A Population-based Study from Electronic Health Records on the Comorbidities of Dementia Older Adults in Hong Kong

J.P. Zhong, N.X. Dong, L.H. Chen, Kevin K.F. Yuen, Arnold Y.L. Wong, Sam C.C. Chan

The Hong Kong Polytechnic University
Kowloon, Hong Kong SAR
{Joni.zhong, pearl.chen, kevin.yuen, arnold.wong, samcc.chan} @polyu.edu.hk
nanxi.dong@connect.polyu.hk

***Abstract*—Dementia is a condition characterized by a group of symptoms that affect memory, learning, and cognitive function. It negatively impacts patients' daily functional skills and independence, particularly in the later stages of the disease when comorbidities are often present. With the global aging problem leading to an increasing number of dementia cases, there is a significant burden on healthcare systems worldwide. Despite growing research on risk mitigation, early diagnosis, and intervention of dementia, few studies have focused on the developmental trajectory of the disease. In this study, we utilized the Hospital Authority (HA) Electronic Clinical Record to analyze structured clinical data of dementia patients in Hong Kong. Using mixed methods, we created a population-based case-control cohort to determine the association between dementia and other disease diagnoses based on ICD-10 codes. We identified significant associations with comorbidities in the categories of "Endocrine, nutritional and metabolic diseases" and "Mental, Behavioral and Neurodevelopmental disorders". Further, we plan to employ machine learning models to predict patient comorbidities data and understand the complex short-term and long-term dependencies in dementia progression.**

Keywords-dementia; comorbidities;clinical data;machine learning.

I. INTRODUCTION

Dementia includes a group of symptoms affecting memory, learning, and at least one other cognitive domain (such as aphasia, apraxia, agnosia, or executive function), which may be part and all impaired in different stages of the disease. Dementia negatively affects the patient's daily functional skills and independence, especially when they enter the later stages with comorbidities. Therefore, the increasing number of dementia with the worldwide ageing problem results in a significant burden for general health care in many countries, especially since several comorbidities are usually with dementia. There is growing research focus on risk mitigation, early diagnosis and intervention of dementia, but few have been conducted to determine the developmental trajectory of dementia (For review, please see [1]).

II. MOTIVATION

To better understand the comorbidity of dementia in the Hong Kong ageing population, a population-based study was conducted to analyze the structured clinical data of dementia patients using the real world clinical data. Our objective is to determine whether we can discover the higher risk of comorbidity of dementia through their dementia journey, since the patients were diagnosed with dementia. To achieve this, we utilized the Hospital Authority (HA) Electronic Clinical Record with a sample of 200,000 patients' digital files in the years of 2007 to 2017.

III. METHODOLOGY AND INTERIM RESULTS

The current study used mixed method to conduct the data analysis. First, our objective in this stage was to determine which disease diagnosis (in the format of structured ICD-10 codes) had a significance association with the diagnosis of dementia. We created a population-based case-control cohort of patients based on the presence (cases) and absence of dementia (controls) during this period in the database. We then filtered the data with the diagnosis of dementia in any year using The International Classification of Diseases, Tenth Version (ICD-10) codes (F00-F03). The related clinical data can be extracted, including their comorbidities throughout their disease journey. As in the control group, a set of random samples of individuals was matched by those based on the age of diagnosis of dementia and gender with a 1:1 ratio. The dependent variables are the comorbidity in the format of ICD-10 coding. By going through the categories of A to Z in ICD-10, we identified the essential dependent variables (i.e. comorbidity) that had a significant association with dementia. The significant association is tested using Chi-Square Test. We found that significant associations exist in "Endocrine, nutritional and metabolic diseases (Category E)" and "Mental, Behavioral and Neurodevelopmental disorders (category F)" in the Hong Kong population.

After the major categories of comorbidity with significant association with dementia have been identified at the first stage, we will further employ machine learning models to learn the development of comorbidity after the diagnosis of dementia. In this part, we will build and validate predictive machine learning models to predict patient comorbidities data (Fig 1). We will primarily focus on the multivariate and temporal nature of patient comorbidities data. Using this data-driven method [2], which can build relationships underlying all the available factors, we aim to create inherent relationships at multiple levels and scales, including 1) the short-term scale recording, indicating the correlations between consecutive medical records during one clinical episode within a short timeframe (e.g., a few days) and clinical events such as necessary lab tests involving a

 ISBN: 978-1-68558-080-3

wide range of vital sign measurements or medical notes indicating the progression over a short period of time; and 2) the long-term dependence, which usually happens among different clinical episodes and indicates the trend in patients' progression of comorbidities. Such trends also rely on the summary of each clinical episode. The integration of short and long-term dependence on clinical data is required to provide more precise predictions. For example, the trend of a patient's vital signs during the previous 24 hours on the prognosis, some of the comorbidities (e.g., diabetes, cardiovascular disease, stroke, hip fracture, etc.) and the demographic data (e.g., older age, male gender, or lower socioeconomic position) may indicate a higher probability of mortality [3]. In this view, a predictive model should follow and capture the complex short-term changes of clinical events and their influences over time (i.e., long-term dependencies) to understand the context of dementia progression (e.g., a hip fracture incident and its recovery influence on the dementia progression).

REFERENCES

[1] Melis, René JF, Miriam L. Haaksma, and Graciela Muniz-Terrera. "Understanding and predicting the longitudinal course of dementia." Current opinion in psychiatry 32.2 (2019): 123.

[2] Liu, Yuxi, Zhenhao Zhang, Antonio Jimeno Yepes, and Flora D. Salim. "Modeling long-term dependencies and short-term correlations in patient journey data with temporal attention networks for health prediction." In Proceedings of the 13th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, pp. 1-10. 2022.

[3] Rajamaki, Blair, Sirpa Hartikainen, and Anna-Maija Tolppanen. "The effect of comorbidities on survival in persons with Alzheimer's disease: a matched cohort study." BMC geriatrics 21 (2021): 1-9.

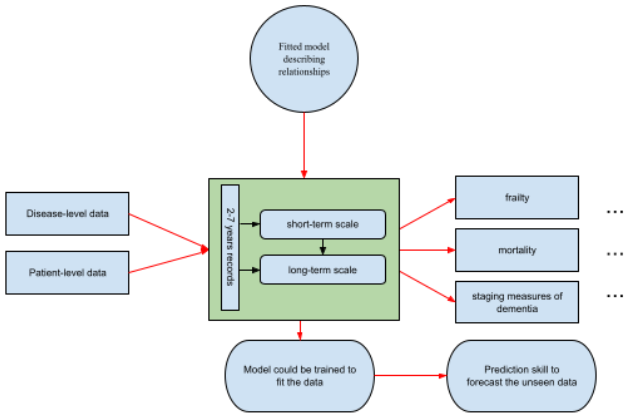

Figure 1. Input and Output of the Machine Learning Model