Developing a risk prediction model for heart failure (HF) incidence in patients with Diabetes Mellitus using AI & ML techniques

Verma V, Brooks L, Krebs B, Markan R, Sharma S ,Kumar S, Musle A, Nayyar A , Gaur A, Kukreja I, Bhargava S

Optum Life Sciences

Introduction

- Heart Failure (HF) is a progressive condition where heart muscles loses ability to pump enough blood to meet body's requirements. With time there has been substantial progress in the treatment paradigm of different forms of heart disease, but HF remains an ever-growing problem in the US. Estimates suggests that ~6.5 million[1] adults in the US have HF. Heart disease is the leading cause of death (697,000[2] deaths in 2020) across all ethnic groups in the US, accounting for 1 in every 5 deaths[2]. The overall economic burden for heart disease was estimated to be $229 billion[2] in 2018, including cost of health care services, medicines & lost productivity due to death.

- Studies suggest that clinical symptoms of HF patients are subtle and progress slowly. Integrating EHR and claim data will help providers with access to patient information in early disease identification and timely intervention.

Objective

Developing a risk prediction model for Heart Failure incidence in patients with Diabetes Mellitus using AI & ML techniques.

Methods

- A retrospective study using Optum® de-identified Market Clarity Dataset (linked claims and electronic health records or EHR of patients) was done among adults (>18 years) in the US.

- Based on ICD-9 and ICD-10 codes, 165,963 diabetics, aged 18 years and above were identified using Optum's Market Clarity Database. Patients with continuous eligibility for 5 years (2015 to 2019), having at least one yearly claim for Diabetes were included in the analysis.

- Of the overall cohort, 23,104 diabetic patients were found to have 2 outpatient claims/encounters at least 60 days apart OR at least one inpatient diagnosis for HF diagnosis recorded between Jan 01, 2019, and Dec 31, 2019, and not having any claim/encounter history for HF in 5-year Pre index period were selected.

- A multivariate analysis was conducted to develop a prediction model by incorporating 28 variables as comorbidities and Sign & Symptoms during pre-index period using feature selection techniques.

- Training and evaluation of Logistic Regression, XGBoost and Random Forest Classifier were executed.

- The models were trained and tested using 80:20 ratio of total subjects.
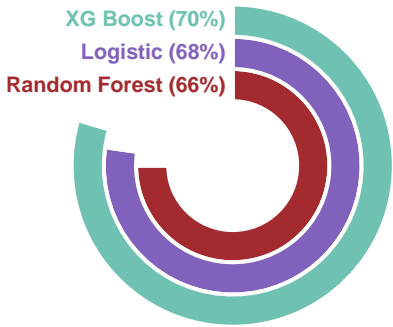
Results

Fig 1: AUROC Score

XG Boost (70%)
Logistic (68%)
Random Forest (66%)

Fig 2: Logistic

Specificity (86%)
Accuracy (73%)
Sensitivity (50%)

Fig 3: Odds Ratio

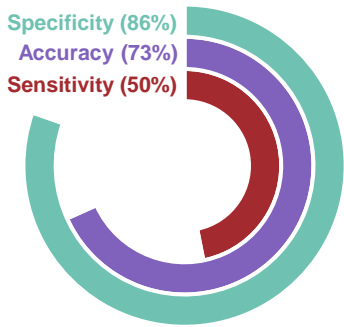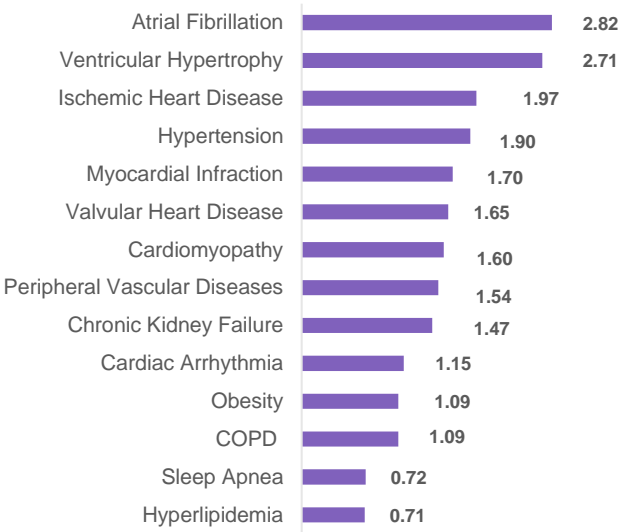|	
Atrial Fibrillation	2.82
Ventricular Hypertrophy	2.71
Ischemic Heart Disease	1.97
Hypertension	1.90
Myocardial Infraction	1.70
Valvular Heart Disease	1.65
Cardiomyopathy	1.60
Peripheral Vascular Diseases	1.54
Chronic Kidney Failure	1.47
Cardiac Arrhythmia	1.15
Obesity	1.09
COPD	1.09
Sleep Apnea	0.72
Hyperlipidemia	0.71

- AUROC for Logistic Regression, XGBoost and Random Forest Classifier were 68%, 70% and 66%, respectively.(Fig 1)

- Logistic regression identified Heart Failure cases with 66% precision as compared to 76% in non-HF control group.

- Overall, Logistic Regression model exhibited 73% accuracy and 86% specificity. (Fig 2)

- Odds ratio (OR) indicates that the probability of diabetic patient getting Heart Failure with risk factors like Atrial fibrillation (2.82), Ventricular Hypertrophy (2.71), Hypertension (1.90), Ischemic Heart Disease (1.97), Valvular Heart Disease (1.65), Peripheral Vascular Diseases (1.54), Chronic Kidney Failure (1.47) and COPD (1.09) were higher.

Conclusion

- Integrating clinical data (EHR data) with claim data helps in identifying a greater number of patients during early phases of diagnosis as compared to use of claim data only.

- With integrated patient data (Claim + EHR) at disposal, a provider can leverage it to conciliate risk factors in individuals during early phases of disease thus helps in avoiding or delaying heart failure condition.

References:1 https://hfsa.org/patient
2. https://www.cdc.gov/heartdisease/facts.htm

Optum