

# An app to classify a 5-year survival in patients with breast cancer using the convolutional neural networks (CNN) in Microsoft Excel

## Development and usability study

Cheng-Yao Lin, MD<sup>a,b,c</sup>, Tsair-Wei Chien, MBA<sup>d</sup> , Yen-Hsun Chen, MD<sup>e</sup>, Yen-Ling Lee, MD<sup>f,g</sup>, Shih-Bin Su, MD PhD<sup>h,\*</sup>

### Abstract

**Background:** Breast cancer (BC) is the most common malignant cancer in women. A predictive model is required to predict the 5-year survival in patients with BC (5YSPBC) and improve the treatment quality by increasing their survival rate. However, no reports in literature about apps developed and designed in medical practice to classify the 5YSPBC. This study aimed to build a model to develop an app for an automatically accurate classification of the 5YSPBC.

**Methods:** A total of 1810 patients with BC were recruited in a hospital in Taiwan from the secondary data with codes on 53 characteristic variables that were endorsed by professional staff clerks as of December 31, 2019. Five models (i.e., convolutional neural network [CNN], artificial neural network, Naïve Bayes, K-nearest Neighbors Algorithm, and Logistic regression) and 3 tasks (i.e., extraction of feature variables, model comparison in accuracy [ACC] and stability, and app development) were performed to achieve the goal of developing an app to predict the 5YSPBC. The sensitivity, specificity, and receiver operating characteristic curve (area under ROC curve) on models across 2 scenarios of training (70%) and testing (30%) sets were compared. An app predicting the 5YSPBC was developed involving the model estimated parameters for a website assessment.

**Results:** We observed that the 15-variable CNN model yields higher ACC rates (0.87 and 0.86) with area under ROC curves of 0.80 and 0.78 (95% confidence interval 0.78–82 and 0.74–81) based on 1357 training and 540 testing cases an available app for patients predicting the 5YSPBC was successfully developed and demonstrated in this study.

**Conclusion:** The 15-variable CNN model with 38 parameters estimated using CNN for improving the ACC of the 5YSPBC has been particularly demonstrated in Microsoft Excel. An app developed for helping clinicians assess the 5YSPBC in clinical settings is required for application in the future.

**Abbreviations:** 5YSPBC = the 5-year survival in patients with BC, ACC = accuracy, AI = artificial intelligence, ANN = artificial neural network, AUC = area under ROC curve, BC = breast cancer, CI = confidence interval, CNN = convolutional neural network,

Editor: Jorddy Neves Cruz.

Ethics approval and consent to participate: Not applicable.

All data were downloaded from PubMed database at pubmed.com.

Consent to publish: This study was approved and monitored by the institutional review board (11004-L04) from Liuying Chi-Mei hospital. All patient identifiers were stripped before conducting this study.

Availability of data and materials: All data used in this study are available in Supplemental Digital Content files.

Competing interests: The authors declare that they have no competing interests.

This study was supported by research grant from Chi Mei Medical Center, Liuying, Tainan, Taiwan, with ID number of CLFHR11006.

The authors have no conflicts of interest to disclose.

Supplemental Digital Content is available for this article.

The datasets generated during and/or analyzed during the current study are publicly available.

<sup>a</sup> Division of Hematology-Oncology, Department of Internal Medicine, Chi Mei Medical Center, Liuying, Tainan, Taiwan, <sup>b</sup> Department of Senior Welfare and Services, Southern Taiwan University of Science and Technology, Tainan, Taiwan, <sup>c</sup> Department of Environmental and Occupational Health, National Cheng Kung University, Tainan, Taiwan, <sup>d</sup> Department of Medical Research, Chi-Mei Medical Center, Tainan, Taiwan, <sup>e</sup> Division of Hematology-Oncology, Department of Internal Medicine, Chi Mei Center, Liuying, Tainan, Taiwan, <sup>f</sup> Department of Oncology, Tainan Hospital, Ministry of Health and Welfare, Tainan, Taiwan, <sup>g</sup> Department of Environmental and Occupational Health, College of Medicine, National Cheng Kung University, Tainan, Taiwan, <sup>h</sup> Department of Occupational Medicine, Chi Mei Medical Center, Tainan, Taiwan.

\* Correspondence: Shih-Bin Su, Department of Occupational Medicine, Chi Mei Medical Center, Tainan, Taiwan (e-mail: shihbin1029@gmail.com).

Copyright © 2022 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Lin CY, Chien TW, Chen YH, Lee YL, Su SB. An app to classify a 5-year survival in patients with breast cancer using the convolutional neural networks (CNN) in Microsoft Excel: Development and usability study. *Medicine* 2022;101:4(e28697).

Received: 3 July 2021 / Received in final form: 9 December 2021 / Accepted: 4 January 2022

<http://dx.doi.org/10.1097/MD.00000000000028697>

KNN = k-nearest neighbors algorithm, LR = logistic regression, ML = machine learning, NB = Naïve Bayes, SD = standard deviation, SENS = sensitivity, SPEC = specificity.

**Keywords:** 5-year survival in patients with breast cancer, artificial neural network, Microsoft Excel, neural revolution network, receiver operating characteristic curve, sensitivity, specificity

### Key Points

- Using the convolutional neural networks (CNN) to develop an app to classify a 5-year survival in patients with breast cancer has not been reported in literature.
- For the purpose of shared decision-making in medicine (SDM) effectively implemented in healthcare setting, the study provides an app for helping physicians and technicians discuss procedures with patients and their family members as a communication tool used in hospital.
- The way to evaluate the survivals in patients with breast cancer can be applied to other cancer-related patients using CNN-based app to predict the 5YSPBC in the future.

## 1. Introduction

Breast cancer (BC) is the most common malignant cancer in women. Over 173,541 articles were found in search for “breast cancer” in the title and 70 documents addressing prediction and survival rate as of September 18, 2020.<sup>[1]</sup> However, none were using the neural network technique to predict survival of patients with BC up to several years (e.g., 5 years) after its initial diagnosis.

### 1.1. The 5-year survival in patients with BC

BC is the leading cause of cancer deaths in women, accounting for >1.6% of deaths worldwide,<sup>[2]</sup> and seems to be the most common cause of death in female patients aging 40 to 55 years.<sup>[3]</sup> Around 5.8 million women will die from BC by 2025, which is equivalent to 28 women being diagnosed every day and nearly 2000 death from BC every day in the United States.<sup>[4]</sup> Based on the study<sup>[5]</sup> proposing BC prediction and diagnosis system based on the Rough Set, whether a predictive model can be developed and designed along with an app for classifying 5-year survival in patients with BC (5YSPBC) used in clinical settings should be verified.

### 1.2. Factors are complex in the 5YSPBC prediction

BC incidence increases with age; however, patients with comorbidities would significantly influence their treatment alternatives and overall survival.<sup>[6]</sup> Prognostic factors and survival for patients with BC are important, but its cause is difficult to understand. BC survival causes are worthy of investigation and exploration, and app development for clinicians and patients.

Correct diagnosis and treatment would be interacted on a good prognosis.<sup>[7]</sup> Thus, no medical guarantee for the success rate of each treatment reaching 100% of effect for survival. An accurate prediction of survival classification for patients with BC surgery (and/or other factors) after 5 years is required to develop shared decision-making in medicine (SDM). In which both the patient and the physician contribute to the SDM process and agree on treatment decisions.<sup>[8]</sup> As such, healthcare providers can explain treatments and alternatives to patients and help patients and their family

members choose appropriate treatments that best align their preferences as well as their unique cultural and personal beliefs.<sup>[9,10]</sup> Accordingly, the development of app is required to assist the communication of SDM between patients and professional clinicians.

In recent years, with the progress of information technology, computer science, particularly with machine learning techniques, has been widely launched onto medicine.<sup>[11]</sup> Computer algorithms with high accuracy (ACC) in prediction can be designed and used by professionals for the timely detection of BC survival.<sup>[12]</sup> Information technology application to construct empirical medicine, establish clinical medical treatment guidelines, and apply disease prognosis prediction (or survival as proposed in this study) are urgently essential and necessary to assist patients and physicians in clinical decision-making accordingly.<sup>[9,10]</sup>

### 1.3. Online classification using smartphones

The use of sophisticated modeling approaches using machine learning (ML),<sup>[12]</sup> such as artificial neural network (ANN),<sup>[13,14]</sup> convolutional neural network (CNN),<sup>[15–17]</sup> Naïve Bayes (NB),<sup>[18]</sup> K-nearest Neighbors Algorithm (KNN),<sup>[19]</sup> and logistic regression (LR)<sup>[20–24]</sup> are required to extract meaningful information from biology datasets and play an important role in healthcare settings.<sup>[25–28]</sup> Whether the CNN can improve the prediction ACC on the classification of the 5YSPBC higher than other algorithms as reported in the increased CNN of 7.14%<sup>[29]</sup> is worthy of study and verification. The first hypothesis of CNN with higher predictive ACC than other model algorithms was made for verification.

As with all forms of web-based technology, advances in mobile health communication technology are rapidly increasing.<sup>[30]</sup> Until now, no smartphone app to classify the 5YSPBC using the ML technique has been proposed in the literature. Therefore, ML is used to create a model to predict the 5YSPBC and improve the treatment quality of increasing the survival rate of patients with BC. The second hypothesis of app developed for helping clinicians assessing the 5YSPBC in hospital settings was proposed in this study.

### 1.4. Study aims

This study aimed to build an ML model to develop an app to automatically and accurately classify the 5YSPBC and improve the treatment quality and increase the survival rate of patients with BC.

Two hypotheses were made for verification, including the CNN has higher predictive ACC than other model algorithms and an app that can be developed and designed to help clinicians in assessing the 5YSPBC.

## 2. Methods

### 2.1. Data source

A total of 1810 patients with BC were included in this study. Secondary data came from the registration cancer center of Liuying Chi-Mei hospital in southern Taiwan with complete

variable assessment evaluated by professional staff clerks as of December 31, 2019; see those 53 characteristic variables illustrated in Supplemental Digital Content 1, <http://links.lww.com/MD/G601>. Wherein several feature variables would be extracted from 53 eligible variables (shown in Supplemental Digital Content 1, <http://links.lww.com/MD/G601>). The coding scheme (assigned by professional staff clerks) was illustrated in Table 1 and Supplemental Digital Content 1, <http://links.lww.com/MD/G601>. All missing data were filled out with further discussion in the study team with proper judgment.

This study was approved and monitored by the institutional review board (11004-L04) from Liuying Chi-Mei hospital. All patient identifiers were stripped before conducting this study.

**2.2. Feature variables (Task 1)**

A multidisciplinary team was established, including physicians and specialists in BC, data scientists, information engineers, nurse practitioners (e.g., case-management clerks for patients with BC), and quality managers for this study on artificial intelligence implementation of BC-survival-prediction model development. Inclusion criteria of cases were determined by the multidisciplinary team—patients without complete records in the first 5 years after diagnosis were excluded.

Feature variables were extracted from 53 items (shown in Supplemental Digital Content 1, <http://links.lww.com/MD/G601>) that were determined by the multidisciplinary team using the Weka software<sup>[31]</sup> (Orlando, FL, USA) via the following steps: standardize each variable to the mean (0) and standard deviation (i.e., SD = 1), use the search method ([Select Attributes]/[InfoGainAttributeEval] [Attribute Evaluator]/[Ranker][Search Method]), use full training sets, and click on the suggested feature items.<sup>[13]</sup>

Forest plots<sup>[32–34]</sup> were drawn to present the extracted feature variables. The standard mean difference (SMD) method was utilized to compare differences in variables alone (such as the *t* test) and with hospital types (such as an analysis of variance) using the forest plot.<sup>[13]</sup>

The Chi-square test was conducted to assess the heterogeneity between variables. The forest plots (confidence interval [CI] plot) were drawn to display the effect estimates and their CIs for each study.

**2.3. Comparison of model accuracy and prediction stability in prediction and classification (Task 2)**

**2.3.1. Comparison of 5 proposed models.** Five predictive models (i.e., ANN,<sup>[13,14]</sup> CNN,<sup>[15–17]</sup> NB,<sup>[18]</sup> KNN,<sup>[19]</sup> and LR<sup>[20–24]</sup>) along with 2 scenarios (i.e., accuracies in training and testing sets, accounting for 70% and 30%, respectively) were compared. All 5 modules, their algorithmic process, and MP4 abstract video are illustrated and deposited in Supplemental Digital Content 1, <http://links.lww.com/MD/G601>.

**2.3.2. Model building and scenarios in comparison.** This study is focused on model ACC (e.g., >0.8) and prediction stability (or generalizability, e.g., the consistency between training and test sets, e.g., >0.7) out of various perspectives, such as model feasibility, efficacy, and efficiency, using the maximum receiver operating characteristic curve (area under ROC curve [AUC]) between models using the training cases to predict the learning cases; see the following steps to create the prediction models and design the scenarios in comparison:

**Table 1**  
**Illustration of the coding schema of response in this study.**

Variable and coding schema	n	%
Age		
1. <40 yrs	261	14.42
2. 41–50 yrs	614	33.92
3. 51–60 yrs	512	28.29
4. >60 yrs	423	23.37
Pathologic M		
1: M0	985	54.42
2: M1	33	1.82
3: M1b uncertain	706	39.01
4: M1c (found during or after surgery but not confirmed by pathology M1c)	1	0.06
5: Null	85	4.7
Clinical N		
1: N0	1144	63.2
2: N1	381	21.05
3: N2	50	2.76
4: N2a	4	0.22
5: N3	19	1.05
6: N3a	2	0.11
7: N3b	3	0.17
8: N3c	12	0.66
9. Null	195	10.77
Clinical M		
1: M0	1533	84.7
2: M1	105	5.8
3. Null	172	9.5
Cancer status		
1: There is no evidence of this primary cancer	1488	82.21
2: There is such primary cancer clinically	322	17.79
Vital status		
0: Non-survival	298	16.46
1: Survival	1512	83.54

Note. The coding schema is referred to as Supplemental Digital Content 1, <http://links.lww.com/MD/G601>.

**2.3.3. Models in comparison.** These five models were analyzed using 2 scenarios previously mentioned. A total of 1810 BC cases were randomly split into training and test sets in a proportion of 70% (n = 1267) to 30% (n = 543), where the training set was used to predict the testing set. The ACC (e.g., sensitivity [SENS], specificity [SPEC], ACC, precision [PREC], and AUC [>0.80]) and stability (or generalizability, e.g., using the training set to predict the test set evaluated by observing the AUC) were verified (e.g., AUC >0.70). The training and testing sets in models are provided in Supplemental Digital Content 1, <http://links.lww.com/MD/G601>.

**2.3.4. Data presentations of model accuracy in models.** Accuracy was determined by observing the high AUC along with indicators of SENS, SPEC, and ACC, etc, in models. The definitions are listed as follows:

$$\text{True positive(TP)} = \frac{\text{the number of predicted survival to the true survival}}{\text{the true survival}} \quad (1)$$

$$\text{True negative(TN)} = \frac{\text{the number of predicted Non – survival to the true Non – survival}}{\text{the true Non – survival}} \quad (2)$$

False positive (FP) = the number of Non – survival – TN (3)

False negative (FN) = the number of survival – TP (4)

Sensitivity (SENS) = true positive rate (TPR)  
= TP ÷ (TP + FN) (5)

Specificity (SPEC) = true negative rate (TNR)  
= TN ÷ (TN + FP) (6)

ACC = accuracy = (TP + TN) ÷ N (7)

N = TP + TN + FP + FN (8)

AUC = (1 – Specificity) × Sensitivity ÷ 2 + (Sensitivity + 1) × Specificity ÷ 2 (9)

SE for AUC =  $\sqrt{(AUC \times (1 - AUC) \div N)}$  (10)

95% CI = AUC ± 1.96 × SE for AUC (11)

Accuracy rate = (TP + TN) / (TP + TN + FP + FN) (12)

**2.3.5. Selection of prediction models referring to accuracy in algorithms.** Comparison of accuracies in training and testing sets was made by observing high indicators of AUC among the 5 models to better understand its effectiveness and efficacy. That is, all indicators are based on high AUC rather than ACC in Eq. (12) since imbalanced-class data exist in this study (e.g., 83.54% vs 16.46% of the ratio between survival and non-survival shown in Table 1). High ACC rates with imbalanced SENS and SPEC are expected in imbalanced-class data using traditional approaches.<sup>[35–38]</sup> Accordingly, we applied the minimization of average model residuals in both classes to obtain balanced SENS and SPEC and overcome the disadvantage of high ACC rates (i.e., minimum residuals minimized by the formula of average [residuals in survival group] + average [residuals in the non-survival group]) as did in the previous study.<sup>[13]</sup> Notably, a balanced SENS and SPEC using professional ML software are hard to gain when an imbalanced number of classes exist, unless the method of minimizing model residuals is controlled by the user as the module provided in Supplemental Digital Content 1, <http://links.lww.com/MD/G601>.

**2.4. Developing an app for classifying the 5YSPBC (Task 3)**

An app for classifying the 5YSPSC was designed and developed. Model parameters were embedded in the computer module. Results of classification (i.e., survival and non-survival) instantly appear on smartphones. The visual representation with binary (i.e., survival and non-survival) categorical probabilities is shown on a dashboard displayed on Google Maps.

**2.5. Statistical tools and data analysis**

International Business Machines Statistical Package for the Social Sciences 22.0 for Windows (SPSS Inc., Chicago, IL) and MedCalc 9.5.0.0 for Windows (MedCalc Software, Ostend, Belgium) was used to obtain the descriptive statistics and frequency distributions among groups and compute the model prediction indicators expressed in Eqs. (1)–(12). The significance level of type I errors was set at 0.05. All those proposed models were performed on MS Excel; see Supplemental Digital Content 1, <http://links.lww.com/MD/G601>.

A visual representation of the classification was plotted using 2 curves based on the probability theory of the Rasch model.<sup>[39,40]</sup> Three tasks of data representations under 5 proposed models are involved in obtaining results; see the study flowchart in Fig. 1.

**3. Results**

**3.1. Demographical patient characteristics**

A total of 1810 patients with BC (5YSPBC survival at 83.54%, n=1512) were recruited in this study. The most number of patients in groups were 33.92% (n=614) aged 41 to 50 years, 30.22% (n=547) in body mass index group ranging from 18.5 to 24, 92.87% (n=1681) classified as first malignant tumor or carcinoma in patient’s life, 30.61% (n=554) diagnosed with primary site at the upper-outer quadrant of the breast, 22.04% (n=399) with pathologic stage at Stage IIA, and 54.42% (n=985) with pathologic M at M0. Other demographical

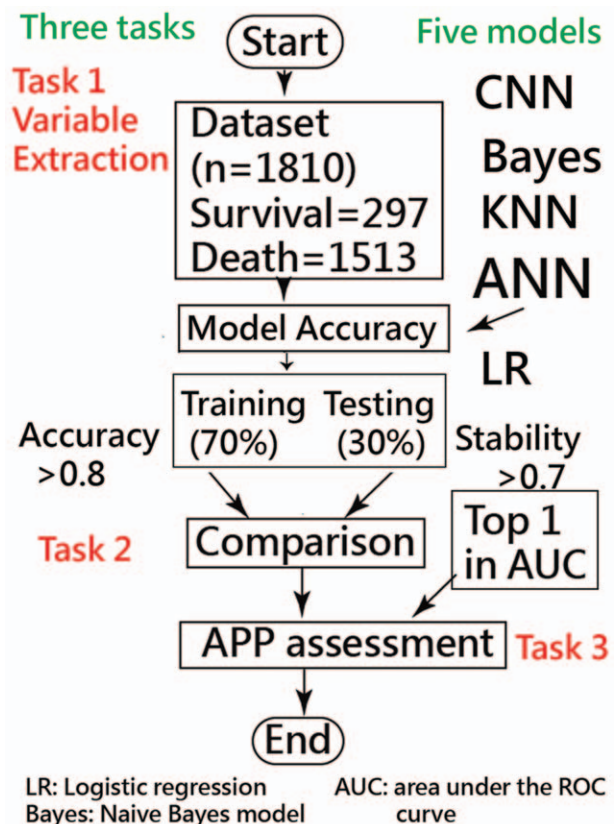


Figure 1. Study flowchart.

**Table 2**  
The 15 variables extracted from the 53 eligible variables.

Stage of variable	No.	Variable	Definition
TNM	1	Pathologic M	Refers to whether there is a remote transfer
Characteristics	2	Body BMI	kg/m <sup>2</sup>
TNM	3	Scope of regional lymph node surgery at this facility	The scope of simultaneous removal, slicing, or aspiration of regional lymph nodes during the primary site operation or another independent operation in the reporting hospital
Treatment	4	Whether the surgery of the primary site	Have clear surgical records and dates
TNM	5	Clinical N	Refers to whether there is regional lymph node metastasis and the extent of metastasis
TNM	6	Regional lymph nodes positive	Total number of regional lymph nodes positive by a pathologist
TNM	7	Clinical M	Refers to whether there is a remote transfer
Treatment	8	Reason for no surgery of primary site	The reason why the case was not operated on the primary site in any medical institution
Cancer	9	Grade/Differentiation	Describe the similarity of tumors and normal tissues
Cancer	10	Tumor size	The largest size or diameter of the primary tumor
Treatment	11	Surgical margins of the primary site	The final state of the surgical margin after resection of the primary tumor
TNM	12	Clinical stage group	Based on clinical T, N, and M to determine the degree of disease invasion on the anatomical site.
Recurrence	13	Date of first recurrence	From the date of the first diagnosis to the date of confirmation of the relapse diagnosis
Recurrence	14	Type of first recurrence	The case records the type of the first relapse after a period of disease-free intermission or remission
Recurrence	15	Cancer status	Whether the case had cancer at the "last contact or death date"

Note. The coding schema is referred to as Supplemental Digital Content 1, <http://links.lww.com/MD/G601>.

characteristics are included in Supplemental Digital Content 1, <http://links.lww.com/MD/G601>.

**3.2. Task 1: Feature variables extracted from data**

Of the original 53 variables, 15 feature variables in Table 2 were extracted using the Weka software.<sup>[31]</sup> Figure 2<sup>[41]</sup> shows the SMD method used in the meta-analysis.<sup>[32-34,42]</sup> We can see that 3 variables at the top panel in Fig. 2 (denoted by P in the far-left column) have the characteristics of higher values toward non-survival, 9 variables (denoted by N in the far-left column) are negative toward non-survival (i.e., the smaller, the better), and 3

variables are the delay-effect indicators not used for predicting BC, but for BC survival classification, with negative characteristics toward non-survival as well. The overall effect presents a significant statistical difference between the 2 groups of survival and non-survival.

**3.3. Task 2: Comparisons of accuracies in training and test samples**

When comparing the 5 models with the 1267 and 543 cases in the training and testing sets, the KNN3 (i.e., in the least 3 distances with the highest ACC among all possible selections of least

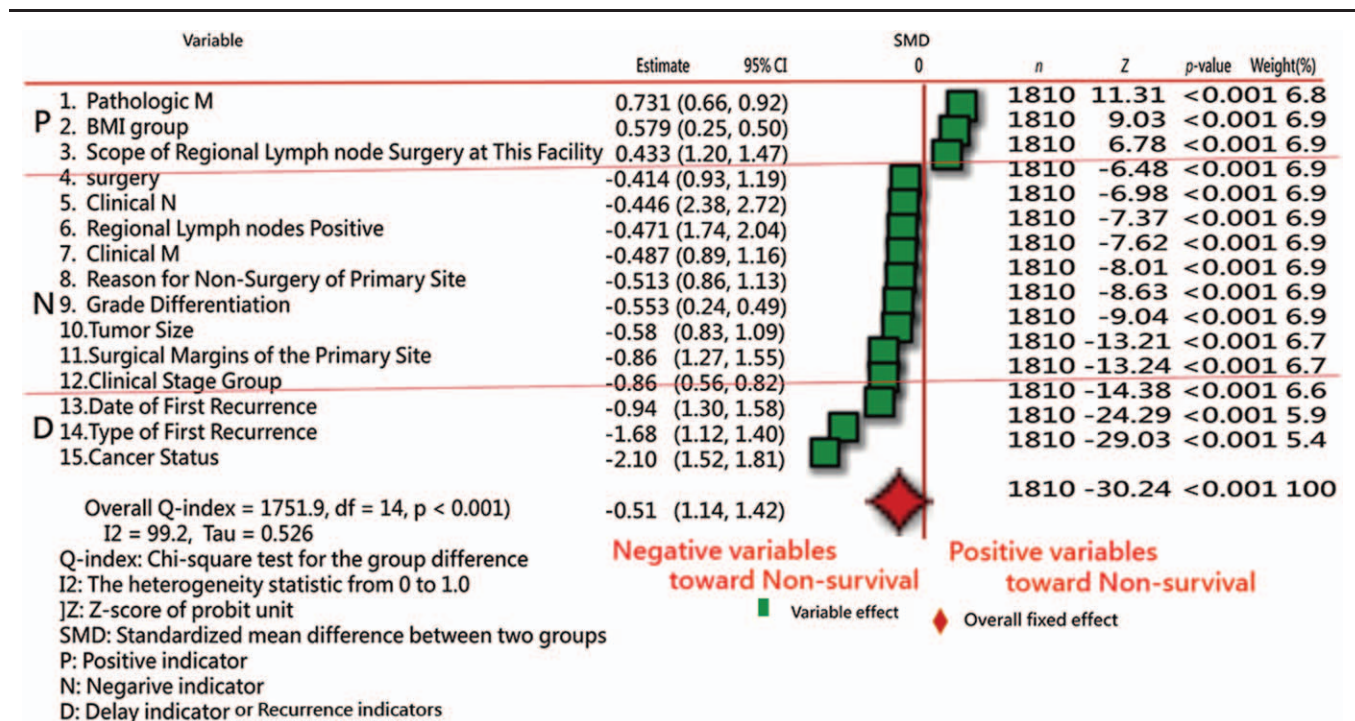


Figure 2. Feature variable comparison between two groups (survival vs non-survival) using the forest plot.

**Table 3**  
**Comparison of predictive models across indicators of accuracy and AUC.**

Model	n	Sensitivity	Specificity	Precision	F <sub>1</sub> score	Accuracy	AUC	95% CI
ANN								
Training	1357	0.87	0.58	0.88	0.87	0.8	0.72	0.7–0.75
Testing	540	0.83	0.49	0.9	0.86	0.78	0.66	0.62–0.70
CNN								
Training	1357	0.92	0.68	0.91	0.78	0.87	0.8	0.78–0.82
Testing	540	0.9	0.66	0.93	0.76	0.86	0.78	0.74–0.81
KNN3								
Training	1357	0.93	0.77	0.94	0.93	0.89	0.85	0.83–0.87
Testing	540	0.87	0.6	0.93	0.9	0.82	0.73	0.69–0.77
LR								
Training	1357	0.97	0.52	0.88	0.68	0.87	0.75	0.72–0.77
Testing	540	0.96	0.54	0.92	0.69	0.89	0.75	0.71–0.79
Bayes								
Training	1357	0.83	0.61	0.88	0.7	0.78	0.72	0.69–0.74
Testing	540	0.72	0.76	0.94	0.74	0.72	0.74	0.70–0.78

The highest model stability (=0.78) from the testing set predicted by the training set is from the CNN model.  
 ANN=artificial neural network, CNN=convolutional neural network, KNN=k-nearest neighbors algorithm, LR=logistic regression.

distances using the mode function in Microsoft Excel to determine the classification of survival or non-survival) model has higher AUC (=0.85) than others, indicating that the KNN3 model has higher ACC in the training set. However, the CNN has the highest prediction stability (i.e., AUC=0.78) than other counterparts (see the AUC in rows of testing in Table 3).

**3.4. Task 3: App created to predict the 5YSPBC**

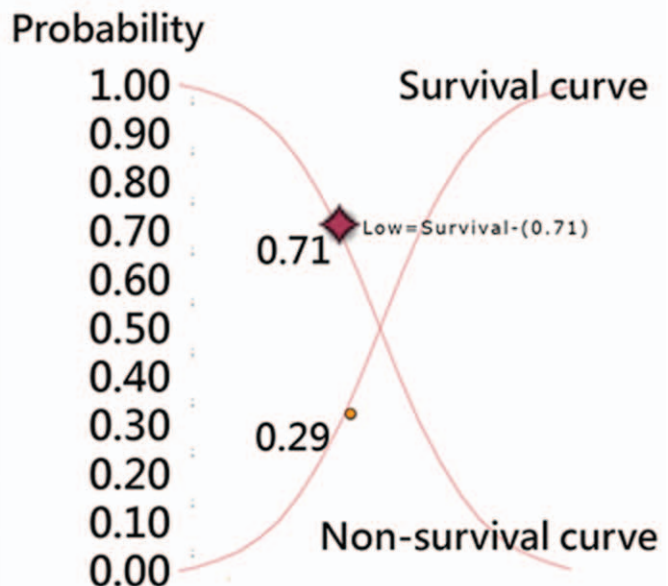
An app to predict the 5YSPBC for clinicians of Liuying Chi-Mei hospital in southern Taiwan was developed (Fig. 3). Readers are invited to click on the link in the reference<sup>[21]</sup> to practice the app on their own. It is worth noting that all 38 model parameters

are embedded in the 15-variable CNN model to classify either non-survival– or survival+ once all 15 items have been responded to.

One resulting example is presented in Fig. 3 (right), from which we can see the high probability survival (=0.71) is shown on the fail curve from the left-top to the right-bottom corner and the survival+ with a low probability (.29) on the success curve. The sum of both probabilities (i.e., survival+ and survival–) equals 1.0. The odds can be computed by the formula  $(P/[1-P]=0.29/0.71=0.41)$ , indicating the patient with a moderately high probability or tendency toward non-survival. An app to predict the 5YSPBC was developed involving 56 models with 56 estimated parameters for website classification.

**A.Snapshot of data entry on an APP**

**B,Classification result with category probabilities**



**Figure 3.** Data entry and assessment result.

### 3.5. Online dashboards shown on google maps

Two QR-codes shown in figures (or links)<sup>[41,43]</sup> are provided for readers who can manipulate the dashboards on their own.

## 4. Discussion

### 4.1. Principle findings

We observed that the 15-variable CNN model yields higher ACC rates (0.87 and 0.86) with AUCs of 0.80 and 0.78 (95% CI 0.78–82 and 0.74–81) on the 1267 training and 543 testing cases and an available app to predict the 5YSPBC was successfully developed and demonstrated in this study.

The aim of this study was thus achieved, and 2 hypotheses were supported, including the CNN has higher predictive ACC than other model algorithms and an app can be developed and designed to help clinicians in assessing the 5YSPBC.

### 4.2. Review of research findings

The incidence of BC increases with age. In patients with BC, comorbidities would significantly influence their treatment alternatives and overall survival.<sup>[6]</sup> As such, the prognostic factors and survival of patients with BC are important to the identification of BC survival.

BC is a fairly complex disease for identifying prognostic factors.<sup>[44]</sup> The initial diagnosis of BC may increase the chance of treatment effect. The early diagnosis of BC is crucial to treatment effect.<sup>[45]</sup> However, understanding the prognostic factors can increase the 5YSPBC as we did in this study. Differences in prognosis are still observed even if patients with the same stage receive the same treatment.<sup>[46]</sup> Accordingly, the study of prognostic cancer factors (e.g., those with 3 types of variables in Fig. 2) for clinical treatment can help not only understand the degree of disease deterioration but also provide patients with information about treatment effects with appropriate treatment strategies, and finally improve the patient survival.

BC mortality improvements can be attributed to diagnosis (e.g., mammographic screening<sup>[47]</sup>) and treatment.<sup>[48]</sup> We merely focused on the feature variables of BC treatment in the first 5 years after diagnosis using the variable assessment evaluated by professional staff clerks in the hospital. Most stage-specific survival improvement in women younger than 70 years old is unexplained by tumor size and estrogen receptor (ER) status; however, suggesting a key role for treatment, tumor size contributed importantly for women  $\geq 70$  years old with local and regional stage and stratification by tumor size and ER status explained, even more, the survival improvement among women age  $\geq 70$  years in the first 5 years after diagnosis.<sup>[49]</sup> Furthermore, 3.3% of deaths were avoided within 5 years of diagnosis among patients with cancer, female BC accounting for most up to 28% (=4822/17,041).<sup>[50]</sup> As such, feature variables extracted from the data are worthy of investigation, as in Fig. 2.

The SDM (i.e., both the patient and the physician contribute to the medical decision-making process and agree on treatment decisions<sup>[8]</sup>) can be informative in helping patients, and their family members choose the treatment option that best aligns with their preferences as well as their unique cultural and personal beliefs<sup>[9,10]</sup> based on the app tool to simulate the alternative and results of patients with BC.

Over 70 published articles adopted tools to predict the survival rates as of June 3, 2021.<sup>[1]</sup> However, none were using neural

network technique to classify the 5YSPBC for patients after an initial diagnosis of BC so far.

Different types of classification algorithms in ML,<sup>[51,52]</sup> such as Logistic Regression, Support Vector Machines,<sup>[44]</sup> NB,<sup>[18]</sup> KNN,<sup>[19,53]</sup> LR,<sup>[20–24]</sup> ANN,<sup>[13,14]</sup> and CNN<sup>[15–17]</sup> are present. The CNN model used in this study was the superior classifier of stability with an AUC of 0.78 based on the 543 testing cases higher than other counterparts.

### 4.3. Implications and applications of this study

CNN can improve the prediction ACC (up to 7.14%).<sup>[15,29]</sup> In this study, CNN was evident of superior classifier than others. No difference was found in these 5 when their 95% CIs were compared; however, anyone using the CNN approach to predict the 5YSPBC in the literature was not seen, which is a breakthrough and the first feature of this study.

Over 1282 articles have been found using the keyword “convolutional neural network” (Title) searched in PubMed Central (PMC) as of June 3, 2021, fewer than those with 2048 articles in PMC using the keyword in search of “ANN” (Title). Until now, no one used Microsoft Excel to perform CNN to classify the 5YSPBC. Interpretations for the CNN concept and process or even the parameter estimations are shown in studies<sup>[15–17]</sup> and Supplemental Digital Content 1, <http://links.lww.com/MD/G601>, which is the second feature of this study and is rarely seen applicable in literature.

The principle concerning more with the vital few and less with the numerous trivial cases is frequently emphasized in the quality control process.<sup>[54]</sup> Therefore the CNN-based app is proposed as the third feature, as shown in Fig. 3.

Furthermore, curves of category probabilities based on the Rasch rating scale model<sup>[39,40]</sup> were used to interpret the classification on an app. The binary categories (e.g., success and failure on an assessment in the psychometric field) have been applied in health-related outcomes.<sup>[13–17]</sup> However, none provided the animation-type dashboard showing on Google Maps in use for patients predicting the 5YSPBC, as we did in Fig. 3.

### 4.4. Strengths of this study

We proposed the CNN algorithm along with the 38-model parameters to design the routine on an app that is used to predict the 5YSPBC in hospitals, which was never seen before in literature.

As with all forms of web-based technology, advances in health communication technology are rapidly emerging.<sup>[30]</sup> The mobile online 5YSPBC app is promising and worth generalizing the model to many other fields of health assessment. The online app can be applied to inform patients, their family members, or physicians with the cancer case manager quickly about the 5YSPBC and prognostic factors for providing patients with information about treatment effect with appropriate treatment strategies, and finally improve the patient survival as mentioned in SDM in previous studies.<sup>[8–10]</sup>

The 5YSPBC app is promising and worth using for promoting patient health literacy by using the animation-type assessment on smartphones. Readers are recommended to click on the link in the reference<sup>[41,43]</sup> and watch the MP4 video in Supplemental Digital Content 1, <http://links.lww.com/MD/G601>. This 5YSPBC app might be an example of SDM<sup>[8–10,55–57]</sup> used in numerous disciplines in hospital settings.

The CNN module on Microsoft Excel is unique and innovative. Users who are not familiar with the CNN software (e.g., Python) can apply our Excel-VBA module to conduct CNN-related research in the future if the sample size is not large (e.g., fewer than 10,000 cases). The module is not limited to binary classification. The multi-classification module can be done by adding the layers on CNN. Any other types of self-assessment, such as predicting the 14-day hospital readmission of patients with pneumonia,<sup>[13]</sup> predicting active NBA players most likely to be inducted into the basketball hall of Famers,<sup>[14]</sup> and screening BC,<sup>[7]</sup> can apply the CNN model to predict and classify the levels of harmfulness and disease in the future.

#### 4.5. Limitations and suggestions

Our study has some limitations. First, the psychometric properties of the variables have been validated for predicting the 5YSPBC after removing invalidate variables from the 53 original categories; however, no evidence can support the 15-variable 5YSPBC viable and validate its use in our predictive models though correlation coefficients of 5YSPBC have been examined before involving them into models. Hence, additional studies are recommended using other selection techniques to select suitable feature variables (i.e., variables correlated to the label of survival and non-survival reaches 0.1 or above<sup>[13]</sup>) in the future.

Second, any improvement variables used to effectively increase the predictive ACC were not discussed. For instance, whether other feature variables applied to the CNN 5YSPBC model can increase the ACC rate is worthy of further study. Future studies are encouraged to look for other feature variables that are not included in the database of cancer registration coded by professional clerks and can also improve the power of the model prediction of the 5YSPBC.

Third, the study was based on patients treated in one single hospital in Taiwan. All data were collected within the period from 2004 to 2019. If any environment or condition is changed (e.g., other countries and hospitals), the result (e.g., the model's parameters) must be different from the current study.

Fourth, based on the model ACC and prediction stability in Table 3, CNN was selected instead of KNN as app development. The reason is that KNN is severely affected by the SENS of the neighborhood size and simple majority voting in the regions of neighborhoods, especially in the case of the small sample size with the existing outliers.<sup>[57-59]</sup> The CNN is thus reliable and viable in this study.

Fifth, the sample size (=1180) is not big enough to increase the model prediction accuracy further. The sample size is required to increase in the future research, not limited to the BC only.

Sixth, the AUC is common as one of indicator to evaluate the model accuracy. The Phi  $\phi$  as the larger effect size ( $\geq 0.50$ ) for Chi-square test could be a supplemental alternative to the AUC ( $\geq 0.75$ ) in the future relevant studies; see the result of a small study in Supplemental Digital Content 2, <http://links.lww.com/MD/G602>.

Seventh, readers might raise doubts about why the treatment approach was not included as a feature variable to assess the 5-year survival (e.g., the choice of radiotherapy or chemotherapy drugs that may affect the survival rate significantly). Feature variables extracted from 53 items were determined by using the Weka software<sup>[31]</sup> (see the Section 2.2). Those variables (e.g., #28 to #33 shown in Supplemental Digital Content 1, <http://links.lww.com/MD/G601>) were not selected, implying that those treatment approaches are not relatively important to the 15 selected variables. It is worth validating the results further using the Weka software<sup>[31]</sup> or other tools in the future relevant studies.

Finally, the study examined the prediction ACC in 5 models (i.e., CNN, KNN, LR, NB, and ANN) across 2 different scenarios of training and testing sets. Numerous types of classification algorithms in ML are available. Whether another algorithm has higher ACC than CNN under an equal condition (e.g., 15 feature variables) is expected to be verified in future 5YSPBC-related research.

## 5. Conclusion

Features and contributions in this study were illustrated as follows: CNN performed in Microsoft Excel, online app demonstrated to display results using a visual dashboard on Google Maps, and category probability curves based on Rasch model observed in the CNN prediction model. The novelty of the app with the CNN algorithm improves the predictive ACC of the 5YSPBC. It is expected to help BC-case managers use the app to assess the 5YSPBC in hospital settings in the future.

## Acknowledgments

The authors thank Enago ([www.enago.tw](http://www.enago.tw)) for the English language review of this manuscript.

## Author contributions

Tsair-Wei Chien developed the study concept and design. Cheng-Yao Lin, Yen-Hsun Chen, Shih-Bin Su, and Yen-Ling Lee analyzed and interpreted the data. Shih-Bin Su monitored the process of this study and helped in responding to the reviewers' advice and comments. Tsair-Wei Chien drafted the manuscript, and all authors provided critical revisions for important intellectual content. The study was supervised by Shih-Bin Su. All authors read and approved the final manuscript.

**Conceptualization:** Cheng-Yao Lin, Yen-Hsun Chen.

**Data curation:** Yen-Ling Lee.

**Investigation:** Shih-Bin Su.

**Methodology:** Tsair-Wei Chien.

**Supervision:** Shih-Bin Su.

## References

- Chien TW. Topic with keywords breast cancer, prediction, and survival rate in Pubmed Central; 2020. Available at: <https://pubmed.ncbi.nlm.nih.gov/?term=%22breast+cancer%22%5Btitle%5D+and+predict%5Btitle%5D+and+%22survival+rate>. Accessed January 14, 2022.
- Ghoncheh M, Pournamdar Z, Salehiniya H. Incidence and mortality and epidemiology of breast cancer in the world. *Asian Pac J Cancer Prev* 2016;17:43–6.
- Tao L, Zhang C, Zeng L, et al. Accuracy and effects of clinical decision support systems integrated with BMJ best practice-aided diagnosis: interrupted time series study. *JMIR Med Inform* 2020;8:e16912.
- American Cancer Society. 50 Facts: Global health situation and trends 1955-2025; 2020. Available at: [https://www.who.int/whr/1998/media\\_centre/50facts/en/](https://www.who.int/whr/1998/media_centre/50facts/en/). Accessed January 14, 2022.
- Hamouda SKM, Wahed ME, Abo Alez RH, Riad K. Robust breast cancer prediction system based on rough set theory at Na-tional Cancer Institute of Egypt. *Comput Methods Programs Biomed* 2018;153:259–68.
- Domchek SM, Eisen A, Calzone K, Stopher J, Blackwood A, Weber BL. Application of breast cancer risk prediction models in clinical practice. *J Clin Oncol* 2003;21:593–601.



- [7] Agide FD, Sadeghi R, Garmaroudi G, Tigabu BM. A systematic review of health promotion interventions to increase breast cancer screening uptake: from the last 12 years. *Eur J Public Health* 2018;28:1149–55.
- [8] Butler AM, Elkins S, Kowalkowski M, Raphael JL. Shared decision making among parents of children with mental health conditions compared to children with chronic physical conditions. *Matern Child Health J* 2015;19:410–8.
- [9] Légaré F, Witteman HO. Shared decision making: examining key elements and barriers to adoption into routine clinical practice. *Health Aff (Millwood)* 2013;32:276–84.
- [10] Florin J, Ehrenberg A, Ehnfors M. Clinical decision-making: predictors of patient participation in nursing care. *J Clin Nurs* 2008;17:2935–44.
- [11] Chae SH, Kim Y, Lee KS, Park HS. Development and clinical evaluation of a web-based upper limb home rehabilitation system using a smartwatch and machine learning model for chronic stroke survivors: prospective comparative study. *JMIR Mhealth Uhealth* 2020;8:e17216.
- [12] Harvey H, Heindl A, Khara G, et al. Deep learning in breast cancer screening. *Artificial Intelligence in Medical Imaging Cham*: Springer; 2019. 187–215.
- [13] Tey S-F, Liu C-F, Chien T-W, et al. Predicting the 14-Day hospital readmission of patients with pneumonia using artificial neural networks (ANN). *Int J Environ Res Public Health* 2021;18:5110.
- [14] Chou PH, Chien TW, Yang TY, Yeh YT, Chou W, Yeh CH. Predicting active NBA players most likely to be inducted into the basketball hall of famers using artificial neural networks in microsoft excel: development and usability study. *Int J Environ Res Public Health* 2021;18:4256.
- [15] Lee YL, Chou W, Chien TW, Chou PH, Yeh YT, Lee HF. An app developed for detecting nurse burnouts using the convolutional neural networks in Microsoft Excel: population-based questionnaire study. *JMIR Med Inform* 2020;8:e16528.
- [16] Ma SC, Chou W, Chien TW, et al. An app for detecting bullying of nurses using convolutional neural networks and web-based computerized adaptive testing: development and usability study. *JMIR Mhealth Uhealth* 2020;8:e16747.
- [17] Yan YH, Chien TW, Yeh YT, Chou W, Hsing SC. An app for classifying personal mental illness at workplace using fit statistics and convolutional neural networks: survey-based quantitative study. *JMIR Mhealth Uhealth* 2020;8:e17857.
- [18] Tang X, Shu Y, Liu W, Li J, Liu M, Yu H. An optimized weighted Naïve Bayes method for flood risk assessment. *Risk Anal* 2021; doi: 10.1111/risa.13743. Online ahead of print.
- [19] Viana Dos Santos Santana I, CM da Silveira A, Sobrinho Á, et al. Classification models for COVID-19 test prioritization in Brazil: machine learning approach. *J Med Internet Res* 2021;23:e27293.
- [20] Golpour P, Ghayour-Mobarhan M, Saki A, et al. Comparison of support vector machine, Naïve Bayes and logistic regression for assessing the necessity for coronary angiography. *Int J Environ Res Public Health* 2020;17:6449.
- [21] Gholizadeh P, Esmaili B. Developing a multi-variate logistic regression model to analyze accident scenarios: case of electrical contractors. *Int J Environ Res Public Health* 2020;17:4852.
- [22] Nhu VH, Shirzadi A, Shahabi H, et al. Shallow landslide susceptibility mapping: a comparison between logistic model tree, logistic regression, Naïve Bayes tree, artificial neural network, and support vector machine algorithms. *Int J Environ Res Public Health* 2020;17:2749.
- [23] Choi Y, Boo Y. Comparing Logistic regression models with alternative machine learning methods to predict the risk of drug intoxication mortality. *Int J Environ Res Public Health* 2020;17:897.
- [24] Wu L, Deng F, Xie Z, et al. Spatial analysis of severe fever with thrombocytopenia syndrome virus in china using a geographically weighted logistic regression model. *Int J Environ Res Public Health* 2016;13:1125.
- [25] Yildirim P, Majnarić L, Ekmekci OI, Holzinger A. Knowledge discovery of drug data on the example of adverse reaction prediction. *BMC Bioinform* 2014;15(Suppl 6):S7; doi: 10.1186/1471-2105-15-S6-S7.
- [26] Tarekegn A, Ricceri F, Costa G, Ferracin E, Giacobini M. Predictive modeling for frailty conditions in elderly people: machine learning approaches. *JMIR Med Inform* 2020;8:e16678.
- [27] Ravi D, Wong C, Deligianni F, et al. Deep learning for health informatics. *IEEE J Biomed Health Inform* 2017;21:4–21.
- [28] Nair TM. Building and interpreting artificial neural network models for biological systems. *Methods Mol Biol* 2021;2190:185–94.
- [29] Sathyanarayana A, Joty S, Fernandez-Luque L, et al. Sleep quality prediction from wearable data using deep learning. *JMIR Mhealth Uhealth* 2016;4:e125.
- [30] Mitchell SJ, Godoy L, Shabazz K, Horn IB. Internet and mobile technology use among urban African American parents: survey study of a clinical population. *J Med Internet Res* 2014;16:e9.
- [31] Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Ian H. The WEKA data mining software: an update. *SIGKDD Explor* 2009;11:10–8.
- [32] Hamling J, Lee P, Weitkunat R, Ambühl M. Facilitating meta-analyses by deriving relative effect and precision estimates for alternative comparisons from a set of estimates presented by exposure level or disease category. *Stat Med* 2008;27:954–70.
- [33] Chen CJ, Wang LC, Kuo HT, Fang YC, Lee HF. Significant effects of late evening snack on liver functions in patients with liver cirrhosis: a meta-analysis of randomized controlled trials. *J Gastroenterol Hepatol* 2019;34:1143–52.
- [34] Lalkhen AG. *Statistics V: Introduction to clinical trials and systematic reviews*. Contin Educ Anaesth Crit Care Pain 2008;8:143–6. Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Chicago, US: University of Chicago Press; 1980. Accessed January 14, 2022.
- [35] Zhao Z, Peng H, Lan C, Zheng Y, Fang L, Li J. Imbalance learning for the prediction of N6-Methylation sites in mRNAs. *BMC Genom* 2018; 19:574.
- [36] Lei T, Sun H, Kang Y, et al. ADMET evaluation in drug discovery. 18. Reliable prediction of chemical-induced urinary tract toxicity by boosting machine learning approaches. *Mol Pharm* 2017;14: 3935–53.
- [37] Li L, Xiong Y, Zhang ZY, et al. Improved feature-based prediction of SNPs in human cytochrome P450 enzymes. *Interdiscip Sci* 2015;7:65–77.
- [38] Wang X, Jin Y, Zhang Q. DeepPred-SubMito: a novel submitochondrial localization predictor based on multi-channel convolutional neural network and dataset balancing treatment. *Int J Mol Sci* 2020;21:5710.
- [39] Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Chicago, IL: University of Chicago Press; 1980.
- [40] Andrich D. A rating formulation for ordered response categories. *Psychometrika* 1978;43:561–73.
- [41] Chien TW. Figure 2 shown in this study; 2021. Available at: <http://www.healthup.org.tw/gps/breastcancer15items.htm>. Accessed January 14, 2022.
- [42] Yan YH, Chien TW. The use of forest plot to identify article similarity and differences in characteristics between journals using medical subject headings terms: a protocol for bibliometric study. *Medicine (Baltimore)* 2021;100:e24610.
- [43] Chien TW. An app for predicting 5-year survival in patients with breast cancer; 2021. Available at: [http://www.healthup.org.tw/irs/irsin\\_e.asp?type1=93](http://www.healthup.org.tw/irs/irsin_e.asp?type1=93). Accessed January 14, 2022.
- [44] Al-Yousef A, Samarasinghe S. A novel computational approach for biomarker detection for gene expression-based computer-aided diagnostic systems for breast cancer. *Methods Mol Biol* 2021;2190:195–208.
- [45] Aličković E, Subasi A. Breast cancer diagnosis using GA feature selection and Rotation Forest. *Neural Comput Appl* 2017;28:753–63.
- [46] Yao N, Li W, Liu T, et al. Primary tumor removal improves the prognosis in patients with stage IV breast cancer: a population-based study (Cohort Study). *Int J Surg* 2020;83:109–14.
- [47] Baum M. Survival and reduction in mortality from breast cancer. Impact of mammographic screening is not clear. *BMJ* 2000;321:1470–2.
- [48] Peto R, Boreham J, Clarke M, Davies C, Beral V UK and USA breast cancer deaths down 25% in year 2000 at ages 20–69 years. *Lancet* 2000;355:1822.
- [49] Park JH, Anderson WF, Gail MH. Improvements in US breast cancer survival and proportion explained by tumor size and estrogen-receptor status. *J Clin Oncol* 2015;33:2870–6.
- [50] Richards MA, Stockton D, Babb P, Coleman MP. How many deaths have been avoided through improvements in cancer survival? *BMJ* 2000;320:895–8.
- [51] Mellors BOL, Spear AM, Howle CR, Curtis K, Macildowie S, Dehghani H. Machine learning utilising spectral derivative data improves cellular health classification through hyperspectral infra-red spectroscopy. *PLoS One* 2020;15:e0238647.
- [52] Singh VK, Maurya NS, Mani A, Yadav RS. Machine learning method using position-specific mutation based classification out-performs one hot coding for disease severity prediction in haemophilia 'A' [published online ahead of print, 2020 Sep 11]. *Genomics* 2020;S0888-7543: 30819–23.
- [53] Zhang PI, Hsu CC, Kao Y, et al. Real-time AI prediction for major adverse cardiac events in emergency department patients with chest pain. *Scand J Trauma Resusc Emerg Med* 2020;28:93.

- [54] Reh JF. Pareto's Principle - The 80-20 Rule, How the 80/20 rule can help you be more effective; 2020. Available at: [http://management.about.com/cs/generalmanagement/a/Pareto081202\\_2.htm](http://management.about.com/cs/generalmanagement/a/Pareto081202_2.htm). Accessed January 14, 2022.
- [55] Rawlings A, Sillin L, Shaddock P, et al. Informed consent: a shared decision-making process that creates a new professional obligation for care. *Surg Endosc* 2020;34:4713–6.
- [56] Landry JT. Is shared decision-making to blame for the provision of ethically inappropriate treatment? Results of a multi-site study exploring physician understanding of the “shared” model of decision making. *J Eval Clin Pract* 2021;27:826–35.
- [57] Zheng B, Wang X, Lederman D, Tan J, Gur D. Computer-aided detection; the effect of training databases on detection of subtle breast masses. *Acad Radiol* 2010;17:1401–8.
- [58] Gou J, Qiu W, Ou W. Locality constrained representation-based K-nearest neighbor classification. *Knowledge-Based Syst* 2019;167:38–52.
- [59] Gou J, Qiu W, Yi Z, Xu Y, Mao Q, Zhan YA. Local mean representation-based K-Nearest neighbor classifier. *ACM Trans Intell Syst Technol* 2019;10:1–25.