Ibeas, Jose; Suppi Boldrito, Remo; Caravaca Rodríguez, Miriam; [et al.]. «Artificial Intelligence for predicting heart failure in chronic kidney disease : analysis of a 5000-patient cohort». Nephrology Dialysis Transplantation, Vol. 39, sup. 1 (May 2024), art. gfae06906722277. DOI 10.1093/ndt/gfae069.672

Artificial Intelligence for Predicting Heart Failure in Chronic Kidney Disease: Analysis of a 5000-Patient Cohort

Authors: Òscar Gallés[1], Remo Suppi[2], Miriam Caravaca[1], Jordi Comas[3], Elisenda Matinez[4], Tomas Salas[4], Miriam Carles[1], Andres Soto [1], José Ibeas[1,5]

1. Clinical, Interventional and Computational Nephrology group (CICN) – Parc Taulí Hospital Universitari. Institut d'Investigació i Innovació Parc Taulí (I3PT - CERCA), Sabadell, Spain
2. Universitat Autònoma de Barcelona, School of Engineering, Spain
3. The Register of Renal Patients of Catalonia (RMRC) - Catalan Transplant Organization, OCATT, Spain
4. Catalan Agency for Health Quality and Evaluation (AQuAS), Spain
5. Nephrology Department, Parc Taulí University Hospital, Spain

Keywords: Chronic kidney disease; heart failure; cardiovascular disease, prediction models; artificial intelligence; machine learning.

Background and Aims:

Chronic kidney disease (CKD) and heart failure share a complex, bidirectional interaction, and are therefore an area of crucial interest in nephrology due to their prevalence and effect on mortality. It is basic to identify the risk of heart failure in patients with CKD as soon as possible to improve treatment outcomes and complications. It is important to note that while there is wide experience in prediction of heart failure, the use of models based on artificial intelligence is practically non-existent with this objective in mind. Based on the experience of our group in the prediction of mortality in patients with CKD and in CKD progression with machine learning algorithms, this study aims to predict the onset of heart failure in CKD patients. This study addresses this issue using machine learning techniques to offer new opportunities for accurate prediction of cardiac disease. This approach facilitates the individualization of risk assessment and allows the implementation of early and personalized interventions that can significantly influence the management of both diseases.

Method:

Design: Retrospective observational study of a historical cohort from the Registry of Renal Patients of Catalonia (RMRC) and the Catalan Agency for Health Quality and Evaluation.

Sample: 5000 patients with CKD, selected for having the minimum number of missing data.

Follow-up: 10 years, from January 2010 to December 2020.

Inclusion criteria: patients older than 18 years with CKD.

Variables: 333 variables: a) Age, gender, weight, height (4); b) Status on the transplant waiting list (2); c) Renal replacement therapy (3); d) Diagnoses (ICD-10) excluding group I50 diagnoses, used as the outcome (146); e) Laboratory variables (78); f) Pharmacological treatments (100).

Label and models: The algorithm used was the Light Gradient Boosting Machine (LGBM), with which four models were trained to perform a binary classification to predict whether a patient would develop heart failure in a period of 3, 4, 5 or 6 years. Label 0 was used to represent the absence of heart failure, while label 1 was used to represent the presence of heart failure. The methodology used to train the models was as follows:

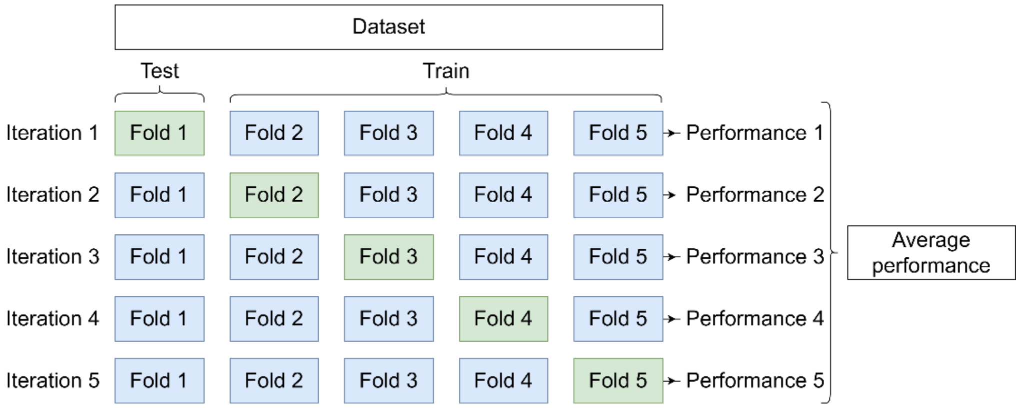
1. Pre-processing of the data to manage missing data and data errors.

Group of variables	Type	Imputation applied
Demographic data	Known value weight and/or height	Simple imputation of the known value
	Unknown weight and height	Median
Laboratory data	Missing values in between known values	Linear interpolation
	Missing values after a known value	Imputation of last known value
	Patients with more than 65% of missing values	Delete instances
	Glomerular filtration rate variable	Estimation equation proposed by Levey et al. 2009, records that cannot be estimated were deleted
	Missing values preceding a known value	Median imputation

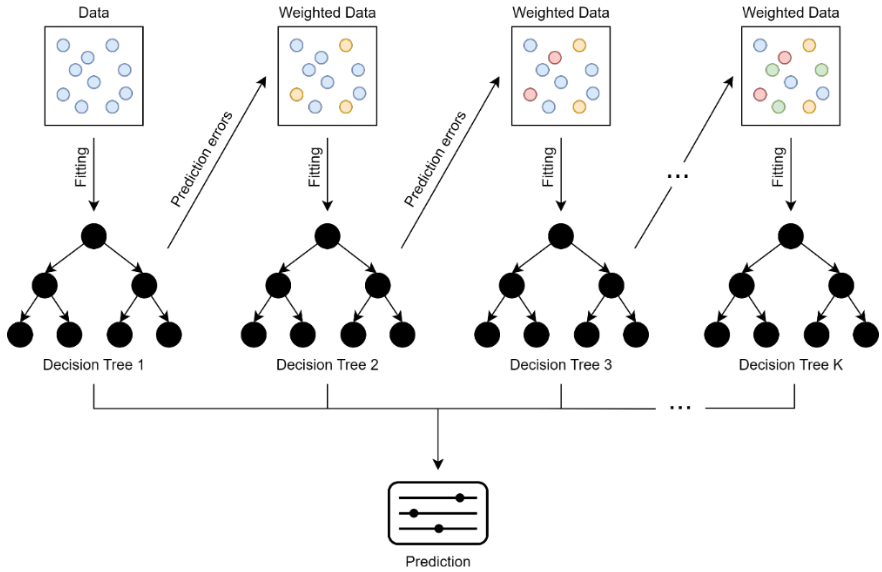
Regarding the outliers, an upper and lower bound (as a credible interval) was defined for each numerical variable. The values out of this interval were considered data errors and deleted.

2. Creating different datasets, one for each prediction horizon.

3. Training and evaluating the LGBM models (with 5-fold CV).



Cross-validation 5-folds . Given that no finetuning have been applied there was no need to split test set out of the cross validation.

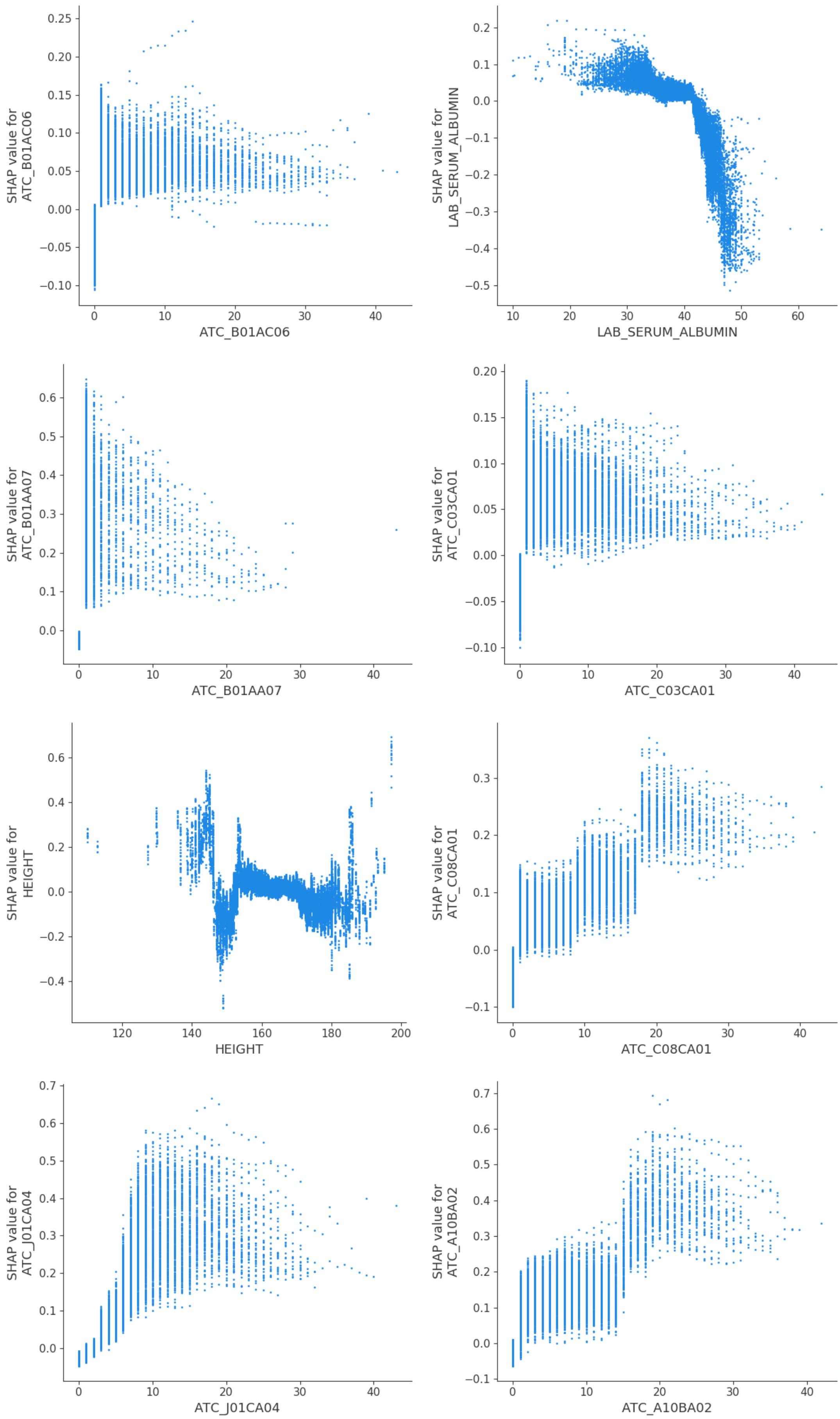


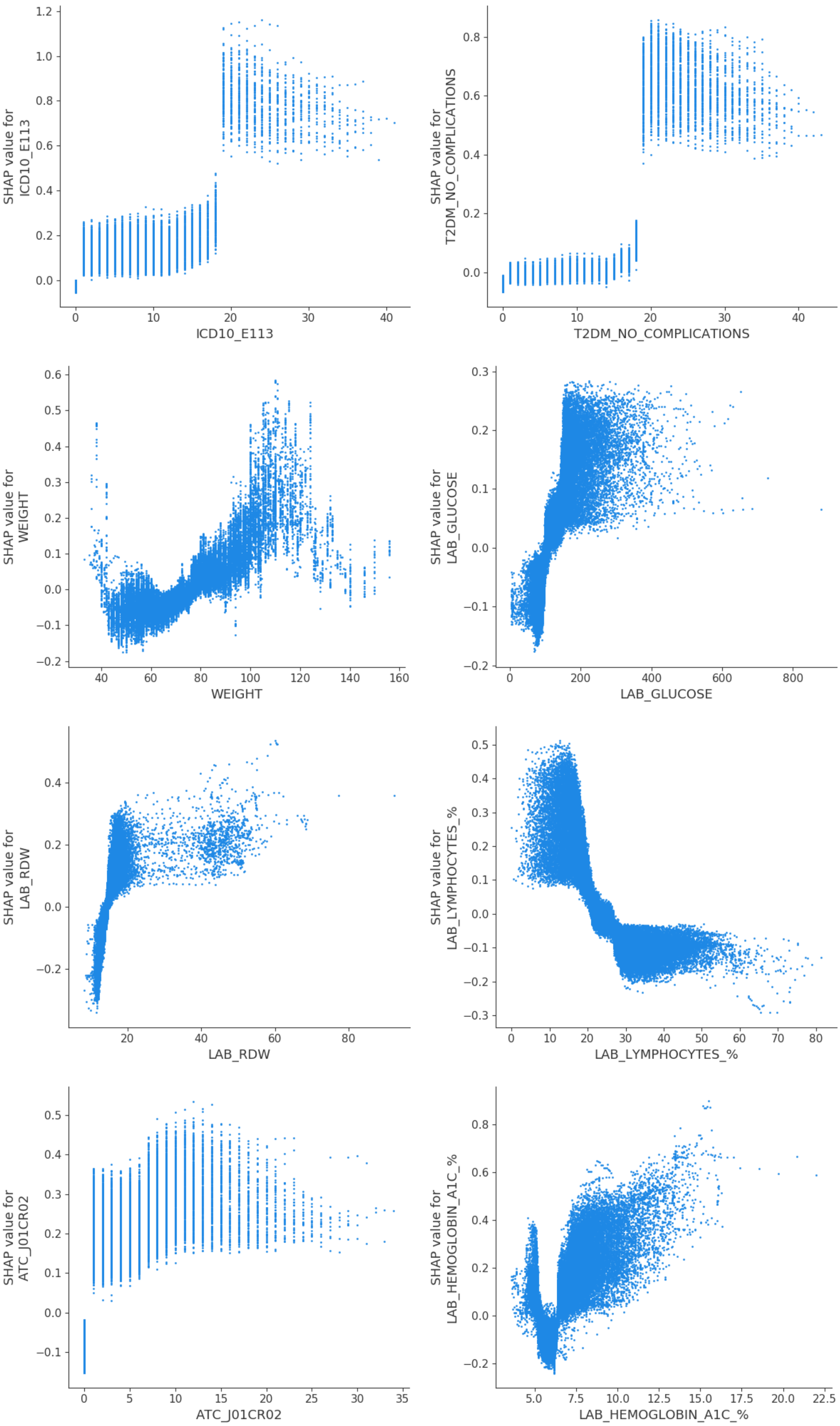
Flow diagram of Gradient Boosting Machine Learning method. The ensemble classifier consists of a set of weak classifiers. The new classifier being added focuses on the correction of the mistakes remaining from the previous models.

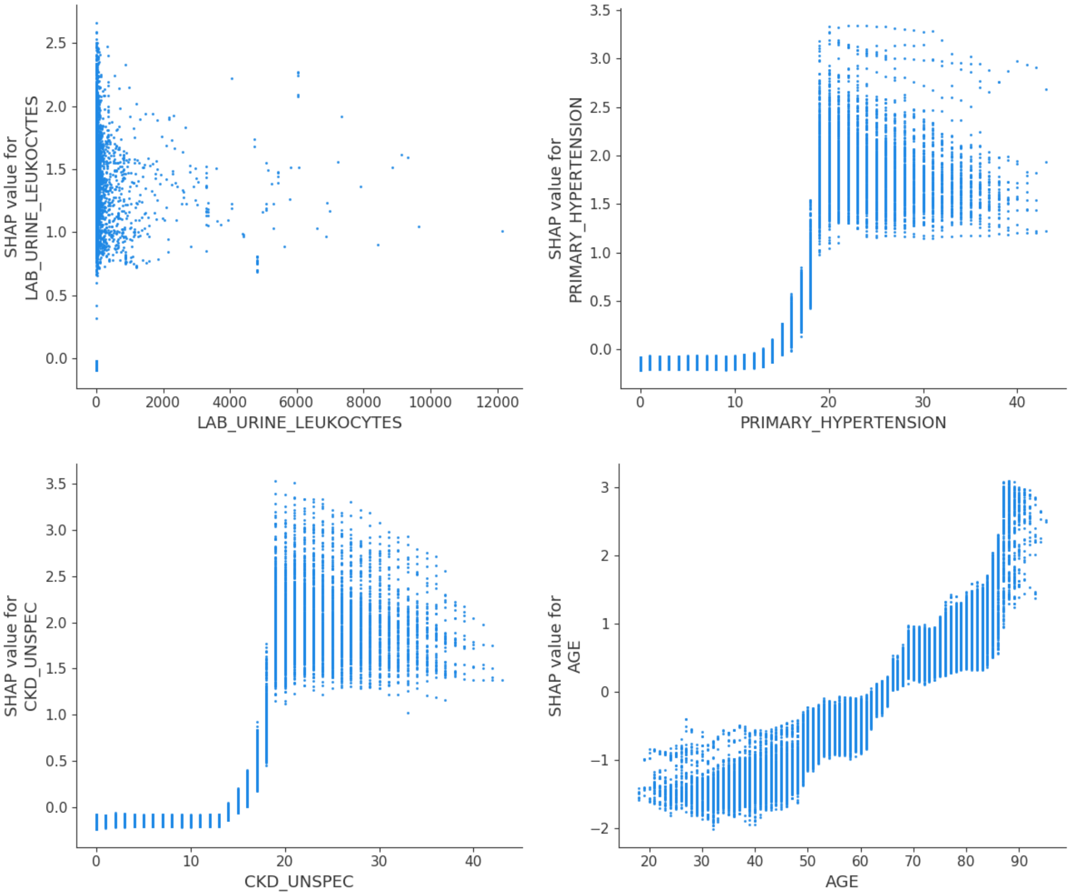
Results:

Age: 63 ± 13 years, Gender: 66% male and 34% were female. Different prediction horizons were tested, and the best results were obtained for 6 years where was achieved an area under the curve of 0.86 and an accuracy of 0.76. The 5 variables with major relevance according to SHAP values (SHapley Additive exPlanations) and in this order are: age, CKD diagnosis, hypertension, haemoglobin, and lymphocytes. The results presented in Figure 1 and Table 1 correspond to the mean obtained for the 5-folds of the Group Cross-Validation.

8 most important features of 6 years prediction:


SHAP value for
ATC_B01AC06
ATC_B01AC06
SHAP value for
LAB_SERUM_ALBUMIN
LAB_SERUM_ALBUMIN
SHAP value for
ATC_B01AA07
ATC_B01AA07
SHAP value for
ATC_C03CA01
ATC_C03CA01
SHAP value for
HEIGHT
HEIGHT
SHAP value for
ATC_C08CA01
ATC_C08CA01
SHAP value for
ATC_J01CA04
ATC_J01CA04
SHAP value for
ATC_A10BA02
ATC_A10BA02


SHAP value for ICD10_E113
ICD10_E113
SHAP value for T2DM_NO_COMPLICATIONS
T2DM_NO_COMPLICATIONS
SHAP value for WEIGHT
WEIGHT
SHAP value for LAB_GLUCOSE
LAB_GLUCOSE
SHAP value for LAB_RDW
LAB_RDW
SHAP value for LAB_LYMPHOCYTES_%
LAB_LYMPHOCYTES_%
SHAP value for ATC_J01CR02
ATC_J01CR02
SHAP value for LAB_HEMOGLOBIN_A1C_%
LAB_HEMOGLOBIN_A1C_%



Conclusion:

- Advanced artificial intelligence algorithms present promising results for the prediction of heart failure in CKD patients.
- This enables early, personalized, and more effective treatments for CKD patients.

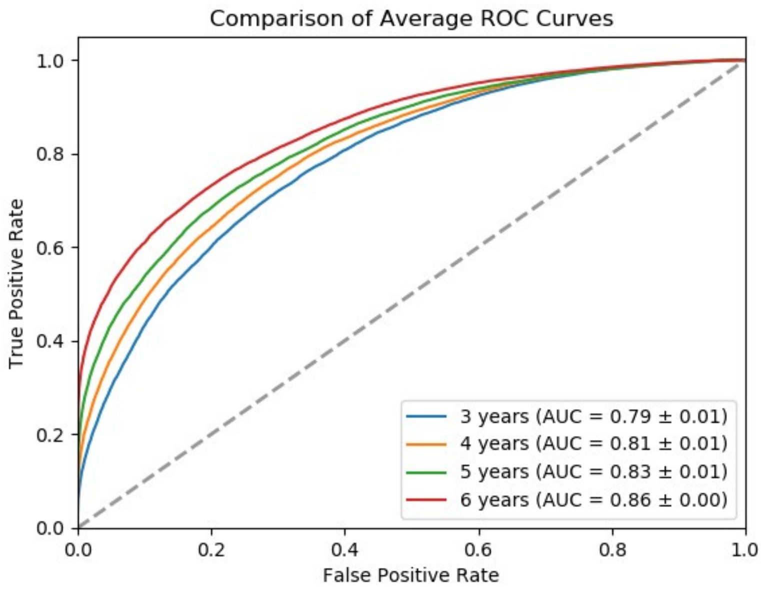


Figure 1. ROC curves for different prediction horizons.

Prediction horizon (years)	HF Balance	Accuracy	F1-Score	Recall	Specificity	AUROC
3	23.66%	0.73	0.54	0.66	0.75	0.79
4	32.55%	0.74	0.63	0.69	0.76	0.81
5	42.63%	0.75	0.71	0.72	0.77	0.83
6	53.76%	0.76	0.77	0.74	0.79	0.86

Table 1. Performance metrics of the models for different prediction horizons.

Confion matrix 6 years model: