, defined as changes in LVEF or all symptomatic cardiac toxic effects after breast cancer therapy (outcome), at any time during follow-up (timing), with no predefined comparator (comparator) or setting (setting). We excluded review articles, conference abstract and studies with only patients with metastatic breast cancer. Two reviewers (E.G.K. and A.K.) independently performed the first screening by title and abstract. The 2 investigators independently undertook all study processes including online database search, study selection, data extraction, and critical appraisal. Any disagreements were discussed to reach a consensus. When a consensus was not obtained, third-party experts were invited to research, discuss, and finally reach a decision.

Search Methods for the Identification of Studies

On September 22, 2021, we searched the electronic bibliographic databases Embase, Medline (PubMed), and Cochrane with no restriction on publication date. We developed a search strategy from a combination of free text and Medical Subject Heading terms for BC, cardiotoxicity, and HF (eAppendix 1 in Supplement 1).

Data Collection and Analysis

We adapted the standardized data extraction form containing items based on the general guidelines of the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) checklist.¹³ Data were extracted for general information, source of data, participants, predictors, outcome, and analysis (eAppendix 2 in Supplement 1).

Risk of Bias and Reporting Transparency Assessment

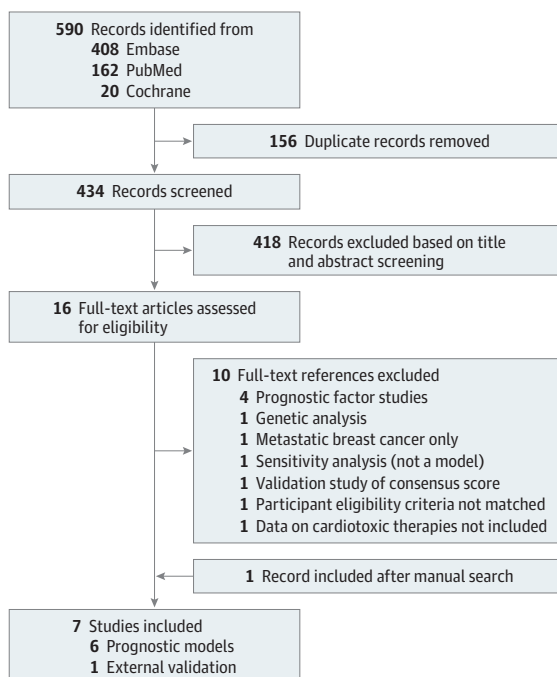
We used PROBAST (Prediction model Risk Of Bias Assessment Tool) to assess the risk of bias of the individual prognostic models investigated.¹⁴ Risk of bias (ROB) was assessed according to the 4 PROBAST domains (participants, predictors, outcome, and analysis) and applicability according to 3 domains (participants, predictors, and outcome). The ROB was assessed to be either high, low, or unclear based on signaling questions (eAppendix 3 in Supplement 1).

We used the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) reporting guideline to assess good reporting of studies developing or validating multivariable prediction models (eAppendix 4 in Supplement 1).¹⁵ In the absence of sufficient data for a meta-analysis, we have used a narrative synthesis instead.

Results

The search identified 590 potentially relevant citations related to prognostic models for cardiotoxicity in breast cancer. After removal of 156 duplicates, the titles and abstracts of the remaining 434 references were screened. We identified 16 studies for full-text assessment that would possibly fulfill our predefined inclusion criteria. Of these, we excluded 10 references (eTable 3 in Supplement 1). One study was included following a manual search.¹⁶ The PRISMA flowchart of the search strategy and the selection process is illustrated in Figure 1. Of the 7 studies included in this review, 6 focused on model development¹⁶⁻²¹ and 1²² focused on external validation (Tables 1 and 2).

Figure 1. Flowchart of Study Inclusion



Model Development Studies

Of the 6 model development studies, 3 produced scoring systems,^{17,18,20} while the other 3^{16,19,21} were logistic regression prediction models (eAppendix 4 and 5 in Supplement 1). The prognostic models were mostly developed after 2019.

Participants

Of 6 studies, 3 used retrospective data^{17,18,20} and 3 used prospective data.^{16,19,21} Three studies included patients with localized breast cancer,^{16,17,21} 1 study included patients with localized and metastatic breast cancer,²⁰ and 2 studies did not report the cancer stage.^{18,19} The numbers of participants in the 6 studies ranged from 217 to 8068 patients, and the mean age ranged between 49 and 73 years. Anthracyclines, alone or combined with trastuzumab, were used in all women in 4 studies.^{16,18,19,21}

Follow-up Duration

Follow-up ranged from 1 to 7 years. Two studies assessed the 1-year risk of cardiotoxicity.^{19,21} The other studies reported had a longer follow-up period: 3-year risk of HF or cardiomyopathy¹⁷; 3-to-10-year risk of HF¹⁸; risk of major adverse cardiovascular events (MACE) at 1, 3, and 7 years²⁰; and 5-year risk of cardiac death or congestive HF.¹⁶

Table 1. Model Development Studies With Their Predictors

Prediction model	Predictors (or risk factors) in the final model ^a	Outcome	Discrimination	Calibration
Ezaz et al, ¹⁷ 2014; development model	Age 67-74 y, age 75-79 y, age 80-94 y, anthracycline chemotherapy, nonanthracycline chemotherapy, no identified chemotherapy, coronary artery disease, atrial fibrillation/flutter, diabetes, hypertension, and kidney failure	HF or CM defined as ICD-9-CM codes in ≥1 inpatient claim or 2 outpatient claims ≥30 d apart	Not reported	HL P = .76
Fogarassy et al, ¹⁸ 2019; development model	Age 40-49 y, age 50-59 y, age 60-69 y, age ≥70 y, diabetes, hypertension, CAD without myocardial infarction or revascularization, CAD with myocardial infarction or revascularization, previous stroke, regional invasion, distant metastasis, epirubicin dose 451-540 mg/m ² , dose 541-709 mg/m ² , dose >709 mg/m ² , docetaxel dose ≤510 mg/m ² , docetaxel dose >510 mg/m ² , capecitabine, gemcitabine, bevacizumab, ACEi/ARB	HF defined as (1) hospital discharge following the diagnosis of the ICD code I50, (2) hospitalization that ended in death and an I50 code issued as a primary or secondary diagnosis or as the underlying cause of death, (3) autopsy report with the I50 code	C-index, 0.79	HL P = .29
Low risk TRC; Goel et al, ¹⁹ 2019; development model	Baseline LVEF and LVEF change ^b	Any of (1) death due to HF failure, AMI, or arrhythmia; (2) grade III or IV cardiac arrhythmia or ischemia or infarction; (3) NYHA III or IV; (4) an asymptomatic decrease in LVEF of >15%; and (5) an asymptomatic decrease in LVEF of >10% to an absolute value of <50%	C-index, 0.87 (95% IC, 0.77-0.96)	Not reported
CHEMO-RADIAT; Kim et al, ²⁰ 2021; development model	Prior congestive HF, Hypertension, age ≥60 y, prior myocardial infarction/peripheral artery occlusive disease, obesity, kidney failure, abnormal lipid profile, diabetes, radiation to left breast with ≥30 Gy dose, anthracycline dose, TIA/stroke	MACE: composite of HF, myocardial infarction, stroke, cardiovascular deaths	C-index, 0.87 (95% IC, 0.78-0.96)	HL P = .09
CRS; Romond et al, ¹⁶ 2012; development model	Age and baseline LVEF	Definite or probable cardiac death or congestive HF manifested by dyspnea with normal activity or at rest and associated with an absolute decrease in LVEF of >10% from baseline to a value <55% or a decrease of >5% to a value below the lower limit of normal	C-index, 0.72	Plot between observed and predicted probability
Upshaw et al, ²¹ 2019; development model	Age, BMI, hypertension, and baseline LVEF	Composite of reduction from baseline in LVEF ≥10% with a resultant LVEF <50% and/or a clinical diagnosis of heart failure through the first year of follow-up	C-index, 0.70 (95% IC, 0.62-0.77)	Not reported

Abbreviations: ACEi/ARB, angiotensin-converting enzyme inhibitors/angiotensin receptor blocker; AMI, acute myocardial infarction; BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); CAD, coronary artery disease; CM, cardiomyopathy; GLS, global longitudinal strain; HF, heart failure; HL, Hosmer-Lemeshow; ICD, international classification of disease; LVEF, left ventricular ejection fraction; MACE, major adverse cardiovascular events; NYHA, New York Heart Association; TIA, transient ischemic attack.

^a Risk factors for Goel et al.¹⁹

^b Baseline LVEF is LVEF preanthracycline, and LVEF change is baseline LVEF minus postanthracycline LVEF.

Outcome Definition

The definition of cardiac events was not uniform. Three of the included studies used clinical criteria^{17,18,20}; 2 of them used administrative data with *International Classification of Diseases, Ninth Revision (ICD-9)* and *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10)* codes to define HF,^{17,18} and 1 of them used MACE as the outcome, which was a composite of HF, myocardial infarction, stroke, cardiovascular deaths.²⁰ In 3 studies, clinical criteria were used in conjunction with left ventricular parameters.^{16,19,21}

Method, Analysis, and Presentation

The included models used between 2 and 20 predictors (Table 1). Most often used predictors were age^{16-18,20,21} and comorbidities. Comorbidities included: diabetes^{17,18,20}; hypertension^{17,18,20,21}; body

Table 2. Study Characteristics of Model Development Studies

Source	Source of data	Participant characteristics				Outcomes			Validation	
		No. of persons	Age, y	Stage	Treatment	Participants with cardiotoxicity, No. (%)	Follow-up	Presentation	Development stage	External
Ezaz et al, ¹⁷ 2014; development model	Retrospective cohort, US, registry data, January 1, 2000-December 31, 2009	1664	73.6 (5.3)	Stage 1: 26.8%; stage 2: 45.6%; stage 3: 27.6%	ANTH chemotherapy: 35.9%; non-ANTH chemotherapy: 47.7%; No identified chemotherapy: 16.4%	318 (19.1%)	3 y	Scoring system	Yes	Yes
Fogarassy et al, ¹⁸ 2019; development model	Retrospective study, Hungary, registry data, January 1, 2004-December 31, 2016	8068	Not reported	Not reported (all stages)	100% Epirubicin-treated breast cancer; 20% of women with targeted therapies	557 (6.9%)	Median 5.89 (3-10) y	Scoring system	Yes	No
Low risk TRC; Goel et al, ¹⁹ 2019; development model	Prospective cohort study, Australia, multicenter, recruitment period not reported	217	52 (28-77)	Not reported (early stage, ERBB2 positive)	Conventional adjuvant ANTH-based chemotherapy followed by taxane chemotherapy given with TRZ, followed by TRZ alone to complete a total of 52 weeks	18	52 wks	Logistic regression	No	No
CHEMO-RADIAT; Kim et al, ²⁰ 2021; development model	Retrospective cohort, South Korea, multicenter, November 2005-September 2015	1256	51.4 (10.7)	Stage 0: 1%; Stage 1: 43.1%; stage 2: 38.5%; stage 3: 16.8%; Stage 4: 0.7%	Use of ANTH:70%; TRZ: 11.9%	21	48.7 mos (range, 25.8-71.8 mos)	Scoring system	Yes	No
CRS; Romond et al, ¹⁶ 2012; development model	Prospective cohort, US, multicenter, February 21, 2000-April 29, 2005	1830	49.0	Not reported (histologically node-positive, HER-2 positive)	Arm 1: doxorubicin and cyclophosphamide for four cycles followed by paclitaxel for four cycles. Arm 2: same chemotherapy plus TRZ starting with the first dose of paclitaxel at a loading dose of 4 mg/kg followed by 2 mg/kg for 51 weeks.	37 of 944 patients vs 10 of 743 patients	87 mos	Logistic regression	Yes	No
Upshaw et al, ²¹ 2019; development model	Prospective cohort, US, multicenter, November 2007-February 2011	967	51.5 (9.6)	Stage 1: 8%; stage 2: 36%; stage 3: 56%	Doxorubicin (cumulative dose, 240 mg/m ²) and cyclophosphamide for 4 cycles (classical [every 3 weeks] or dose-dense [every 2 weeks] followed by 12 doses of weekly paclitaxel	51	47.6 mos	Logistic regression	Yes	No

Abbreviations: ANTH, anthracycline; TRZ, trastuzumab.

mass index^{20,21}; atrial fibrillation¹⁷; prior coronary artery disease^{17,18,20}; kidney failure^{17,20}; and prior stroke.^{18,20} Some studies explored other predictors such as: disease characteristics (regional or distant invasion)¹⁸; therapeutic methods used such as dose of chemotherapy agents (specifically anthracycline doses)^{18,20} or dose of irradiation to left breast²⁰; biochemical predictors such as lipid profile²⁰; and echocardiographic parameters such as baseline LVEF^{16,19,21} and LVEF change.¹⁹ The models and their calculations are described in eTables 4 and 5 in Supplement 1.

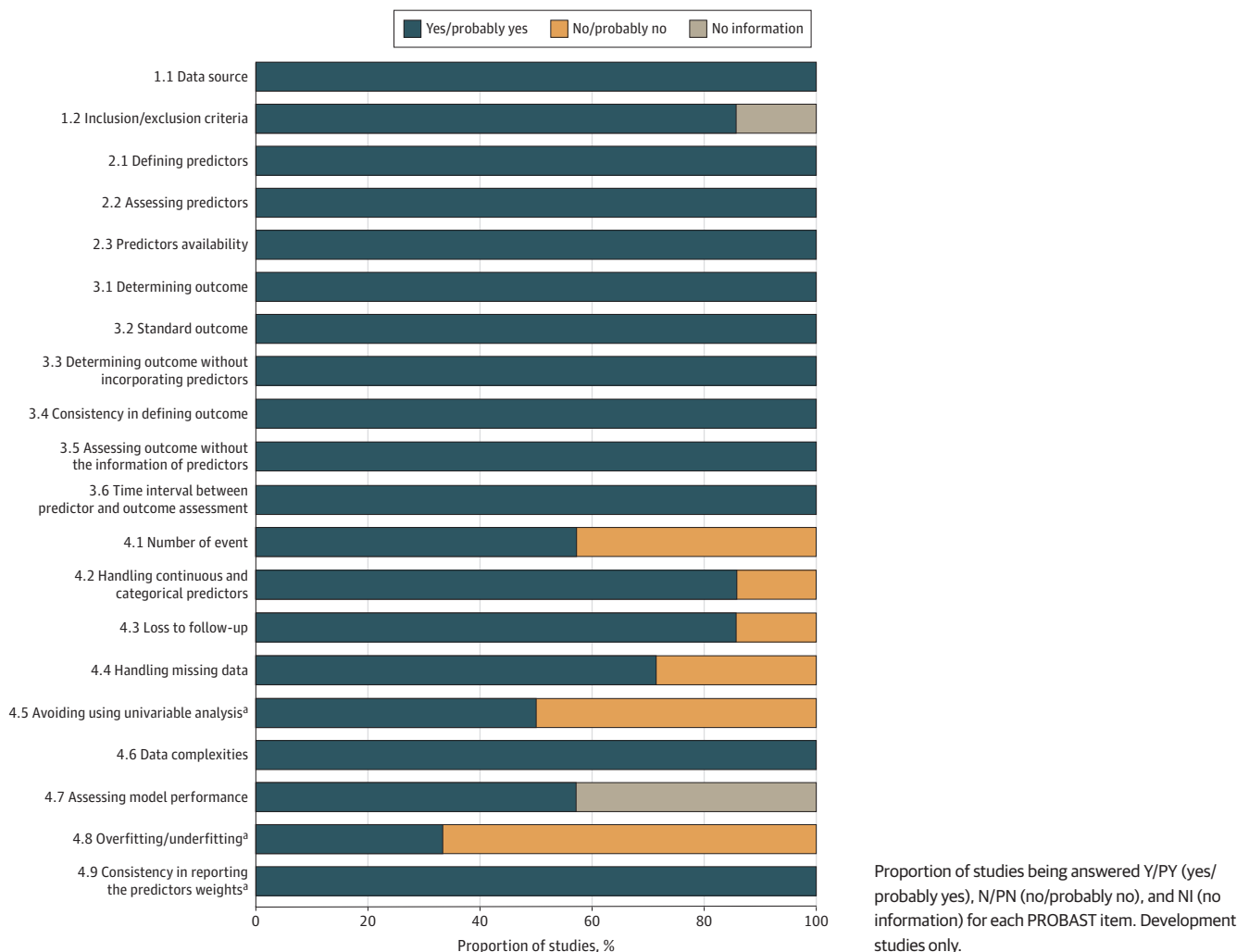
Performance

Calibration was reported in 4 studies.^{16-18,20} One study did not report discrimination.¹⁷ Of the 5 prognostic models for which the C index was reported, 2 models^{19,20} had excellent discrimination (0.80-0.89) and 3^{16,18,21} had acceptable discrimination (0.70-0.79), irrespective of cardiac event definition (Table 1).

Risk of Bias Assessment

In the PROBAST quality analysis, 1 study was rated as low ROB,¹⁸ 1 other as unclear ROB,²¹ and the remaining model development studies had high ROB due to concerns in the analysis domain (Figure 2; eFigure 1 in Supplement 1). Lack of overfitting consideration (item 4.8: internal validation techniques consists of random split-sample of participants^{17,18,20} or no internal validation has been

Figure 2. PROBAST Quality Analysis Results of 7 Studies



performed¹⁹), reasonable number of participants with the outcome (item 4.1: fewer than 10 events per variable^{19,20}), and participants lost to follow-up or with missing data handled inappropriately (item 4.3 and 4.4: participants with missing data^{16,19} or loss to follow up¹⁹ were excluded from analysis) were identified as bias related to analysis. We rated as low ROB the analysis domain for the prediction model of Fogarassy et al¹⁸ because the development study was based on a large data set and a high number of events per variable (EPV) stratum. In Upshaw and al,²¹ the development study was rated as unclear ROB due to insufficient model performance assessment (item 4.7: calibration not reported). We rated concern for applicability of all the development studies as low across all domains (eTable 6 in Supplement 1).

Transparent Reporting Assessment

TRIPOD scores varied from 25 to 28 (of 31) in the 6 development studies (Figure 3; eFigure 2 in Supplement 1). All studies met more than 80% of the items in the TRIPOD checklist. Two studies^{16,17} were published before the TRIPOD's publication in 2015. The items that were most often reported included the abstract (item 2), introduction (item 3), and source of data (item 4) for all studies. Lack of reporting on participant selection (item 13a) or models' predictive performance (item 10d) were noted in 4 studies.^{17,18,20,21} Only 1 study¹⁹ reported unadjusted association between each candidate predictor and outcome (item 14b). No study reported masked assessment of outcome (item 6b) and predictors (item 7b). Missing data handling (item 9) was not reported in 1 study.¹⁸

Model Validation Study

We identified 1 study (Milks et al²²) externally validating the prediction model by Ezaz et al.¹⁷ It was derived from data from a retrospective, monocentric cohort study. The recruitment period for the study was from 2010 to 2014, and follow-up was not reported. Patients with all stage breast cancer treated by doxorubicin alone or in association with trastuzumab were included. A total of 183 patients were included with 33 events. The reported discrimination was excellent (AUC = 0.86). Calibration was not reported (item 4.7 rated "no information"). We rated the ROB as high due to analysis concerns, including item 4.1 (the number of participants with the outcome was under 100) and item 4.2 (categorical variables like age were categorized using cut-points that were different from those of the development study), and a lack of specified exclusion criteria (Figure 2). The TRIPOD score was 21 (total 31). Insufficient reporting of the title (item 1), source of data (item 4a), participants (item 5c, and 13), missing data handling (item 9), and model performance were identified.

In their study, Milks et al²² also suggest that superior prediction assessment for cancer-therapy-related cardiac dysfunction can be achieved by combining left ventricular global longitudinal strain (LV-GLS) and the clinical risk score from Ezaz et al.¹⁷ The combined model had superior receiver-operating characteristics (C = 0.9629).

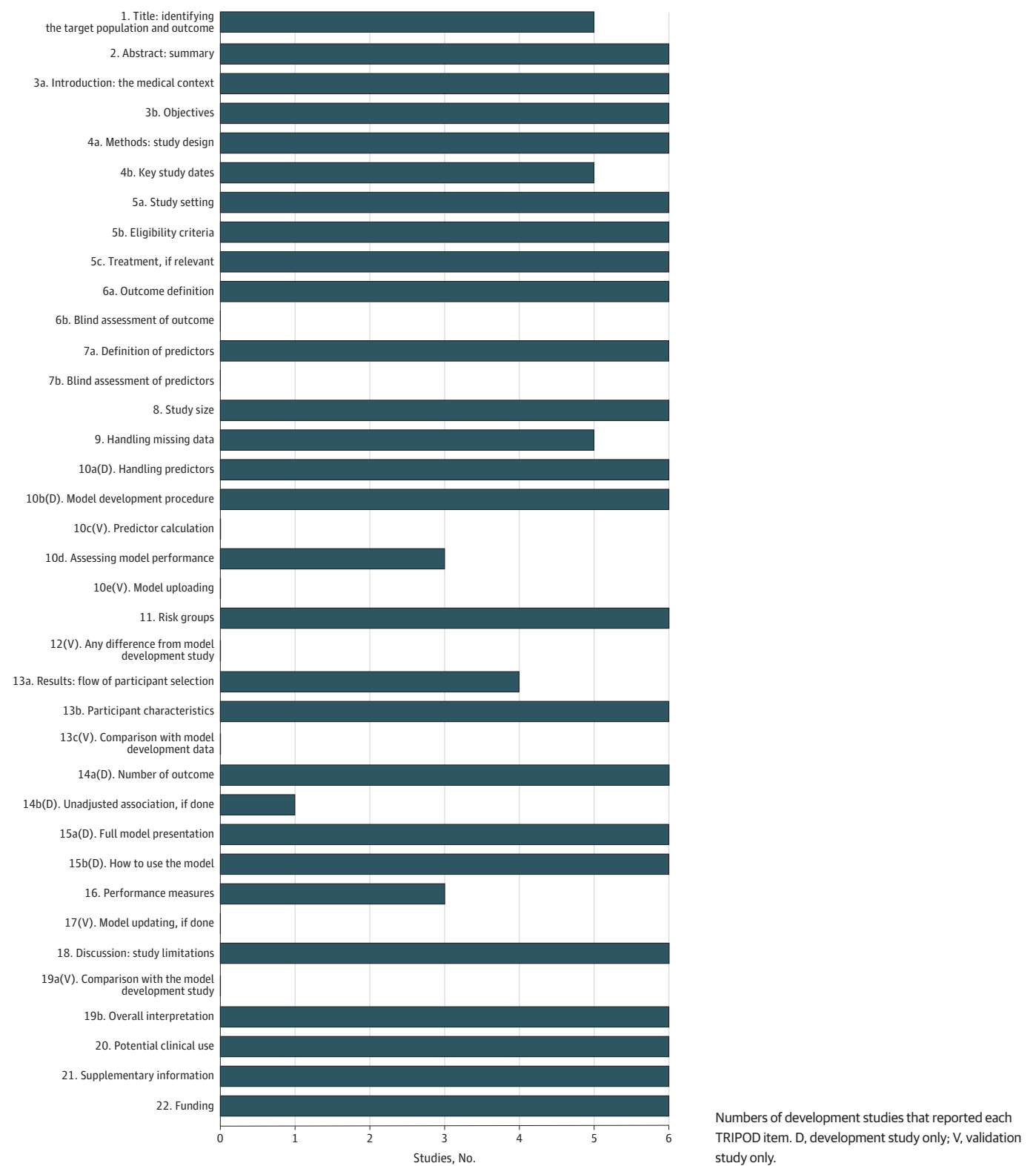
Presentation of Usability of the Models

Fogarassy et al¹⁸ was the only included study that had a low overall risk of bias. However, neither study was externally validated. None of the included studies explored net benefit analysis with respect to the developed models.

Discussion

We conducted a systematic literature search that identified 7 studies and six outcome prediction models for cardiotoxicity in different breast cancer populations, including patients with breast cancer at any stage, *ERBB2* (formerly *HER2*)-negative patients, *ERBB2*-positive patients and patients eligible for trastuzumab. One of the models¹⁷ underwent external validation in a separate external validation study.²² The performance of these models and their applicability were assessed. Although most prognostic models reported good to excellent discrimination, of the 6 models, only 1 had low ROB and 1

Figure 3. TRIPOD Checklist



other unclear ROB; the other 4 models had high ROB due to statistical analysis concerns, particularly for sample size, handling of missing data and not presenting appropriate performance statistics.

Suggestions for Future Research

Two studies included in this review used registry data to develop their model. Registry data are commonly used in clinical prediction models because they provide large sample sizes and good representativeness while at the same time being relatively inexpensive and easy to obtain.²³ Ezaz et al¹⁷ used the SEER-Medicare registry to identify 1664 American women diagnosed with early stage breast cancer, and Fogarassy et al¹⁸ used the Hungarian National Cancer Registry to include 8068 patients. Nevertheless, we should be cautious because the amount of detail in data collection and whether outcome assessment was conducted by protocol may be a limitation for prognostic analyses in patient registries.²⁴

Two models included patients with advanced cancer stage (ie, metastasis),^{18,20} whereas other models included early breast cancer stage. As expected, in the study by Fogarassy et al,¹⁸ advanced cancer with distant metastases was confirmed as a significant predictor associated with heart failure. This likely reflects the choice of chemotherapy in addition to its potential impact on cardiotoxicity progression. Pinder et al²⁵ published a large analysis of the Medicare database for patients with breast cancer older than 65 years, and identified advanced cancer stage as significant predictor for heart failure. In addition, Cho et al²⁶ found that metastasis increased doxorubicin-induced cardiotoxicity by 2.66-fold and concomitant trastuzumab increased it by 4.08-fold. In the metastatic setting, *ERBB2*-targeted therapies are used until disease progression or toxicity.²⁷ These findings suggest that cancer stage may be relatively important predictor in cardiotoxicity prediction models.

There was significant heterogeneity in the definition of cardiac events and follow-up. One study used a MACE outcome to develop a model to estimate the risk of MACE in women with breast cancer. Kim et al¹⁸ utilized a 4-point MACE outcome. Although MACE is an increasingly common and standardized primary outcome of interest in randomized controlled trials, there are potential limitations in studies using administrative databases. A systematic review reported substantial heterogeneity for the MACE composite endpoints used in studies based on administrative databases.²⁸ "Acute myocardial infarction and stroke" and "acute myocardial infarction, stroke, and all-cause death" were the 2 most common composite MACE definitions. This diversity made it challenging to compare findings across studies or to aggregate multiple study results for meta-analyses or systematic reviews when considering different treatment or research questions.²⁸

The diagnosis of cardiotoxicity is currently defined as a greater than 10% reduction in the ejection fraction (EF) and an absolute value of less than 53% according to the American Society of Echocardiography (ASE) and the European Society of Cardiovascular Imaging (EACVI),²⁹ while the cutoff is 50% according to the European Society of Cardiology.⁵ In the past, the definition of cardiotoxicity related to chemotherapy was restricted to ventricular dysfunction. However, the introduction of new treatments in the last few years has widened the concept of cardiotoxicity to a broader definition of cardiovascular toxicity, which incorporates arterial hypertension, ischemia, cardiomyopathy, myocarditis, arrhythmic complications, long QT, and arterial and venous thrombosis.³⁰ Future studies should therefore include a uniform and established definition of cardiotoxicity and a longer follow-up period to determine the effects of cancer treatments on long-term cardiotoxicity risk.

Most predictors included in the 6 prognostic models of this review were clinical variables. Incorporating a greater number of variables such as biomarkers (cardiac troponin and natriuretic peptides) or genetic information (microRNAs, proteomics and metabolomics) is another potential way to improve the prediction tool.³¹ Moreover, because in different patients, the same biomarker level could have a different meaning depending on whether it is stable or rising, repeated measurements of multiple biomarkers reflecting different pathophysiological pathways may convey incremental prognostic value to single or even repeated biomarker measurements of established biomarkers.³² The cumulative dose of anthracyclines has been proven to be an important factor

associated with cardiotoxicity in patients receiving anthracycline-containing therapy.^{33,34} In our study, models that included both the cumulative dose of anthracyclines and several comorbidities showed a better predictive performance than models that did not include these variables, except for the study by Goel et al,¹⁹ that had excellent discrimination using only baseline and change in LVEF. This suggests that the cumulative dose of anthracyclines could play a large role in the risk of cardiotoxicity in patients with breast cancer.

Model overfitting and the resulting optimism are important concerns in prediction models. We can define overfitting specific to prediction models as fitting a statistical model with too many degrees of freedom in the modeling process.²⁴ Although 5 studies conducted internal validations, only 2 studies^{16,21} applied internal validation properly by using bootstrapping techniques to overcome overfitting. In addition, although the demonstration of the performance of a model in an independent population is a necessary step before recommending its widespread use, only one of these models was externally validated.³⁵ Inappropriate EPV for sample size can cause high risks of overfitting and biased predictions.³⁶ EPV above 20 is recommended as the minimum sample size for model development³⁷ and validation samples should contain at least 100 individuals who develop the outcome of interest.³⁸ In this review, the sample size of 3 development models (EPV below 20) and the external model validation (less than 100 events) was not appropriate.

Applicability of Findings to Clinical Practice

While the key messages were consistent, 4 included studies were at high overall ROB. One predictive model was only validated once, and we did not consider there to be enough information to assess its suitability in different situations. This review reveals that there is insufficient evidence to provide personalized risk prediction of cardiotoxicity related to breast cancer treatment in clinical practice.

Strengths and Limitations

To the best of our knowledge, this is the first systematic review to focus on prediction models of cardiotoxicity in patients with breast cancer, including a thorough evaluation of the ROB and reporting quality. A main strength of this study is the adoption of the PROBAST and TRIPOD guidelines as benchmarks.

However, we acknowledge that there were several limitations to this systematic review. We only included studies published in English and did not search gray literature. However, it is unlikely that any missed studies would have significantly changed our main findings. Another weakness was that we were unable to perform our planned meta-analysis due to the lack of eligible studies.

Conclusions

This review identified 6 prognostic models developed to predict risk of cardiotoxicity following treatment in women with breast cancer. The prediction models were designed in diverse clinical settings and populations, and with a range of included predictors. We are not yet able to accurately predict cardiotoxicity outcomes for a given patient with breast cancer based on their clinical parameters. Most of the models identified in this review fell short in addressing bias due to methodological weaknesses. To improve the overall quality in future predictive models, researchers should adhere to best practices in the development process. Investigators should also be aware of the potential benefits and weaknesses of various methods for model creation and internal validation. More external validation studies are needed to assess the performance and clinical applicability of existing models before they can successfully be used in clinical practice.

ARTICLE INFORMATION

Accepted for Publication: December 20, 2022.

Published: February 23, 2023. doi:10.1001/jamanetworkopen.2023.0569

Open Access: This is an open access article distributed under the terms of the [CC-BY License](#). © 2023 Kaboré EG et al. *JAMA Network Open*.

Corresponding Author: Elisé G. Kaboré, MD, Health across Generations Team, Inserm U1018, Centre for Research in Epidemiology and Population Health (CESP), 114 rue Edouard Vaillant, Villejuif 94805, France (kabelisee@yahoo.fr; elise.kabore@gustaveroussy.fr).

Author Affiliations: Health across Generations Team, Inserm U1018, Centre for Research in Epidemiology and Population Health, Villejuif, France (E. G. Kaboré, Macdonald, Boutron-Ruault); Université Joseph Ki-Zerbo, Ouagadougou, Burkina Faso (A. Kaboré, Meda); Department of Cardiology, CHU Dijon-Bourgogne, Dijon, France (Didier, Guenancia); Center for Primary Care and Public Health, Unisanté, University of Lausanne, Lausanne, Switzerland (Arveux).

Author Contributions: Drs E. Kaboré and Guenancia had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: E. Kaboré, Macdonald, A. Kaboré, Arveux, Meda, Guenancia.

Acquisition, analysis, or interpretation of data: E. Kabore, A. Kaboré, Didier, Boutron-Ruault.

Drafting of the manuscript: E. Kaboré, A. Kaboré.

Critical revision of the manuscript for important intellectual content: Macdonald, A. Kaboré, Didier, Arveux, Meda, Boutron-Ruault, Guenancia.

Statistical analysis: Kabore.

Administrative, technical, or material support: E. Kaboré, A. Kaboré, Meda.

Supervision: Macdonald, Arveux, Meda, Boutron-Ruault, Guenancia.

Conflict of Interest Disclosures: Dr Boutron-Ruault reported receiving personal fees from Mayoli-Spindler, Gilead, and ViiV for presentations outside the submitted work. No other disclosures were reported.

Data Sharing Statement: See [Supplement 2](#).

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209-249. doi:10.3322/caac.21660
2. DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. *CA Cancer J Clin*. 2019;69(6):438-451. doi:10.3322/caac.21583
3. Hovaldt HB, Suppli NP, Olsen MH, et al. Who are the cancer survivors? a nationwide study in Denmark, 1943-2010. *Br J Cancer*. 2015;112(9):1549-1553. doi:10.1038/bjc.2015.68
4. Rubin G, Berendsen A, Crawford SM, et al. The expanding role of primary care in cancer control. *Lancet Oncol*. 2015;16(12):1231-1272. doi:10.1016/S1470-2045(15)00205-3
5. Zamorano JL, Lancellotti P, Rodriguez Muñoz D, et al; ESC Scientific Document Group. 2016 ESC position paper on cancer treatments and cardiovascular toxicity developed under the auspices of the ESC Committee for Practice Guidelines: the task force for cancer treatments and cardiovascular toxicity of the European Society of Cardiology (ESC). *Eur Heart J*. 2016;37(36):2768-2801. doi:10.1093/eurheartj/ehw211
6. Guglin M, Hartlage G, Reynolds C, Chen R, Patel V. Trastuzumab-induced cardiomyopathy: not as benign as it looks? a retrospective study. *J Card Fail*. 2009;15(8):651-657. doi:10.1016/j.cardfail.2009.04.011
7. Bodai BI, Tuso P. Breast cancer survivorship: a comprehensive review of long-term medical issues and lifestyle recommendations. *Perm J*. 2015;19(2):48-79. doi:10.7812/TPP/14-241
8. Jawa Z, Perez RM, Garlie L, et al. Risk factors of trastuzumab-induced cardiotoxicity in breast cancer: a meta-analysis. *Medicine (Baltimore)*. 2016;95(44):e5195. doi:10.1097/MD.0000000000005195
9. Lyon AR, Dent S, Stanway S, et al. Baseline cardiovascular risk assessment in cancer patients scheduled to receive cardiotoxic cancer therapies: a position statement and new risk assessment tools from the Cardio-Oncology Study Group of the Heart Failure Association of the European Society of Cardiology in collaboration with the International Cardio-Oncology Society. *Eur J Heart Fail*. 2020;22(11):1945-1960. doi:10.1002/ehf.1920
10. Kaboré EG, Guenancia C, Vaz-Luis I, et al. Association of body mass index and cardiotoxicity related to anthracyclines and trastuzumab in early breast cancer: French CANTO cohort study. *PLoS Med*. 2019;16(12):e1002989. doi:10.1371/journal.pmed.1002989

11. Guenancia C, Lefebvre A, Cardinale D, et al. Obesity as a risk factor for anthracyclines and trastuzumab cardiotoxicity in breast cancer: a systematic review and meta-analysis. *J Clin Oncol*. 2016;34(26):3157-3165. doi:10.1200/JCO.2016.67.4846
12. Debray TP, Damen JA, Snell KI, et al. A guide to systematic review and meta-analysis of prediction model performance. *BMJ*. 2017;356:i6460. doi:10.1136/bmj.i6460
13. Moons KG, de Groot JA, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med*. 2014;11(10):e1001744. doi:10.1371/journal.pmed.1001744
14. Wolff RF, Moons KGM, Riley RD, et al; PROBAST Group. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med*. 2019;170(1):51-58. doi:10.7326/M18-1376
15. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015;350:g7594. doi:10.1136/bmj.g7594
16. Romond EH, Jeong JH, Rastogi P, et al. Seven-year follow-up assessment of cardiac function in NSABP B-31, a randomized trial comparing doxorubicin and cyclophosphamide followed by paclitaxel (ACP) with ACP plus trastuzumab as adjuvant therapy for patients with node-positive, human epidermal growth factor receptor 2-positive breast cancer. *J Clin Oncol*. 2012;30(31):3792-3799. doi:10.1200/JCO.2011.40.0010
17. Ezaz G, Long JB, Gross CP, Chen J. Risk prediction model for heart failure and cardiomyopathy after adjuvant trastuzumab therapy for breast cancer. *J Am Heart Assoc*. 2014;3(1):e000472. doi:10.1161/JAHA.113.000472
18. Fogarassy G, Vathy-Fogarassy Á, Kenessey I, Kásler M, Forster T. Risk prediction model for long-term heart failure incidence after epirubicin chemotherapy for breast cancer—a real-world data-based, nationwide classification analysis. *Int J Cardiol*. 2019;285:47-52. doi:10.1016/j.ijcard.2019.03.013
19. Goel S, Liu J, Guo H, et al. Decline in left ventricular ejection fraction following anthracyclines predicts trastuzumab cardiotoxicity. *JACC Heart Fail*. 2019;7(9):795-804. doi:10.1016/j.jchf.2019.04.014
20. Kim DY, Park MS, Youn JC, et al. Development and validation of a risk score model for predicting the cardiovascular outcomes after breast cancer therapy: the CHEMO-RADIAT score. *J Am Heart Assoc*. 2021;10(16):e021931. doi:10.1161/JAHA.121.021931
21. Upshaw JN, Ruthazer R, Miller KD, et al. Personalized decision making in early stage breast cancer: applying clinical prediction models for anthracycline cardiotoxicity and breast cancer mortality demonstrates substantial heterogeneity of benefit-harm trade-off. *Clin Breast Cancer*. 2019;19(4):259-267.e1. doi:10.1016/j.clbc.2019.04.012
22. Milks MW, Velez MR, Mehta N, et al. Usefulness of integrating heart failure risk factors into impairment of global longitudinal strain to predict anthracycline-related cardiac dysfunction. *Am J Cardiol*. 2018;121(7):867-873. doi:10.1016/j.amjcard.2017.12.022
23. Chen L. Overview of clinical prediction models. *Ann Transl Med*. 2020;8(4):71. doi:10.21037/atm.2019.11.121
24. Steyerberg EW. *Clinical prediction models: A Practical Approach to Development, Validation, and Updating*. Springer; 2009.
25. Pinder MC, Duan Z, Goodwin JS, Hortobagyi GN, Giordano SH. Congestive heart failure in older women treated with adjuvant anthracycline chemotherapy for breast cancer. *J Clin Oncol*. 2007;25(25):3808-3815. doi:10.1200/JCO.2006.10.4976
26. Cho H, Lee S, Sim SH, et al. Cumulative incidence of chemotherapy-induced cardiotoxicity during a 2-year follow-up period in breast cancer patients. *Breast Cancer Res Treat*. 2020;182(2):333-343. doi:10.1007/s10549-020-05703-5
27. Giordano SH, Temin S, Kirshner JJ, et al; American Society of Clinical Oncology. Systemic therapy for patients with advanced human epidermal growth factor receptor 2-positive breast cancer: American Society of Clinical Oncology clinical practice guideline. *J Clin Oncol*. 2014;32(19):2078-2099. doi:10.1200/JCO.2013.54.0948
28. Bosco E, Hsueh L, McConeghy KW, Gravenstein S, Saade E. Major adverse cardiovascular event definitions used in observational analysis of administrative databases: a systematic review. *BMC Med Res Methodol*. 2021;21(1):241. doi:10.1186/s12874-021-01440-5
29. Plana JC, Galderisi M, Barac A, et al. Expert consensus for multimodality imaging evaluation of adult patients during and after cancer therapy: a report from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging*. 2014;15(10):1063-1093. doi:10.1093/ehjci/jeu192
30. Bisceglia I, Cartoni D, Petrolati S. Concepts in cardiac oncology. *Eur Heart J Suppl*. 2020;22(suppl L):L19-L23. doi:10.1093/eurheartj/suaa127

31. Ananthan K, Lyon AR. The role of biomarkers in cardio-oncology. *J Cardiovasc Transl Res*. 2020;13(3):431-450. doi:10.1007/s12265-020-10042-3
32. Klimczak-Tomaniak D, de Bakker M, Bouwens E, et al. Dynamic personalized risk prediction in chronic heart failure patients: a longitudinal, clinical investigation of 92 biomarkers (Bio-SHIFT study). *Sci Rep*. 2022;12(1):2795. doi:10.1038/s41598-022-06698-3
33. Lotrionte M, Biondi-Zoccai G, Abbate A, et al. Review and meta-analysis of incidence and clinical predictors of anthracycline cardiotoxicity. *Am J Cardiol*. 2013;112(12):1980-1984. doi:10.1016/j.amjcard.2013.08.026
34. Cardinale D, Colombo A, Bacchiani G, et al. Early detection of anthracycline cardiotoxicity and improvement with heart failure therapy. *Circulation*. 2015;131(22):1981-1988. doi:10.1161/CIRCULATIONAHA.114.013777
35. Moons KG, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart*. 2012;98(9):691-698. doi:10.1136/heartjnl-2011-301247
36. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ*. 2020;368:m441. doi:10.1136/bmj.m441
37. Ogundimu EO, Altman DG, Collins GS. Adequate sample size for developing prediction models is not simply related to events per variable. *J Clin Epidemiol*. 2016;76:175-182. doi:10.1016/j.jclinepi.2016.02.031
38. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol*. 2005;58(5):475-483. doi:10.1016/j.jclinepi.2004.06.017

SUPPLEMENT 1.

eTable 1. PRISMA Statement, Information Flow Through the Systematic Search Process

eTable 2. PRISMA Statement for Abstract Checklist

eAppendix 1. Search Strategy Conducted on September 22, 2021

eAppendix 2. Data Collection Questionnaire

eAppendix 3. PROBAST Questionnaire

eAppendix 4. TRIPOD Checklist

eTable 3. Papers Excluded on Full-text

eTable 4. Overview of Calculation per Included Model

eTable 5. Regression Models

eTable 6. PROBAST Quality Analysis Results

eFigure 1. PROBAST Quality Analysis Results of 07 Studies (06 Model Development Studies and 01 Model Validation Studies)

eFigure 2. TRIPOD Checklist

SUPPLEMENT 2.

Data Sharing Statement