Original Paper

Development of a Prognostic App (iCanPredict) to Predict Survival for Chinese Women With Breast Cancer: Retrospective Study

Zhuo Ma[1*], BSc; Sijia Huang[1*], BSc; Xiaoqing Wu[2*], BSc; Yinying Huang[3*], MSc; Sally Wai-Chi Chan[4], PhD; Yilan Lin[2], MSc; Xujuan Zheng[5], PhD; Jiemin Zhu[1], PhD

[1]Department of Nursing, School of Medicine, Xiamen University, Xiamen, China

[2]Department of Chronic Non-infectious Diseases and Endemic Diseases Control, Xiamen Center for Disease Control and Prevention, Xiamen, China

[3]Department of Nursing, Women and Children's Hospital, School of Medicine, Xiamen University, Xiamen, China

[4]President Office, Tung Wah College, Hong Kong, China

[5]School of Nursing, Health Science Centre, Shenzhen University, Shenzhen, China

[*]these authors contributed equally

Corresponding Author:
Jiemin Zhu, PhD
Department of Nursing
School of Medicine
Xiamen University
Room 222, Alice Lee Building, Xiangan Campus, Xiamen University
Xiang An South Road
Xiamen, 361005
China
Phone: 86 15960212649
Email: jieminzhu@xmu.edu.cn

Abstract

Background: Accurate prediction of survival is crucial for both physicians and women with breast cancer to enable clinical decision making on appropriate treatments. The currently available survival prediction tools were developed based on demographic and clinical data obtained from specific populations and may underestimate or overestimate the survival of women with breast cancer in China.

Objective: This study aims to develop and validate a prognostic app to predict the overall survival of women with breast cancer in China.

Methods: Nine-year (January 2009-December 2017) clinical data of women with breast cancer who received surgery and adjuvant therapy from 2 hospitals in Xiamen were collected and matched against the death data from the Xiamen Center of Disease Control and Prevention. All samples were randomly divided (7:3 ratio) into a training set for model construction and a test set for model external validation. Multivariable Cox regression analysis was used to construct a survival prediction model. The model performance was evaluated by receiver operating characteristic (ROC) curve and Brier score. Finally, by running the survival prediction model in the app background thread, the prognostic app, called iCanPredict, was developed for women with breast cancer in China.

Results: A total of 1592 samples were included for data analysis. The training set comprised 1114 individuals and the test set comprised 478 individuals. Age at diagnosis, clinical stage, molecular classification, operative type, axillary lymph node dissection, chemotherapy, and endocrine therapy were incorporated into the model, where age at diagnosis (hazard ratio [HR] 1.031, 95% CI 1.011-1.051; P=.002), clinical stage (HR 3.044, 95% CI 2.347-3.928; P<.001), and endocrine therapy (HR 0.592, 95% CI 0.384-0.914; P=.02) significantly influenced the survival of women with breast cancer. The operative type (P=.81) and the other 4 variables (molecular classification [P=.91], breast reconstruction [P=.36], axillary lymph node dissection [P=.32], and chemotherapy [P=.84]) were not significant. The ROC curve of the training set showed that the model exhibited good discrimination for predicting 1- (area under the curve [AUC] 0.802, 95% CI 0.713-0.892), 5- (AUC 0.813, 95% CI 0.760-0.865), and 10-year (AUC 0.740, 95% CI 0.672-0.808) overall survival. The Brier scores at 1, 5, and 10 years after diagnosis were 0.005, 0.055, and 0.103 in the training set, respectively, and were less than 0.25, indicating good predictive ability. The test set externally validated

model discrimination and calibration. In the iCanPredict app, when physicians or women input women's clinical information and their choice of surgery and adjuvant therapy, the corresponding 10-year survival prediction will be presented.

Conclusions: This survival prediction model provided good model discrimination and calibration. iCanPredict is the first tool of its kind in China to provide survival predictions to women with breast cancer. iCanPredict will increase women's awareness of the similar survival rate of different surgeries and the importance of adherence to endocrine therapy, ultimately helping women to make informed decisions regarding treatment for breast cancer.

Introduction

Accurate prediction of survival will help both physicians and women with breast cancer to determine the type of surgery and the adjuvant therapy that are beneficial for these women [1]. Physicians usually make survival predictions and adjuvant therapy formulations according to the clinical characteristics of women with breast cancer, such as age, tumor size, clinical stage, hormone receptor status, and molecular characteristics, in addition to their own clinical experience [1-3]. The predicted adjuvant benefit is achieved by calculating the percentage improvement in the predicted survival rate after adjuvant treatment [1].

Currently, some survival prediction tools are widely used, such as PREDICT [1,4] and Adjuvant! Online [5-9]. PREDICT was developed based on clinical data of women with breast cancer collected from East Anglia (UK) between 1999 and 2003 [1,4], and Adjuvant! Online was developed based on data at the Netherlands Cancer Institute from 1987 to 1998 [5-9]. Studies have shown that such prediction tools underestimated or overestimated the survival rate of women with breast cancer, especially in Asia [10-12]. Differences in access to health care and the quality of existing health care among countries are possible reasons [13].

In China, the practice of surgery for breast cancer markedly differs from that in Western countries. A national survey reported that breast-conserving surgery accounted for 22% of surgeries in China [14], in contrast to 64.5% in the United States [15], despite the similar survival rates of breast-conserving surgery and mastectomy for women with early stage breast cancer [16,17]. Furthermore, some recent studies have shown that breast-conserving surgery has higher breast cancer–specific and overall survival rates than mastectomy [18]. The insufficient and unbalanced medical resources, information disparity between physicians and women with breast cancer, distrust of breast-conserving surgery by some physicians and women, and fear of recurrence are important reasons for the low rate of breast-conserving surgery in China [14]. In addition, for women with breast cancer in China who had received adjuvant endocrine therapy, the nonpersistence rates ranged from 5.9% to 22.6% in years 1-5, and the rate of low compliance ranged from 2.1% to 11.2% [19]. The low compliance with endocrine therapy may be partially explained by inadequate information and social support [19-23]. Therefore, demonstrating the similar survival rate of different surgeries and the efficacy of adherence to endocrine therapy may help physicians and women with breast cancer make better clinical decisions.

In 2020, the popularity of mobile phones in China exceeded 1 per person [24], and 920 million people could surf the internet using mobile phones [25]. Further, WeChat, a popular mobile phone platform, has over 1 billion monthly active users in China [26]. The Mini app, one of the functionalities of WeChat, provides an easily accessible platform to present the Chinese survival prediction model for women with breast cancer.

To the best of our knowledge, no prognostic app has been designed and applied to women with breast cancer in China. Such survival prediction tools would be beneficial to help physicians and women with breast cancer obtain and understand the survival rate of different surgeries and adjuvant therapy, thereby enabling better clinical decision making. The aim of this study is to develop and validate a prognostic app to predict the overall survival of women with breast cancer in China.

Methods

Study Design

A retrospective design was applied. All samples were randomly divided (7:3 ratio) into a training set for model construction and internal validation, and a test set for model external validation [27].

Study Setting and Population

Clinical data collection was conducted at 2 hospitals in Xiamen, a tertiary hospital and a hospital for women and children health. Data on death were collected from the Xiamen Center of Disease Control and Prevention (CDC), which retains all death records for Xiamen residents.

The inclusion criteria were women with breast cancer (1) who received treatment including surgery and adjuvant therapy in these 2 hospitals from January 1, 2009, to December 31, 2017; and (2) who were Xiamen residents. Women who lacked surgical and treatment information were excluded from the analysis.

Sample Size Calculation

The pmsampsize package for Stata and R written by Riley et al [28] was used to calculate the sample size. The 10-year mortality rate in women with breast cancer was 0.214, and the median follow-up was 11.7 years [6]. To target a margin of error in the estimate of death risk of ≤0.05, a total of 680 samples were

needed with Cox-Snell R^2 statistic of 0.1 [28], assumed outcome event rate of 0.214, follow-up of 11.7 years [6], timepoint of interest for 10 years' prediction, and 8 parameters (variables).

Procedure

The researchers collected clinical data of all patients with breast cancer who received treatment at 2 participating hospitals from January 2009 to December 2017. Further, the population information on death for Xiamen residents from January 2009 to December 2020 was obtained from the Xiamen Center CDC. The citizen ID number was used as the basis for matching. If the match was successful, the death dates of those women were recorded. If not successful, women with breast cancer were assumed to be alive. Finally, a data set was developed for model construction and validation.

Ethical Consideration

Ethical approvals were obtained from Research Ethics Committees at School of Medicine, Xiamen University (number XDYX2019008), First Hospital Affiliated to Xiamen University (number XMDY-2020-019), and Women and Children's Hospital affiliated to Xiamen University (number KY-2019-058). To protect the privacy of women with breast cancer, data collection, entry, and processing were carried out by 2 professional researchers. During the use of iCanPredict, the users' input information can only be accessed by authorized researchers, who needed to log-in to the iCanPredict website background thread with the protected account number and passport. Accordingly, unauthorized personnel was prohibited from viewing and using the clinical data and users' input information.

Data Collected

Outcomes data were obtained from the medical records (January 2009-December 2017) of the 2 participating hospitals including age at diagnosis, clinical stage (stage 0-IV), molecular classification (luminal A, luminal B, Her-2 overexpression, basal-like), operative type (breast conservative surgery, mastectomy), breast reconstruction, axillary lymph node dissection, chemotherapy, and endocrine therapy. These 2 participating hospitals used the TNM clinical stages from the American Joint Committee on Cancer, which are determined by tumor size (T), lymph node metastasis (N), and distant metastasis (M) [29].

Xiamen residents' death information from January 1, 2009, to December 15, 2020, was collected from the Xiamen Center CDC. Women's citizen ID number was retrieved from the hospital to link women's medical records to Xiamen residents' death information at the Xiamen Center CDC.

Outcomes

The 5-year (or 10-year) mortality was defined as the number of deaths divided by the total number of breast cancer cases from the diagnosis to 5-year (or 10-year) follow-up. The outcome event was the death of women with breast cancer from January 1, 2009, to December 15, 2020. The survival of women with breast cancer at the end of follow-up was considered as a truncated event [30].

Data Analysis

Data analysis was performed using SPSS version 25.0 (IBM) [31] and R version 4.0.4 (R Foundation). Demographic data were summarized using descriptive statistics. Because all variables satisfied the proportional hazards (PH) assumption, multivariate Cox regression analysis was used to develop a PH regression model for time-to-event data (eg, happening of death) and produce an equation to predict an individual's outcome risk on the condition of her values of multiple predictors [28].

Model Construction

The training set was used for model construction. On the basis of the hazard ratio (HR) of each prognosis factor estimated by multivariable Cox regression, the survival prediction model was expressed as

$$S(t) = [S_0(t)]^{\exp(\Sigma\beta_i \ast X_i - \Sigma\beta_i \ast \overline{X}_i)}$$

where $S(t)$ indicates the predicted survival rate of women with breast cancer after treatment, $S_0(t)$ represents the baseline survival rate of women with breast cancer, β is the coefficient for each prognostic factor in the multivariable Cox regression, X is the value of the prognosis factor, and $\overline{X}$ is the mean of the prognosis factor.

Validation

The training set was analyzed for model internal validation and the test set was analyzed for model external validation. Model discrimination was evaluated using the receiver operating characteristic (ROC) curve. In general, the larger the area under the curve (AUC), the higher the model discrimination. Further, an AUC greater than 0.7 indicates that the model has a certain distinctive ability. Model calibration was based on the Brier score; the lower the Brier score, the higher the calibration degree. When the Brier score was less than 0.25, the model could be predicted.

Design of a Prognostic App to Predict Survival

The research team comprised researchers, statisticians, breast cancer specialists, and app specialists at Xiamen University, the participating hospitals, and Tung Wah College of Hong Kong. The Xiamen Quanwu Information Service Company (Xiamen, People's Republic of China) undertook the technical development and maintenance of the iCanPredict app.

iCanPredict is mainly divided into a website background thread (back end) and the user side (front end). The website background thread was developed to approve the users' applications for registration, run the survival prediction model on the basis of the users' input information, and track the app usage frequency and duration. Only researchers are able to access the website background thread with protected account number and passport. The users' side was developed to interact with users and present the 10-year survival rate for women with breast cancer.

According to the Technology Acceptance Model proposed by Davis [32], perceived ease of use and perceived usefulness are 2 decisive variables of the user's reception of the information system. Perceived ease of use refers to the degree of difficulty a user perceives when using an information system [33].

Perceived usefulness refers to the fact that users can obtain certain values when using an information system [33]. Perceived ease of use affects perceived usefulness and further affects users' actual use behavior by affecting their attitudes and behaviors [33]. To promote ease of use, the iCanPredict app was incorporated into the functionality of WeChat, the most popular mobile phone platform in China. Prior to implementing the iCanPredict app, the perceived ease of use and perceived usefulness were tested by women with breast cancer and surgeons.

Results

Participants' Demographic and Clinic Characteristics

The clinical data of 1686 participants were collected, of which 97 lacked treatment information, leaving data on 1592 participants for model construction and validation. The training set comprised 1114 individuals and the test set comprised 478 individuals. By December 15, 2020, 147/1592 participants died (9.23%) and 1445/1592 (90.77%) were still alive. In this study, the mean follow-up was 6.38 years, ranging from 0.32 to 12.35 years. The mean age at diagnosis was 49.92 (SD 11.59) years, and most participants were diagnosed with stage II breast cancer (733/1592, 46.04%). Luminal B-type (1059/1592, 66.52%) was the most common molecular classification. Only 95/1592 participants (5.97%) underwent breast-conserving surgery, while 1497/1592 (94.03%) underwent mastectomy. A total of 453/1497 participants underwent breast reconstruction, accounting for 30.26% of the participants who underwent mastectomy and 28.45% (453/1592) of all participants. Most participants underwent axillary lymph node dissection (1143/1592, 71.80%), received adjuvant chemotherapy (1305/1592, 81.97%), and received adjuvant endocrine therapy (1007/1592, 63.25%). The demographic and clinical characteristics of the participants in the training and test sets are presented in Table 1.

Table 1. Demographic and clinical characteristics of participants in the training set and the test set.[a]

Characteristics	Total (n=1592)	Training set (n=1114)	Test set (n=478)
Deaths, n (%)	147 (9.23)	103 (9.25)	44 (9.21)
Patients alive, n (%)	1445 (90.77)	1011 (90.75)	434 (90.79)
Follow-up (years), mean (SD)	6.38 (2.68)	6.40 (2.68)	6.32 (2.68)
Age at diagnosis (years), mean (SD)	49.92 (11.59)	49.68 (11.48)	50.50 (11.82)
Clinical stage, n (%)			
0	26 (1.63)	21 (1.89)	5 (1.05)
I	524 (32.91)	363 (32.59)	161 (33.68)
II	733 (46.04)	508 (45.60)	225 (47.07)
III	286 (17.96)	204 (18.31)	82 (17.15)
IV	23 (1.44)	18 (1.62)	5 (1.05)
Molecular classification, n (%)			
Luminal A	215 (13.51)	160 (14.36)	55 (11.51)
Luminal B	1059 (66.52)	723 (64.90)	336 (70.29)
HER-2 (+)	232 (14.57)	168 (15.08)	64 (13.39)
Basal like	86 (5.40)	63 (5.66)	23 (4.81)
Operative type, n (%)			
Breast-conserving surgery	95 (5.97)	68 (6.10)	27 (5.65)
Mastectomy	1497 (94.03)	1046 (93.90)	451 (94.35)
Breast reconstruction, n (%)			
Yes	453 (28.45)	321 (28.82)	132 (27.62)
No	1139 (71.55)	793 (71.18)	346 (72.38)
Axillary lymph node dissection, n (%)			
Yes	1143 (71.80)	810 (72.71)	333 (69.67)
No	449 (28.20)	304 (27.29)	145 (30.33)
Chemotherapy, n (%)			
Yes	1305 (81.97)	919 (82.50)	386 (80.75)
No	287 (18.03)	195 (17.50)	92 (19.25)
Endocrine therapy, n (%)			
Yes	1007 (63.25)	707 (63.46)	300 (62.76)
No	585 (36.75)	407 (36.54)	178 (37.24)

[a]Percentage=number/group total number.

Initial Model Fit

The training set was used to fit the initial model. The graph based on the standardized Schoenfeld residual method and the –ln(–ln[survival]) test indicated that all 8 variables met the PH assumption. A survival prediction model was constructed for women with breast cancer (Table 2). Age at diagnosis (P=.002), clinical stage (P<.001), and endocrine therapy (P=.02) were identified as significant prognostic factors. The risk of death increased by 0.031 times per year with an increase in age at diagnosis for women with breast cancer (HR 1.031, 95% CI 1.011-1.051). When the clinical stage increased, the risk of death in women increased by 2.044 times per stage (HR 3.044, 95% CI 2.347-3.928). Women with breast cancer receiving endocrine therapy had a 0.407 (1–0.592) times lower risk of death than those that did not receive endocrine therapy (HR 0.592, 95% CI 0.384-0.914). The other 5 variables, including molecular classification (P=.91), operative type (P=.81), breast reconstruction (P=.36), axillary lymph node dissection (P=.32), and chemotherapy (P=.84), were not significant.

Table 2. Hazard ratios and model coefficients for prognostic factors included in the models (n=1114).

Prognostic factors	β	Standard error	Hazard ratio (95% CI)	P value
Age at diagnosis	0.031	0.01	1.031 (1.011-1.051)	.002
Clinical stage	1.113	0.133	3.044 (2.347-3.928)	<.001
Molecular classification	0.016	0.145	1.017 (0.765-1.351)	.91
Operative type	0.127	0.52	1.136 (0.410-3.145)	.81
Breast reconstruction	–0.318	0.343	0.728 (0.372-1.426)	.36
Axillary lymph node dissection	0.42	0.422	1.521 (0.665-3.478)	.32
Chemotherapy	–0.067	0.34	0.935 (0.480-1.821)	.84
Endocrine therapy	–0.524	0.222	0.592 (0.384-0.914)	.02

Validation Results

All statistically significant and nonsignificant factors are included in the prediction model because all factors are clinically significant. Further, the prediction model including nonsignificant factors presented model discrimination and calibration similar to those presented by the prediction model excluding nonsignificant factors. The AUC and Brier score of the training and test sets are presented in Tables 3 and 4, respectively. For internal validation with the training set, the ROC curve of the training set indicated that the model exhibited good discrimination for predicting 1- (AUC 0.802, 95% CI 0.713-0.892), 5- (AUC 0.813, 95% CI 0.760-0.865), and 10-year (AUC 0.740, 95% CI 0.672-0.808) overall survival (Figure 1A,C,E). The Brier scores at 1, 5, and 10 years after diagnosis were 0.005, 0.055, and 0.103, respectively. For external validation with the test set, except that the AUC at 10 years after diagnosis was 0.685, the AUC values 1 and 5 years after diagnosis were greater than 0.7, indicating that the model has a certain distinctive ability (Figure 1B,D,F). The Brier scores 1, 5, and 10 years after diagnosis were less than 0.25, indicating that the model calibration was also good. Therefore, the prediction model could be used to predict survival.

Table 3. Internal validation and external validation: model discrimination (AUC[a]) at 1, 5, and 10 years after diagnosis.

Year	Internal validation		External validation	
	AUC	95% CI	AUC	95% CI
1	0.802	0.713-0.892	0.857	0.725-0.988
5	0.813	0.760-0.865	0.738	0.634-0.841
10	0.740	0.672-0.808	0.685	0.580-0.790

[a]AUC: area under the curve.

Table 4. Internal validation and external validation: calibration (Brier) at 1, 5, and 10 years after diagnosis.

Year	Internal validation		External validation	
	Brier score	95% CI	Brier score	95% CI
1	0.005	0.001-0.010	0.014	0.004-0.025
5	0.055	0.043-0.067	0.057	0.038-0.075
10	0.103	0.083-0.124	0.120	0.084-0.156

Figure 1. Receiver operator characteristic (ROC) curves for breast cancer overall survival rates. (A) Training set at 1 year; (B) test set at 1 year; (C) training set at 5 years; (D) test set at 5 years; (E) training set at 10 years; (F) test set at 10 years.
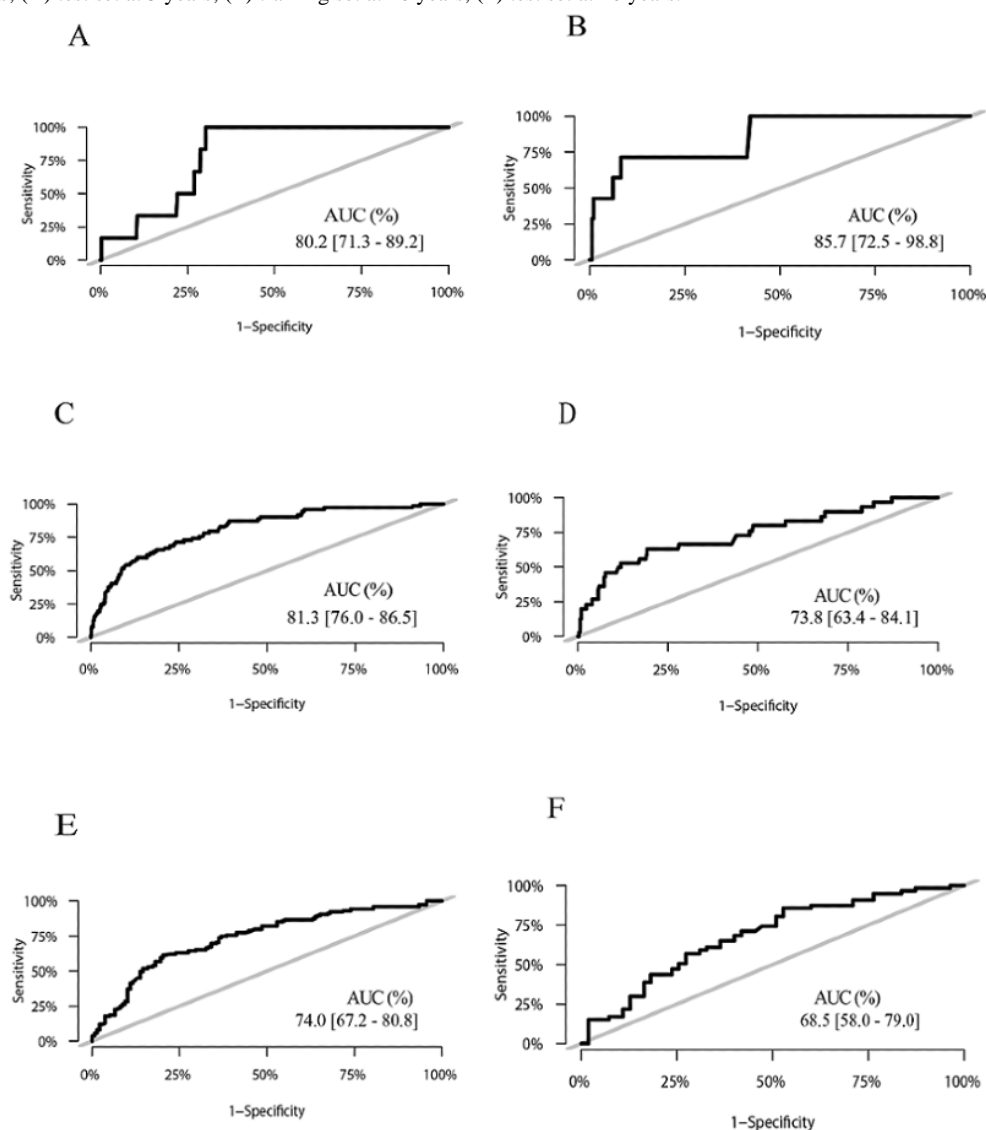
Development of Preliminary Survival Prediction for Breast Cancer Survivors

The development of iCanPredict was based on a uniAPP open-source free framework with separation technology for the back end and the front end. The back end was developed using the Java language 8 environment, IntelliJ IDEA for code development, and Maven for service dependency management. The front end utilized an Ant-design-VUE component for page interaction construction under the VUE framework. Each function in the iCanPredict app was tested multiple times to ensure usability and stability.

A QR code was generated for the users to scan with their own WeChat account. After downloading the iCanPredict app into their WeChat platform, the users need apply for registration. The researcher will approve the users' application from the website background thread. Each user's mobile phone number is set up as the unique username and an automated passport is sent to the user's mobile phone (changeable later). The users do not need to pay for the access to the iCanPredict app.

iCanPredict includes 2 function modules: survival prediction and a personal home page for women with breast cancer. The purpose and data sources are displayed on the home page with a statement as follows: " iCanPredict aims to be used for scientific research and provide the survival probability of different treatment options. Neither the model author nor the relevant hospitals can guarantee the calculation accuracy for any specific patient. Any decision related to cancer treatment should consult your physician". Women need to click the button "I agree" under the statement before using iCanPredict (Figure 2A).

Figure 2. Screenshots of the iCanPredict app.

A. Home page   B. Information input interface   C. Survival prediction

In the survival prediction model module, there are interactive page and report page. In the interactive page, users need to input personal information, including the year of birth, the year of diagnosis, breast cancer clinical stage, and molecular classification. In addition, users need to choose surgery type (breast-conserving surgery or mastectomy), whether they will undergo reconstruction if they choose mastectomy, whether they will undergo axillary lymph node dissection, and whether they will receive chemotherapy or endocrine therapy (Figure 2B). In addition, to make it easier and faster for users to learn how to input information, there are tutorials and concept explanations beside each concept in the input information interface. Tutorials and concept explanations can resolve doubts in the use process. After the completion of information input, users need to click the button "generate report" at the bottom of the interactive page.

In the report page, iCanPredict will present the 10-year survival rates after surgery only, surgery plus chemotherapy, and surgery plus chemotherapy plus endocrine therapy based on the user's input information. The survival rates are presented as tables and diagrams for women with breast cancer. The tables present 5- and 10-year survival rates, while the diagrams describe the continuous survival rates within 10 years after primary treatment. The predicted benefits of adjuvant chemotherapy and endocrine therapy are represented by the percentage increase in the predicted survival rate with differently colored curves in the diagram for women with breast cancer who received adjuvant therapy, compared with those who did not receive adjuvant therapy [1] (Figure 2C).

Users can make different choices on different surgeries and adjuvant therapies, and the corresponding survival prediction data will be saved on their personal home page. The personal home page module includes personal centers, problems feedback, and history. Personal center refers to the account-related information (nickname, avatar, basic information, etc.), which is set up by the users themselves. Problems feedback is an interactive platform for communication between users and the iCanPredict developer group. Users can provide feedback on the problems encountered in the use of iCanPredict or provide suggestions for app improvement. History summarizes user input information and survival prediction data are expressed in tables and diagrams and arranged in chronological order for users to compare the survival rates of different surgeries and adjuvant treatments. The flowchart of the use of iCanPredict is presented in Figure 3.

A total of 6 women with breast cancer and 2 surgeons were invited to assess the perceived ease of use and perceived usefulness. Of these, 2 women were aged over 60 years and were reluctant to use the iCanPredict app. One women got confused about her own value of molecular classification and needed to refer to the laboratory test report to input information. With help from the researchers to download the iCanPredict app and explain the laboratory test, all 6 women showed great interest in the prediction model. They appraised the iCanPredict app's convenience, ease of use, and ability to provide valuable information they wanted. The 2 surgeons indicated that it was easy to use the iCanPredict app. The surgeons expressed surprise that mastectomy and breast conservative surgery resulted in similar survival rates. They also suggested to add more information to predict survivals following other treatment options, such as immunotherapy or gene therapies. Such therapies have been emphasized to treat women with breast cancer in the recent years, but were not commonly applied before. Both women with breast cancer and surgeons indicated that it was feasible and acceptable to use the iCanPredict app in the clinics.

Figure 3. Flowchart of the use of the iCanPredict app.
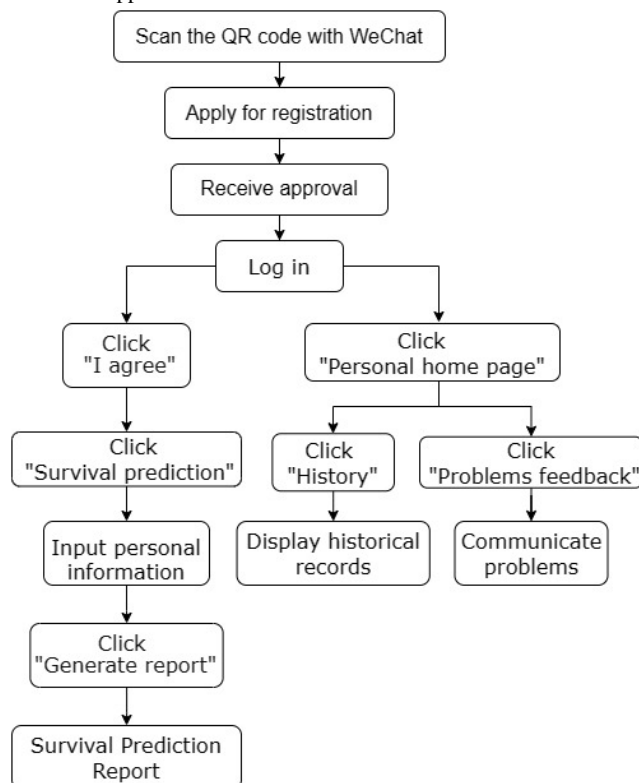
Discussion

Principal Findings

In this study, we constructed a well-calibrated prediction model on the basis of 9 years of clinical and death information for women with breast cancer in Xiamen. The survival prediction model was run in the app background thread, and the iCanPredict prognostic app was designed and developed for women with breast cancer in China.

iCanPredict was developed on the popular mobile phone platform in China (WeChat), which is easily accessible and user-friendly for both physicians and women with breast cancer [34]. iCanPredict provides individual survival predictions and calculates the benefits of the corresponding treatments. The tables display the percentage of predicted survival rates for physicians and women to compare the predicted results among different surgeries and adjuvant therapies. The use of colored diagrams to display the improvement in the predicted survival rate will facilitate sophisticated and difficult discussions on corresponding adjuvant therapies for patients [1].

Comparison With Prior Work

In our prediction model, the type of surgery was not an independent prognostic factor for breast cancer overall survival, which is consistent with some previous studies [16,17]. When physicians discuss the choices of breast surgeries with women, they may input women's clinical data and attempt different types of surgery in the iCanPredict app, which presents the survival prediction rate for different surgeries in an easy-to-understand manner. In combination with value preferences, women may thus make better informed decisions on surgeries [35,36].

In our prediction model, endocrine therapy was a significant prognostic factor for overall survival, which was also supported by the literature [37]. iCanPredict will present the benefit of endocrine therapy based on the percentage of improvement in the predicted survival rate if physicians or women choose endocrine therapy. iCanPredict will provide a quick, visual presentation, such as tables and diagrams, to inform women about the beneficial effects of endocrine therapy.

For clinical implementation, it is recommended that women with breast cancer use the iCanPredict together with surgeons when making decisions on treatments. Further, cautions should be exercised when using this iCanPredict app to guide the decision making, considering new therapies that are constantly emerging in recent years such as immunotherapy [38-41] or gene-based therapies [42,43], which may affect the survival rate of women with breast cancer.

Strengths and Limitations

iCanPredict is the first tool of its kind in China to provide survival predictions to women scheduled to undergo breast surgeries and corresponding adjuvant therapies. iCanPredict was proved to have good model discrimination and calibration. This study had several limitations. First, the sample size was relatively small to build a prediction model, as there were only limited number of eligible cases in the 2 participating hospitals, which may limit the generalization of the iCanPredict app. Second, death data from the Xiamen Center CDC did not include specific causes of death; therefore, this study did not have specific breast cancer survival rates, and only contained overall survival rates. Therefore, the iCanPredict app may overestimate the mortality of breast cancer. Third, due to the incomplete clinical records of adjuvant endocrine therapy and adjuvant chemotherapy treatment, we can only predict the benefits of

adjuvant therapy by determining whether adjuvant therapy was administered (yes vs no). In future studies, the adherence to adjuvant endocrine therapy should be emphasized to help clinicians and patients make accurate and individual medical decisions.

Future Directions

Future studies need to combine the results of this study with the financial burden of different treatments [44] and patients' own preferences [35,36] to provide better clinical decision support. Studies with large sample size and more comprehensive clinical data, in combination with the medical costs related to medical treatments and patients' value preference, are warranted to be conducted in different countries and cultures.

Conclusions

The prediction model had good model discrimination and model calibration to predict the overall survival of women with breast cancer in China. iCanPredict will increase women's awareness of the similar survival rate of different surgeries and the importance of adherence to endocrine therapy, ultimately helping them make informed decisions about treatment for breast cancer. In the long run, better choice of surgery and increased adherence to prescribed endocrine therapy may be attained.

Acknowledgments

Authors' Contributions

JZ and XZ designed the study. JZ and SWCC obtained the funding for this study. SH, XW, YL, YH, and XZ collected the data. ZM analyzed and interpreted the data. ZM, JZ, and XZ interpreted the results and wrote the manuscript. All authors read the manuscript.

Conflicts of Interest

None declared.

References

1. 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(1):R1. [doi: 10.1186/bcr2464] [Medline: 20053270]

2. Wang X, Feng Z, Huang Y, Li H, Cui P, Wang D, et al. A Nomogram To Predict The Overall Survival Of Breast Cancer Patients And Guide The Postoperative Adjuvant Chemotherapy In China. Cancer Manag Res 2019;11:10029-10039 [FREE Full text] [doi: 10.2147/CMAR.S215000] [Medline: 31819635]

3. 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 Aug;128(3):827-835. [doi: 10.1007/s10549-011-1366-9] [Medline: 21327471]

4. 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 May 22;19(1):58 [FREE Full text] [doi: 10.1186/s13058-017-0852-3] [Medline: 28532503]

5. Ravdin PM, Siminoff LA, Davis GJ, Mercer MB, Hewlett J, Gerson N, et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. J Clin Oncol 2001 Feb 15;19(4):980-991. [doi: 10.1200/JCO.2001.19.4.980] [Medline: 11181660]

6. Mook S, Schmidt MK, Rutgers EJ, van de Velde AO, Visser O, Rutgers SM, et al. Calibration and discriminatory accuracy of prognosis calculation for breast cancer with the online Adjuvant! program: a hospital-based retrospective cohort study. Lancet Oncol 2009 Nov;10(11):1070-1076. [doi: 10.1016/S1470-2045(09)70254-2] [Medline: 19801202]

7. Campbell HE, Taylor MA, Harris AL, Gray AM. An investigation into the performance of the Adjuvant! Online prognostic programme in early breast cancer for a cohort of patients in the United Kingdom. Br J Cancer 2009 Oct 06;101(7):1074-1084 [FREE Full text] [doi: 10.1038/sj.bjc.6605283] [Medline: 19724274]

8. Olivotto IA, Bajdik CD, Ravdin PM, Speers CH, Coldman AJ, Norris BD, et al. Population-based validation of the prognostic model ADJUVANT! for early breast cancer. J Clin Oncol 2005 Apr 20;23(12):2716-2725. [doi: 10.1200/JCO.2005.06.178] [Medline: 15837986]

9. 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 May;37(5):411-417. [doi: 10.1016/j.ejso.2011.02.001] [Medline: 21371853]

10. Yao-Lung K, Dar-Ren C, Tsai-Wang C. Accuracy validation of adjuvant! online in Taiwanese breast cancer patients--a 10-year analysis. BMC Med Inform Decis Mak 2012 Sep 17;12:108 [FREE Full text] [doi: 10.1186/1472-6947-12-108] [Medline: 22985190]

11. Bhoo-Pathy N, Yip C, Hartman M, Saxena N, Taib NA, Ho G, et al. Adjuvant! Online is overoptimistic in predicting survival of Asian breast cancer patients. Eur J Cancer 2012 May;48(7):982-989 [FREE Full text] [doi: 10.1016/j.ejca.2012.01.034] [Medline: 22366561]

12. de Glas NA, van de Water W, Engelhardt EG, Bastiaannet E, de Craen AJM, Kroep JR, et al. Validity of Adjuvant! Online program in older patients with breast cancer: a population-based study. Lancet Oncol 2014 Jun;15(7):722-729. [doi: 10.1016/S1470-2045(14)70200-1] [Medline: 24836274]

13. Mazo C, Kearns C, Mooney C, Gallagher WM. Clinical Decision Support Systems in Breast Cancer: A Systematic Review. Cancers (Basel) 2020 Feb 06;12(2):1-15 [FREE Full text] [doi: 10.3390/cancers12020369] [Medline: 32041094]

14. Yang B, Ren G, Song E, Pan D, Zhang J, Wang Y, et al. Current Status and Factors Influencing Surgical Options for Breast Cancer in China: A Nationwide Cross-Sectional Survey of 110 Hospitals. Oncologist 2020 Oct;25(10):e1473-e1480 [FREE Full text] [doi: 10.1634/theoncologist.2020-0001] [Medline: 32333626]

15. Kummerow KL, Du L, Penson DF, Shyr Y, Hooks MA. Nationwide trends in mastectomy for early-stage breast cancer. JAMA Surg 2015 Jan;150(1):9-16. [doi: 10.1001/jamasurg.2014.2895] [Medline: 25408966]

16. Sinnadurai S, Kwong A, Hartman M, Tan EY, Bhoo-Pathy NT, Dahlui M, et al. Breast-conserving surgery mastectomy in young women with breast cancer in Asian settings. BJS Open 2019 Feb;3(1):48-55 [FREE Full text] [doi: 10.1002/bjs5.50111] [Medline: 30734015]

17. van Maaren MC, de Munck L, de Bock GH, Jobsen JJ, van Dalen T, Linn SC, et al. 10 year survival after breast-conserving surgery plus radiotherapy compared with mastectomy in early breast cancer in the Netherlands: a population-based study. Lancet Oncol 2016 Aug;17(8):1158-1170. [doi: 10.1016/S1470-2045(16)30067-5] [Medline: 27344114]

18. Lagendijk M, van Maaren MC, Saadatmand S, Strobbe LJA, Poortmans PMP, Koppert LB, et al. Breast conserving therapy and mastectomy revisited: Breast cancer-specific survival and the influence of prognostic factors in 129,692 patients. Int J Cancer 2018 Jan 01;142(1):165-175 [FREE Full text] [doi: 10.1002/ijc.31034] [Medline: 28884470]

19. Xu H, Jin F, Zhang X, Wang D, Yu S, Wang A. Adherence status to Adjuvant Endocrine Therapy in Chinese Women with Early Breast Cancer and its influencing factors: A cross-sectional survey. Cancer Med 2020 Jun;9(11):3703-3713 [FREE Full text] [doi: 10.1002/cam4.3017] [Medline: 32237070]

20. Xu H, Zhang X, Wang D, Xu L, Wang A. Factors influencing medication-taking behaviour with adjuvant endocrine therapy in women with breast cancer: A qualitative systematic review. J Adv Nurs 2020 Feb;76(2):445-458. [doi: 10.1111/jan.14253] [Medline: 31657028]

21. Simon R, Latreille J, Matte C, Desjardins P, Bergeron E. Adherence to adjuvant endocrine therapy in estrogen receptor-positive breast cancer patients with regular follow-up. Can J Surg 2014 Feb;57(1):26-32 [FREE Full text] [doi: 10.1503/cjs.006211] [Medline: 24461223]

22. Heisig SR, Shedden-Mora MC, von Blanckenburg P, Schuricht F, Rief W, Albert U, et al. Informing women with breast cancer about endocrine therapy: effects on knowledge and adherence. Psychooncology 2015 Feb;24(2):130-137. [doi: 10.1002/pon.3611] [Medline: 24953538]

23. Friese CR, Pini TM, Li Y, Abrahamse PH, Graff JJ, Hamilton AS, et al. Adjuvant endocrine therapy initiation and persistence in a diverse sample of patients with breast cancer. Breast Cancer Res Treat 2013 Apr;138(3):931-939 [FREE Full text] [doi: 10.1007/s10549-013-2499-9] [Medline: 23542957]

24. Number of mobile cell phone subscriptions in China from December 2019 to December 2020 (in millions). Statista. URL: https://www.statista.com/statistics/278204/china-mobile-users-by-month/ [accessed 2021-09-19]

25. Statista. Number of mobile internet users in China from 2015 to 2019 with a forecast until 2023 (in millions). Statista. URL: https://www.statista.com/statistics/558731/number-of-mobile-internet-user-in-china/ [accessed 2021-09-19]

26. Xiao S, Li T, Zhou W, Shen M, Yu Y. WeChat-based mHealth intention and preferences among people living with schizophrenia. PeerJ 2020;8:e10550 [FREE Full text] [doi: 10.7717/peerj.10550] [Medline: 33362979]

27. Zhu J, Hao J, Ma Q, Shi T, Wang S, Yan J, et al. A Novel Prognostic Model and Practical Nomogram for Predicting the Outcomes of Colorectal Cancer: Based on Tumor Biomarkers and Log Odds of Positive Lymph Node Scheme. Front Oncol 2021;11:661040 [FREE Full text] [doi: 10.3389/fonc.2021.661040] [Medline: 33937076]

28. Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020 Mar 18;368:m441. [doi: 10.1136/bmj.m441] [Medline: 32188600]

29. Giuliano AE, Connolly JL, Edge SB, Mittendorf EA, Rugo HS, Solin LJ, et al. Breast Cancer-Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA Cancer J Clin 2017 Jul 08;67(4):290-303 [FREE Full text] [doi: 10.3322/caac.21393] [Medline: 28294295]

30. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part I: basic concepts and first analyses. Br J Cancer 2003 Jul 21;89(2):232-238 [FREE Full text] [doi: 10.1038/sj.bjc.6601118] [Medline: 12865907]

31. IBM Corporation. IBM SPSS Statistics for Windows (25th edition). 2017. URL: https://www.ibm.com/in-en/analytics/spss-statistics-software [accessed 2022-03-04]

32. Davis F. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep 28;13(3):319-6333. [doi: 10.2307/249008] [Medline: 32716422]

33. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]

34. Li Y, Ye S, Zhou Y, Mao F, Guo H, Lin Y, et al. Web-Based Medical Information Searching by Chinese Patients With Breast Cancer and its Influence on Survival: Observational Study. J Med Internet Res 2020 Apr 17;22(4):e16768 [FREE Full text] [doi: 10.2196/16768] [Medline: 32301740]

35. Hawley ST, Newman L, Griggs JJ, Kosir MA, Katz SJ. Evaluating a Decision Aid for Improving Decision Making in Patients with Early-stage Breast Cancer. Patient 2016 Apr;9(2):161-169 [FREE Full text] [doi: 10.1007/s40271-015-0135-y] [Medline: 26178202]

36. Minami CA, King TA, Mittendorf EA. Patient preferences for locoregional therapy in early-stage breast cancer. Breast Cancer Res Treat 2020 Sep;183(2):291-309. [doi: 10.1007/s10549-020-05737-9] [Medline: 32691377]

37. Davies C, Pan H, Godwin J, Gray R, Arriagada R, Raina V, Adjuvant Tamoxifen: Longer Against Shorter (ATLAS) Collaborative Group. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. Lancet 2013 Mar 09;381(9869):805-816 [FREE Full text] [doi: 10.1016/S0140-6736(12)61963-1] [Medline: 23219286]

38. Basu A, Ramamoorthi G, Jia Y, Faughn J, Wiener D, Awshah S, et al. Immunotherapy in breast cancer: Current status and future directions. Adv Cancer Res 2019;143:295-349. [doi: 10.1016/bs.acr.2019.03.006] [Medline: 31202361]

39. Marra A, Viale G, Curigliano G. Recent advances in triple negative breast cancer: the immunotherapy era. BMC Med 2019 May 09;17(1):90 [FREE Full text] [doi: 10.1186/s12916-019-1326-5] [Medline: 31068190]

40. Emens LA. Immunotherapy in Triple-Negative Breast Cancer. Cancer J 2021;27(1):59-66. [doi: 10.1097/PPO.0000000000000497] [Medline: 33475294]

41. Zhu Y, Zhu X, Tang C, Guan X, Zhang W. Progress and challenges of immunotherapy in triple-negative breast cancer. Biochim Biophys Acta Rev Cancer 2021 Dec;1876(2):188593. [doi: 10.1016/j.bbcan.2021.188593] [Medline: 34280474]

42. Jakub JW, Peled AW, Gray RJ, Greenup RA, Kiluk JV, Sacchini V, et al. Oncologic Safety of Prophylactic Nipple-Sparing Mastectomy in a Population With BRCA Mutations: A Multi-institutional Study. JAMA Surg 2018 Feb 01;153(2):123-129 [FREE Full text] [doi: 10.1001/jamasurg.2017.3422] [Medline: 28903167]

43. Diéras V, Han HS, Kaufman B, Wildiers H, Friedlander M, Ayoub J, et al. Veliparib with carboplatin and paclitaxel in BRCA-mutated advanced breast cancer (BROCADE3): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol 2020 Oct;21(10):1269-1282. [doi: 10.1016/S1470-2045(20)30447-2] [Medline: 32861273]

44. Fan L, Strasser-Weippl K, Li J, St LJ, Finkelstein DM, Yu K, et al. Breast cancer in China. Lancet Oncol 2014 Jun;15(7):e279-e289. [doi: 10.1016/S1470-2045(13)70567-9] [Medline: 24872111]

Abbreviations

AUC: area under the curve
HR: hazard ratio
PH: proportional hazards
ROC: receiver operating characteristic
Xiamen Center CDC: Xiamen Center of Disease Control and Prevention

XSL•FO
RenderX