Prediction and external validation of Heart Failure in patients with Type 2 Diabetes Mellitus

Ross D. Williams, MSc[1], Jenna Reps, PhD[2], Peter R. Rijnbeek, PhD[1]
[1]Dept of Medical Informatics, Erasmus University Medical Centre, Rotterdam, The Netherlands; [2]Janssen Research and Development, Raritan, NJ, USA

Background

Type 2 Diabetes Mellitus (T2DM) and Heart Failure (HF) are two major and growing contributors to the global burden of disease both in developed and developing countries.

Patients with diabetes are much more likely to develop congestive heart failure than patients without diabetes (incidence rate 30.9 vs. 12.4 cases per 1,000 person-years)[1]. Heart failure and peripheral arterial disease are the most common initial manifestations of cardiovascular diseases in patient with T2DM [2].

Current best practice for diabetes treatment indicates a stratification based upon incidence of heart failure. For example, Metformin is contraindicated in patients with significant HF, Thiazolidinediones should be used very cautiously in those with or at risk for HF. Patients with diabetes and heart failure may benefit most from glucose-lowering therapies with SGLT2 inhibition [3]. In contrast, no improvement in heart failure, or potential detrimental effects, have been reported for glucose-lowering strategies that directly or indirectly increase the availability of insulin. Meta-analyses showed that intensive glucose lowering is not associated with any significant reduction in cardiovascular risk but results in a significant increase in heart failure risk.

From this it is clear that there is a need to predict incident HF in patients with T2DM.

Methods

In consultation with endocrinologists and cardiologists we developed two definitions of T2DM and three of HF.

Diabetes	
Broad	Narrow
• First recorded condition of T2DM At least 365 days of observation time prior to the index date • No diagnosis of HF preceding the index date	• First recorded condition of T2DM • At least 365 days of observation time prior to the index date • No diagnosis of HF preceding the index date • No Secondary Diabetes before the index date • No Type I Diabetes before the index date

Heart Failure		
All	Heart failure with preserved ejection fraction (HFpEF)	Heart failure with reduced ejection fraction (HFrEF)
This outcome cohort is defined as any occurrence of heart failure	Heart failure where the muscles in the left ventricle fail to relax properly. This means that it fails to fill with blood effectively after contraction	Heart failure where the muscles in the left ventricle fail to contract properly. This means that less oxygenated blood is expelled upon contraction.

Using these definitions we trained both Lasso Logistic regression and Gradient Boosting Machines (GBM) to predict outcomes of HF from a target cohort of T2DM. These were developed in one of three databases and then externally validated in the other two databases. In this way we get information about the best performing models and also information on their robustness and transportability.

Database	T2DM Broad	T2DM Narrow	HF	HFpEF	HFrEF
CCAE	2690798	2458483	1133125	55281	96292
MDCD	469497	415376	853203	81056	102613
MDCR	784071	693412	1634606	158085	185404

Table 1. Overview of the individual cohort sizes in the three databases (January 23th 2018)

Contact: r.williams@erasmusmc.nl

Results

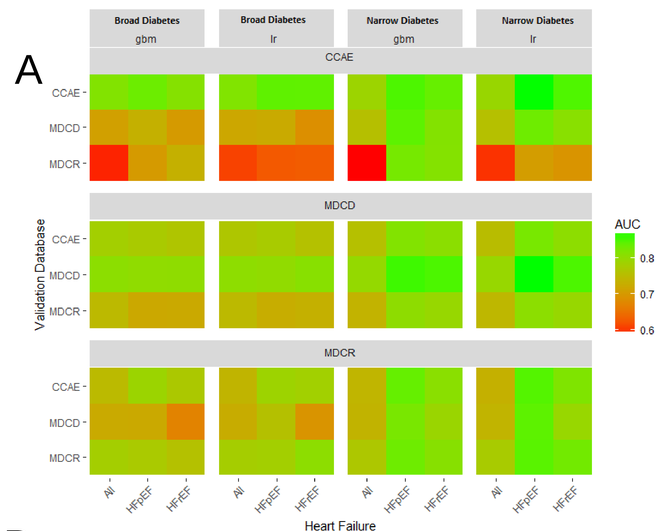

Figure A shows the area under the curve (AUC) achieved in the test set as well as in the external databases. Figure B shows the ROC curves for the Broad-All training for both algorithms in each of the three databases. As can be seen there is a stable performance across the models and databases.

The results show that the algorithms perform similarly and that generally the AUC is preserved when validated externally. There is also acceptable calibration of the models, although this is not shown here.

Conclusions

We leveraged the PatientLevelPrediction package to train, test and externally validate GBMs and LR for 12 different predictions of HF in T2DM. The results generally show good performance as well as good transportability of the models. Where this is not the case it tends to occur for all the models in a specific database. This suggests there is something about this database that is causing the models performance to drop off.

This study has been designed to be performed on the OHDSI network so we hope that other researchers will perform further external validations of the developed models on their own datasets.

References

1. Nichols GA, Hillier TA, Erbey JR, Brown JB. Congestive heart failure in type 2 diabetes: prevalence, incidence, and risk factors. Diabetes Care. 2001;24(9):1614-9.
2. Dinesh Shah A, Langenberg C, Rapsomaniki E, Denaxas S, Pujades-Rodriguez M, Gale CP, et al. Type 2 diabetes and incidence of a wide range of cardiovascular diseases: a cohort study in 1.9 million people. Lancet. 2015;385 Suppl 1:S86.
3. Lehrke M, Marx N. New antidiabetic therapies: innovative strategies for an old problem. Curr Opin Lipidol. 2012;23(6):569-75.