MEDINFO 2023 — The Future Is Accessible
J. Bichel-Findlay et al. (Eds.)

doi:10.3233/SHTI231116

Machine-Learning Based Risk Assessment for Cancer Therapy-Related Cardiac Adverse Events Among Breast Cancer Patients

Quynh T.N Nguyen[a]; Phuc T. Phan[b]; Shwu-Jiuan Lin[a]; Min-Huei Hsu[c]; Yu-Chuan (Jack) Li[d]; Jason C. Hsu[b,c,*]; and Phung-Anh Nguyen[c,1,*]
[a] *Ph.D. Program in School of Pharmacy, College of Pharmacy, Taipei Medical University, Taipei, Taiwan*
[b] *International Ph.D. Program in Biotech and Healthcare Management, College of Management, Taipei Medical University, Taipei, Taiwan*
[c] *Clinical Data Center, Office of Data Science, Taipei Medical University, Taipei, Taiwan*
[d] *Graduate Institute of Biomedical Informatics, College of Medical Science & Technology, Taipei Medical University, Taipei, Taiwan*
[*] *Equal contribution*
ORCID: Phung-Anh Nguyen https://orcid.org/my-orcid?orcid=0000-0002-7436-9041

Abstract. The study aims to develop machine-learning models to predict cardiac adverse events in female breast cancer patients who receive adjuvant therapy. We selected breast cancer patients from a retrospective dataset of the Taipei Medical University Clinical Research Database and Taiwan Cancer Registry between January 2004 and December 2020. Patients were monitored at the date of prescribed chemo- and/or -target therapies until cardiac adverse events occurred during a year. Variables were used, including demographics, comorbidities, medications, and lab values. Logistics regression (LR) and artificial neural network (ANN) were used. The performance of the algorithms was measured by the area under the receiver operating characteristic curve (AUC). In total, 1321 patients (an equal 15039 visits) were included. The best performance of the artificial neural network (ANN) model was achieved with the AUC, precision, recall, and F1-score of 0.89, 0.14, 0.82, and 0.2, respectively. The most important features were a pre-existing cardiac disease, tumor size, estrogen receptor (ER), human epidermal growth factor receptor 2 (HER2), cancer stage, and age at index date. Further research is necessary to determine the feasibility of applying the algorithm in the clinical setting and explore whether this tool could improve care and outcomes.

Keywords. Breast Cancer, Cardiotoxicity, Prediction Model, TMUCRD

1. Introduction

Breast cancer is the site of cancer that has the highest incidence rate regardless of gender, although it is much more common in females than males [1]. Nevertheless, the 5-year

[1] Corresponding Author: Phung-Anh Nguyen, Clinical Data Center, Office of Data Science, Taipei Medical University, Taipei, Taiwan; Email: alex0303@tmu.edu.tw;

survival rate of female breast cancer patients is relatively high (90%) [2]. The number of cancer survivors is increasing worldwide. They have become a fast-growing group of patients in the healthcare system of many countries [3]. Given that cardiovascular disease and cancer share several risk factors (such as diabetes, obesity, smoking, drinking) [4], the population of patients with cancer is facing a higher risk of cardiac problems.

Another big issue is that cancer treatment can also be associated with cardiac diseases. Cardiotoxicity of anticancer drugs has been detected since the 1960s, including anthracycline-induced heart failure and antimetabolites with a higher risk of myocardial infarction [5]. Cardiotoxicity may result in restricted drug indication and impact patients' quality of life. Many machine-learning studies have been conducted to predict this kind of adverse event. However, most of them are based on physiochemical properties (such as molecular structure, atom count, bond count descriptors), which are more appropriate for novel drug discovery and development [6].

In this study, we used clinical features to develop machine-learning models to predict cardiac adverse events in female breast cancer patients who receive adjuvant therapy.

2. Methods

2.1. Data Source

We conducted a retrospective study using data obtained from Taipei Medical University Clinical Research Database (TMUCRD) which is linked to Taiwan Cancer Registry (TCR). TCR database was established in 1979 and is managed by Taiwan Ministry of Health, covering 98% of Taiwanese cancer patients with their related information. TMUCRD retrieved data from three hospitals' electronic health records (EHRs). The database contains data of 3.8 million people from 1998 to 2020, including basic patient information, medical information, laboratory test, treatment process, and other unstructured data. This study has been approved by the Joint Institute Review Board of Taipei Medical University, and the data were anonymized before analysis.

2.2. Study Population

In this study, we selected female patients with a primary diagnosis of breast cancer (ICD-O-3 code: C50) who received chemotherapy or target therapy between 2004 and 2020. Exclusion criteria included individuals under 20 years old, patients with no medical history at the TMU hospital system, and patients who did not receive cancer treatment after they were diagnosed at the three affiliated hospitals.

2.3. Outcome Measurement

We ascertained the study outcome using TMUCRD EHR. The outcome was cardiac adverse events that consisted of myocardial infarction (MI), arrhythmia, conduction disorders, heart failure (HF), and coronary artery diseases (CAD). Index date was when patients were prescribed chemo and/or target therapies. The outcome was any cardiac adverse events occurred within one year after the index date. All outcomes were extracted using the International Classification of Diseases 9 and 10 (i.e., MI, 410, 412,

I21, I22, I25.2; Arrhythmia, 427, I46-I49; Conduction disorders, 426, I44, I45; HF, 402, 404, 425, 428, I09.9, I11.0, I13, I25.5, I42, I43, I50; CAD, 414, I25) in outpatient clinics and hospital admission.

2.4. Features

In this study, we included several clinical features; 1) demographics such as age, body mass index (BMI), smoking, and drinking status; 2) cancer conditions such as cancer stage and tumor size; 3) comorbidities such as hypertension, hyperlipidemia, diabetes, renal disease, liver disease, chronic pulmonary disease, cerebrovascular disease, and pre-existing cardiac disease; 4) concurrent medications such as metformin and other biguanides, HMG CoA reductase inhibitors (statins), beta-blockers, angiotensin II receptor blockers (ARBs), calcium channel blockers (CCBs), antiplatelets, benzodiazepines; 5) laboratory test: HER2 (human epidermal growth factor receptor 2), PR (progesterone receptor), ER (estrogen receptor).

Those comorbidities were considered if they were diagnosed within a year before the index date. We measured patients who had used medications by receiving them for more than a month in a year before the index date.

2.5. Development of the Algorithms

This study used logistic regression (LR) as the basic model compared with artificial neural network (ANN) as the novel machine-learning algorithm. The ANN model is a learning algorithm vaguely inspired by biological neural networks. Computations are structured in terms of an interconnected group of artificial neurons, and these neurons process information using a connectionist approach to analysis. The number of hidden layers with the number of neurons in each layer was set at 3 and 16, respectively. Additionally, the l2 regularization of .01 and the relu activation were used for each layer. We set the softmax activation for the output layer. We used the Adam optimizer, a highly performant stochastic gradient descent algorithm, and binary_crossentropy as the binary classification outcome for the loss function.

2.6. Evaluating the Algorithms

The training dataset contained the data of patients from Taipei Medical University and Wan-Fang hospitals. The stratified 5-fold cross-validation was applied in the training dataset to assess the different machine-learning models' performance and general errors. We recruited data from Shuang-ho hospital and used it for the external testing dataset to generalize the model.

We analyzed the feature's contribution to the best model using SHAP values (Shapley Additive explanation). The performance of the algorithms was measured by the area under the receiver operating characteristic curve (AUC), accuracy, sensitivity (recall), specificity, positive predictive value (PPV), negative predictive value (NPV), and F1-score. All the processing was performed using the MSSQL server 2017, and the model training and testing were performed using Python version 3.8 with Scikit-learn version 1.1 and TensorFlow version 2.0.

3. Results

3.1. Baseline Characteristics of Patients

We identified 6,464 eligible female breast cancer patients diagnosed for the first time and registered at the TCR. We excluded 3,039 patients, including those under 20 years, patients without treatment history at TMUCRD before the cancer diagnosis, and patients who did not receive any cancer therapies at three hospitals. Afterward, only 1,321 patients who underwent chemotherapies or target therapies (i.e., an equal 15,039 visits consisted of 7,642 visits in the training set, whereas 7,451 visits in the testing set) were included in the development and validation of the models (Figure 1).

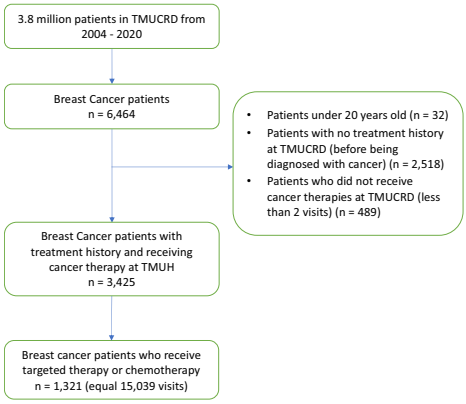

Figure 1. Overview study population

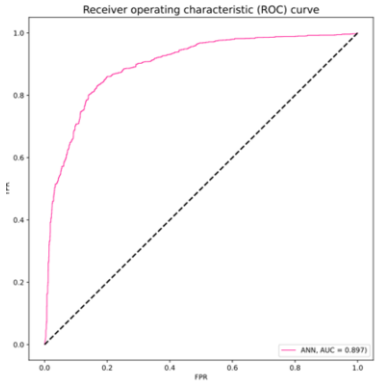

Figure 2. The performance of the ANN model in the testing dataset

The baseline demographic characteristics of patients are described in Table 1. The mean (standard deviation, SD) age and BMI of the cohort were 56.0 (11.3) and 24.4 (4.14), respectively. Most of the patients were in the early stages of breast cancer: stage I (27.5%), stage II (50.0%), and less likely to smoke (6.6%) or drink (5.1%). Almost all of them had surgery (92.3%) and/or radiation (63.2%) before receiving anticancer drugs.

Table 1. Baseline demographic characteristics of the study population

Variables	Overall (n=1,321)	Training (n=738)	Testing (n=583)
Age, Mean (SD), years	56.0 (11.3)	56.4 (11.8)	55.6 (10.6)
BMI, Mean (SD), kg/m2	24.4 (4.14)	24.1 (4.00)	24.8 (4.28)
Smoking, N (%)	87 (6.6)	43 (5.8)	44 (7.5)
Drinking, N (%)	68 (5.1)	46 (6.2)	22 (3.8)
Tumor size, Mean (SD), cm	29.7 (20.1)	29.6 (20.9)	29.7 (19.1)
Cancer stage, N (%)			
0 - 2	1,070 (81)	634 (85.8)	436 (74.8)
3 - 4	178 (13.5)	95 (12.8)	83 (14.2)
HER2, N (%)	357 (27.0)	193 (26.2)	164 (28.1)
PR, N (%)	864 (65.4)	471 (63.8)	393 (67.4)
ER, N (%)	962 (72.8)	546 (74.0)	416 (71.4)
Radiation therapy, N (%)	835 (63.2)	389 (52.7)	446 (76.5)
Surgery, N (%)	1219 (92.3)	677 (91.7)	542 (93.0)
Cardiac Outcomes, N (%)	112 (8.5)	75 (10.2)	37 (6.3)

3.2. The Performances of Prediction Models

The performance of the LR model was observed with an AUC of 0.65, a precision of 0.09, a recall of 0.82, and an F1-score of 0.17. We found better performance with an AUC of 0.897 for the ANN model (i.e., precision, 0.14; recall, 0.83; and F1-score, 0.2). Figure 2 shows the ROC curves of the ANN model.

4. Discussion

In this study, we built predictive ML models for cardiac risk assessment among overall cardiovascular outcomes. Oncologists referred these patients to cardio-oncology services based on the professional assessment of clinical factors such as pre-existing cardiac disease, tumor size, ER, HER2, cancer stage, and age at index date. These highly predictive models offer potential approaches for cardio-oncology clinical practice.

There are limitations to this study. First, although the study used data from various clinical settings (e.g., TMUH and WFH for establishing the prediction model and SHH for conducting an external test) located in the north of Taiwan, the results may not directly apply to breast cancer patients in other regions. Future studies may need to consider validating the model using data from other areas. Second, this study used retrospective data for development and validation. Further experiments with a prospective study design in clinical settings are needed.

5. Conclusions

In summary, we designed an artificial neural network model with high AUC, precision, and recall to observe the expected cardiotoxicity of breast patients during a year. Further research is necessary to determine the feasibility of applying the algorithm in the clinical setting and explore whether this tool could improve care and outcomes.

References

[1] Sung, H.; Ferlay, J.; Siegel, R. L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021, 71 (3), 209–249. https://doi.org/10.3322/caac.21660.

[2] Siegel, R. L.; Miller, K. D.; Fuchs, H. E.; Jemal, A. Cancer Statistics, 2022. CA Cancer J Clin 2022, 72 (1), 7–33. https://doi.org/10.3322/CAAC.21708.

[3] Cancer survival statistics | Cancer Research UK. https://www.cancerresearchuk.org/health-professional/cancer-statistics/survival#heading-Zero (accessed 2022-11-25).

[4] Koene, R. J.; Prizment, A. E.; Blaes, A.; Konety, S. H. Shared Risk Factors in Cardiovascular Disease and Cancer. Circulation 2016, 133 (11), 1104–1114. https://doi.org/10.1161/CIRCULATIONAHA.115.020406.

[5] Levis, B. E.; Binkley, P. F.; Shapiro, C. L. Cardiotoxic Effects of Anthracycline-Based Therapy: What Is the Evidence and What Are the Potential Harms? Lancet Oncol 2017, 18 (8), e445–e456. https://doi.org/10.1016/S1470-2045(17)30535-1.

[6] Yang, M.; Tao, B.; Chen, C.; Jia, W.; Sun, S.; Zhang, T.; Wang, X. Machine Learning Models Based on Molecular Fingerprints and an Extreme Gradient Boosting Method Lead to the Discovery of JAK2 Inhibitors. *J Chem Inf Model* 2019, *59* (12), 5002–5012. https://doi.org/10.1021/ACS.JCIM.9B00798/SUPPL_FILE/CI9B00798_SI_001.PDF.