APAMI2020 Poster Presentation Sessions | APAMI 2020 | Poster Presentation Sessions

Artificial Intelligence

Sun. Nov 22, 2020 3:00 PM - 4:00 PM Room E-2 (Congress center 5F - Conference Room 53)

[AP2-E2-4-06] A Deep Learning Model for Improved Breast Cancer Risk Prediction

*Md. Mohaimenul Islam[1,2], Hsuan-Chia Yang[1,2], Phung-Anh Nguyen[1,2], Yu-Hsiang Wang[3], Tahmina Nasrin Poly[1,2], Yu-Chuan (Jack) Li[1,2] (1. Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taiwan, 2. International Center for Health Information Technology, Taipei Medical University, Taiwan, 3. College of Medicine, Taipei Medical University, Taiwan)

Keywords: Breast Cancer, Deep Learning, Convolutional Neural Network

Breast cancer (BC) is the most commonly diagnosed cancer and the leading cause of cancer death among females. We aimed to develop a deep learning (DL)-based risk stratification system to predict BC patients one early using minimal features from electronic health records. We identified 8,606 patients who underwent a diagnosis of BC between January 1999 and December 2013. The CNN model was developed to predict BC one-year earlier. CNN model demonstrated great performance in predicting BC cancer. For the prediction of BC one year earlier, the areas under the receiver operating characteristic curve was 0.918. The sensitivity, specificity, and positive predictive value were 0.816, 0.848, and 0.541, respectively. The CNN model based on variables available in EHR can be a promising tool to distinguish patients at risk of BC.

A Deep Learning Model for Improved Breast Cancer Risk Prediction

Md. Mohaimenul Islam[a, b], Hsuan-Chia Yang[a, b], Phung-Anh Nguyen[a, b], Yu-Hsiang Wang[c], Tahmina Nasrin Poly[a, b] and Yu-Chuan (Jack) Li[a, b]

[a] *Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taiwan*
[b] *International Center for Health Information Technology, Taipei Medical University, Taiwan*
[c] *College of Medicine, Taipei Medical University, Taiwan*

Abstract

Breast cancer (BC) is the most commonly diagnosed cancer and the leading cause of cancer death among females. We aimed to develop a deep learning (DL)-based risk stratification system to predict BC patients one early using minimal features from electronic health records. We identified 8,606 patients who underwent a diagnosis of BC between January 1999 and December 2013. The CNN model was developed to predict BC one-year earlier. CNN model demonstrated great performance in predicting BC cancer. For the prediction of BC one year earlier, the areas under the receiver operating characteristic curve was 0.918. The sensitivity, specificity, and positive predictive value were 0.816, 0.848, and 0.541, respectively. The CNN model based on variables available in EHR can be a promising tool to distinguish patients at risk of BC.

Keywords:

Breast Cancer, Deep Learning, Convolutional Neural Network

Introduction

Breast cancer (BC) is the most commonly diagnosed cancer and the leading cause of cancer death among females [1]. Based on the report of Global Health Estimates, WHO 2013, approximately 508,000 women died in 2011 due to BC globally [2]. The prevalence rates of BC vary greatly globally from 19.3 per 100,000 women in Eastern Africa to 89.7 per 100,000 women in Western Europe [3-4]. The overall survival rate of BC is ranging from 80% or over in North America, Sweden, and Japan to around 60% in middle-income countries and below 40% in low-income countries. The higher rate of incidence and mortality of BC has prompted efforts in the risk stratification of BC patients based on available clinical features [5-6]. The currently available risk scores have several shortcomings including lack of variables, low performances that are the main hindrances of utilization in everyday clinical practice. A tool that could able to classify patients more correctly and precisely is required for clinical practice [7-8]. The recent improvements in the deep learning algorithms are considered as promising tools to meet this compelling demand. We aimed to develop a deep learning (DL)-based risk stratification system to predict BC patients one early using minimal features from electronic health records.

Methods

We identified 8,606 patients who underwent a diagnosis of BC between January 1999 and December 2013. For each of these patients, clinical characteristics such as demographics, medical history, and laboratory parameters were extracted retrospectively from electronic medical records and entered our structured database. Follow-up data [status (BC), date of BC diagnosis] was obtained for all patients from the Taiwan National Health Insurance database. Patients with less than 1-year follow-up duration were excluded from all analyses. The primary endpoint of our study was to predict BC patients one year earlier. The task of CNN was to predict the probability of BC of each patient based on clinical features. We used a 5-fold cross-validation technique to reduce prediction bias. To quantify the model discriminative capability, receiver operating characteristic (ROC) curve analysis was performed and an area under the curve (AUC) was calculated.

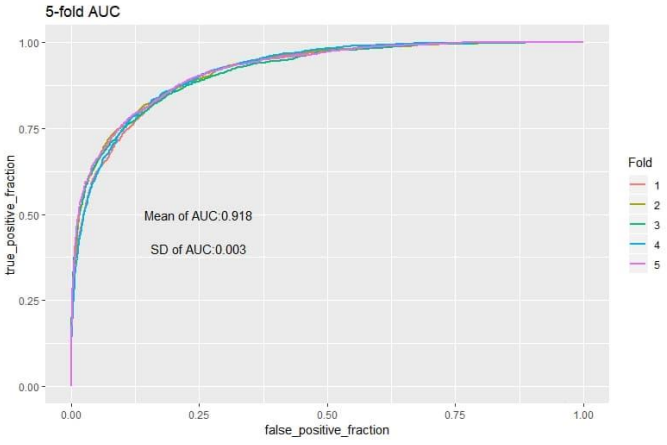

Figure 1- Receiver operating characteristic curve analysis of breast cancer risk.

Results

The final dataset contained a total of 8,606 cases and 34,424 controls. CNN model demonstrated great performance in predicting BC cancer. For the prediction of BC one year earlier, the areas under the receiver operating characteristic curve was 0.918 (Figure 1). The sensitivity, specificity, and positive predictive value were 0.816, 0.848, and 0.541, respectively. The

discriminative ability of our model was superior to other existing prediction models.

Discussion and conclusion

In this study, we applied a deep learning approach to clinical features, including medication history, and comorbidities, to developed a prediction model to stratify BC patients one year earlier. Our CNN-based risk assessment tool outperformed previous findings. By capturing variables from EHR, the CNN model effectively stratify patients at high risk and low risk. Therefore, the integration of the CNN model to daily clinical practice may facilitate optimal candidate selection and prognostication of BC patients.

References

[1] Ghoncheh M, Pournamdar Z, Salehiniya H. Incidence and mortality and epidemiology of breast cancer in the world. *Asian Pac J Cancer Prev*. 2016; 17(S3):43-6.

[2] Akinyemiju TF. Socio-economic and health access determinants of breast and cervical cancer screening in low-income countries: analysis of the World Health Survey. *PLoS One*. 2012; 7(11):e48834.

[3] Parkin DM, Fernández LM. Use of statistics to assess the global burden of breast cancer. *The breast journal*. 2006; 12:S70-S80.

[4] Ferlay J, Héry C, Autier P, Sankaranarayanan R. Global burden of breast cancer. Breast cancer epidemiology: Springer; 2010. p. 1-19.

[5] Azubuike SO, Muirhead C, Hayes L, McNally R. Rising global burden of breast cancer: the case of sub-Saharan Africa (with emphasis on Nigeria) and implications for regional development: a review. *World J Surg Oncol*. 2018; 16(1):1-13.

[6] Ginsburg O, Bray F, Coleman MP, Vanderpuye V, Eniu A, Kotha SR, *et al*. The global burden of women's cancers: a grand challenge in global health. *The Lancet*. 2017; 389(10071):847-60.

[7] Chang S-W, Kareem SA, Kallarakkal TG, Merican A, Abraham MT, Zain RB. Feature selection methods for optimizing clinicopathologic input variables in oral cancer prognosis. *Asian Pac J Cancer Prev*. 2011; 12:2659-64.

[8] Ganesan N, Venkatesh K, Rama M, Palani AM. Application of neural networks in diagnosing cancer disease using demographic data. *International Journal of Computer Applications*. 2010; 1(26):76-85.

Address for correspondence

Yu-Chuan (Jack) Li, MD, PhD
Professor, Graduate Institute of Biomedical Informatics
College of Medicine Science and Technology
Taipei Medical University, Taiwan
International Center for Health Information Technology (ICHIT)
Taipei Medical University, Taiwan
Chair, Department of Dermatology, Wan Fang Hospital, Taiwan
250 Wu-Hsing St., Taipei 110, Taiwan
Email: jack@tmu.edu.tw, jaak88@gmail.com
Phone: +886-2-27361661 ext.7600
Fax: +886-2-6638-7537