Machine learning for prediction of all-cause mortality after transcatheter aortic valve implantation

Jacek Kwiecinski[1,2] MD, PhD, Maciej Dabrowski[1] MD, PhD, Luis Nombela-Franco[3] MD, PhD, Kajetan Grodecki[4] MD, PhD, Konrad Pieszko[2,5] MD, PhD, Zbigniew Chmielak[1] MD, PhD, Anna Pylko[1] MD, Breda Hennessey[3] MD, PhD, Lukasz Kalinczuk[1] MD, PhD, Gabriela Tirado-Conte[3] MD, PhD, Bartosz Rymuza[4] MD, PhD, Janusz Kochman[4] MD, PhD, Maksymilian P. Opolski[1] MD, PhD, Zenon Huczek[4] MD, PhD, Marc R Dweck[6] MD, PhD, Damini Dey[2] PhD, Pilar Jimenez-Quevedo[3] MD, PhD, Piotr Slomka[2] PhD*, Adam Witkowski[1] MD, PhD*

1. Department of Interventional Cardiology and Angiology, Institute of Cardiology, Warsaw, Poland

2. Departments of Medicine (Division of Artificial Intelligence in Medicine) and Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA, USA

3. Cardiovascular Institute. Hospital Clinico San Carlos, IdISSC, Madrid, Spain

4. 1st Department of Cardiology, Medical University of Warsaw, Warsaw, Poland

5. Department of Interventional Cardiology and Cardiac Surgery, University of Zielona Gora, Poland

6. Centre for Cardiovascular Science, University of Edinburgh, Edinburgh, United Kingdom

*Equal contribution as senior author

On-line Supplement

Content:

1. Supplementary Methods. The hyperparameters used for the XGboost model…………..2
2. Supplementary Methods. Missing values …..3
3. Supplementary Results. Prediction of all cause mortality based on preprocedural data…3
4. Supplementary Table 1. Variables used in machine learning……………………………4
5. Supplementary Figure 1. Prediction of 1-year all-cause mortality on external testing…...8
6. Supplementary Figure 2. Further examples of individual prediction of all-cause mortality with explainable artificial intelligence……………………………………………………9
7. Supplementary Figure 3. Feature importance for the machine learning model based on pre-procedural data……………………………………………………………………...10

Supplementary Methods

Standard parameters were Gamma (minimum loss reduction) = 1, standard L2 regularization (lambda)=1. The final hyperparameters found in the grid search for the XGBoost model in the training population were as follows:

Learning rate = 0.001
Maximum depth of a tree = 1
Subsample ratio of the parameters = 0.5
Subsample ratio of the training instances = 0.5
Minimum sum of instance weight (hessian) needed in a child = 1
number of iterations = 5000

Missing values

XGBoost internally provides the missing value default imputation by unique codes for the missing variables to introduce a sparsity pattern and perform sparsity-aware split findings thus it is not necessary to provide explicitly replace missing data. As shown previously this internal XGBoost mechanism is performing comparably to dedicated missing variables replacement methods (1).

1. Rios R, Miller RJH, Manral N, et al. Handling missing values in machine learning to predict patient-specific risk of adverse cardiac events: Insights from REFINE SPECT registry. Comput Biol Med. 2022 Jun;145:105449. doi: 10.1016/j.compbiomed.2022.105449. Epub 2022 Mar 25.

Supplementary Results

A machine-learning model based on 30 pre-procedural variables had an AUC of 0.64 (0.59-0.69). We present a graphical representation of the top predictors of our pre-procedural data model in Supplementary Figure 3.

Supplementary Table 1. Variables used in machine learning.

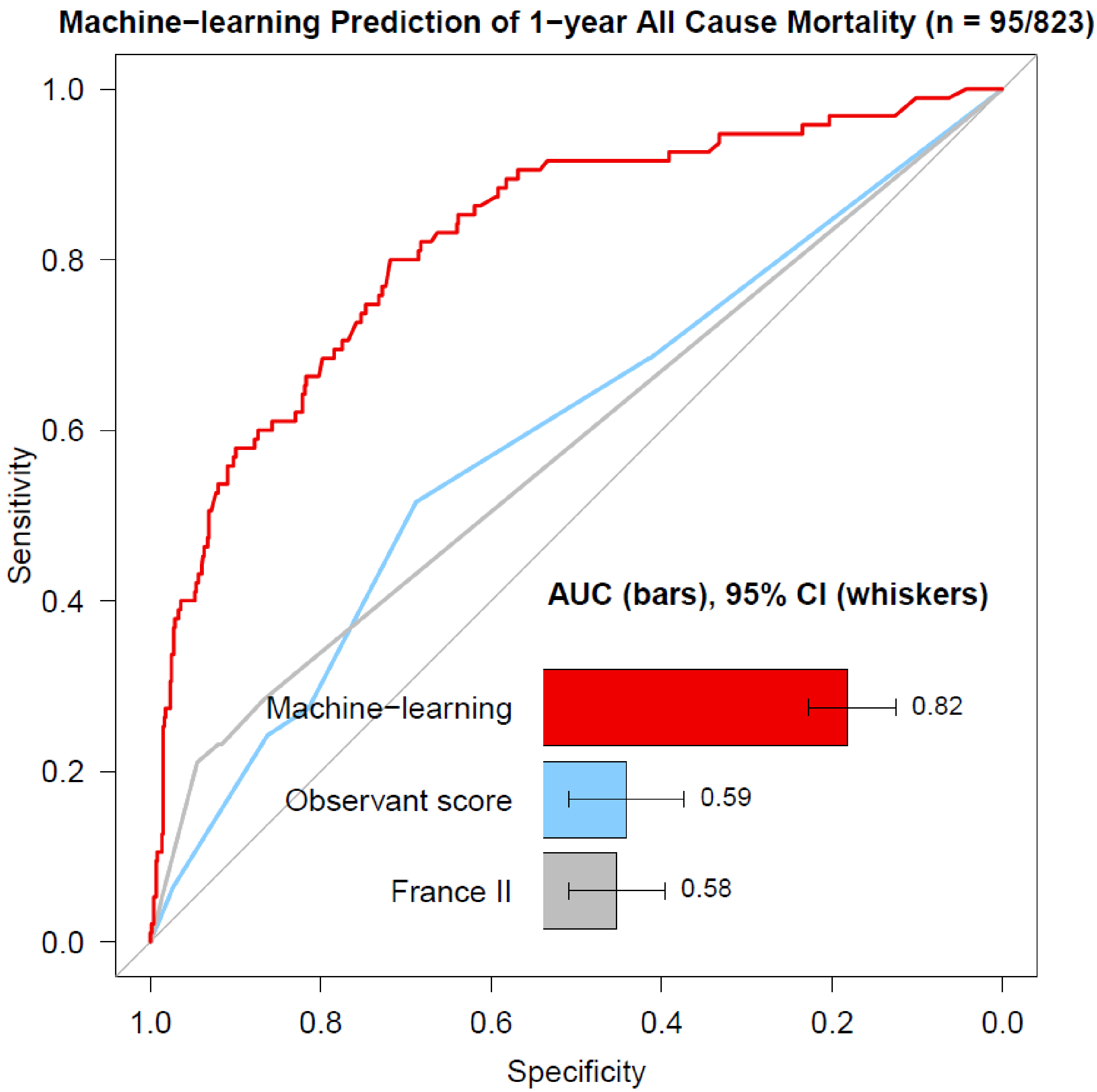
Category	No.	Variable name	Percent of missing values, %
Clinical	1	age (years)	0
	2	Gender (0, 1)	0
	3	height (m)	0
	4	weight (kg)	1
	5	body mass index (kg/m^2)	1
	6	diabetes mellitus (0, 1)	0
	7	past myocardial infarction (0, 1)	0
	8	chronic obstructive pulmonary disease (0, 1)	6
	9	pulmonary hypertension (0, 1)	0
	10	past cerebrovascular accident (0, 1)	0
	11	past percutaneous coronary intervention (0, 1)	1

	12	past coronary artery bypass surgery (0, 1)	0
	13	previous valvular surgery (0, 1)	0
	14	atrial fibrillation (0, 1)	0
	15	past pacemaker implantation (0,1)	0
	16	EuroSCORE II	0
Biomarker	17	Baseline creatine (mg/dL)	0
	18	Baseline Estimated glomerular filtration rate (ml/m2)	0
	19	Baseline hemoglobin (g/dL)	0
	20	Baseline platelets (n/dL)	0
	21	Baseline NT-proBNP	18
Echocardiography	22	Baseline left ventricular ejection fraction (%)	1
	23	Baseline effective orifice area (cm2)	1
	24	Peak transvalvular pressure gradient (mmHg)	1
	25	Mean transvalvular pressure gradient (mmHg)	1
	26	Aortic regurgitation (0,1)	2
	27	Mitral regurgitation (0, 1)	3
	28	Tricuspid regurgitation (0, 1)	8

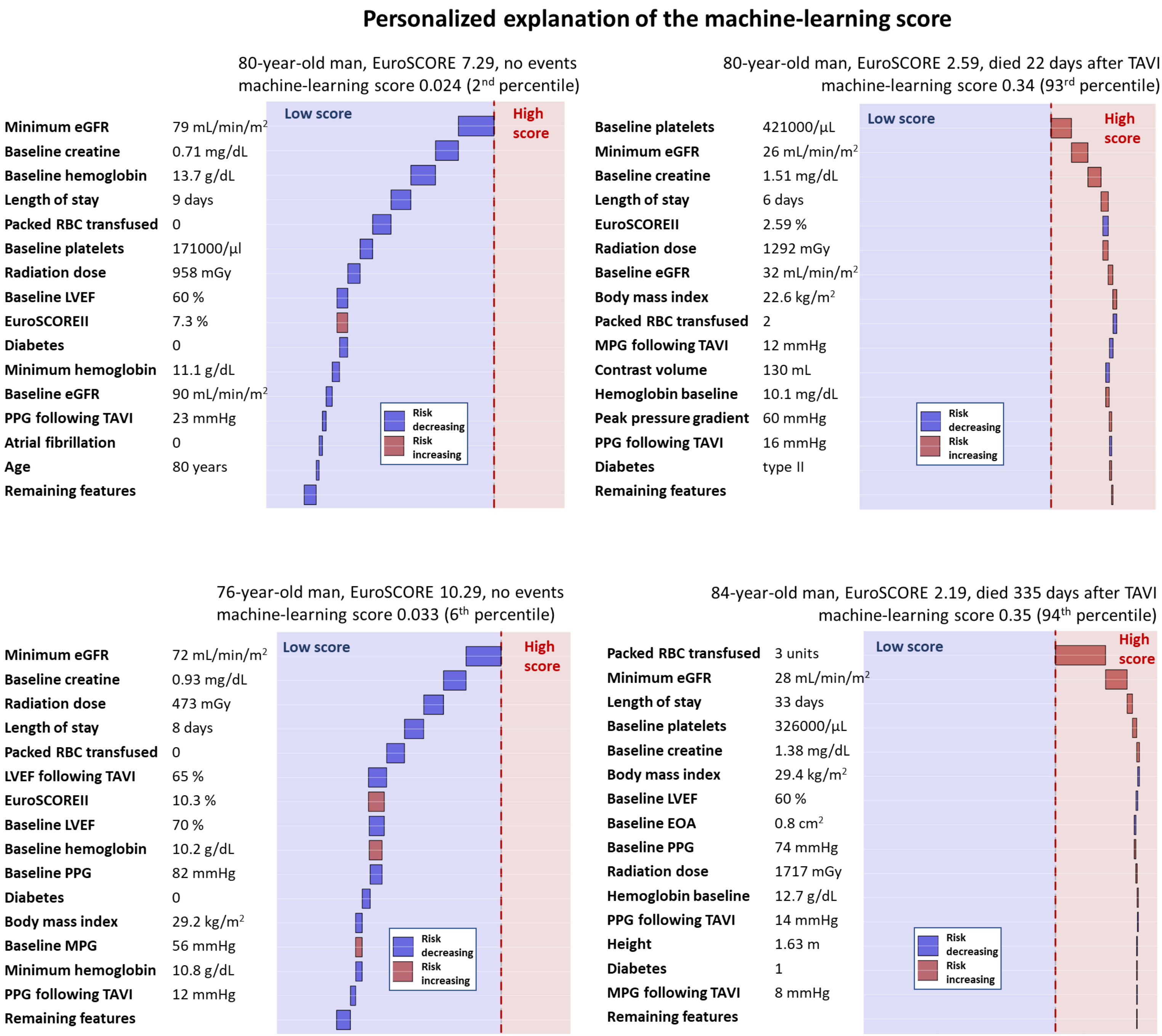
	29	Pulmonary artery systolic pressure (mmHg)	23
	30	Bicuspid aortic valve (0, 1)	7
Procedure	31	General anesthesia (0, 1)	0
	32	Non femoral TAVI access (0, 1)	0
	33	Contrast volume (mL)	6
	34	Fluoroscopy time (min)	3
	35	Radiation dose (mGy)	16
	36	Minimum hemoglobin following TAVI (g/dL)	0
	37	Minimum platelets following TAVI (n/dL)	0
	38	Minimum eGFR following TAVI (ml/m2)	0
	39	Major vascular complication (0, 1)	1
	40	Minor vascular complication (0, 1)	2
	41	Life threatening bleeding (0, 1)	0
	42	Major bleeding (0, 1)	0
	43	Minor bleeding (0, 1)	0
	44	Packed red blood cells transfused (units)	0
	45	Periprocedural myocardial infarction (0, 1)	0
	46	Periprocedural stroke (0, 1)	0

	47	Coronary occlusion (0, 1)	0
	48	Annulus rupture (0, 1)	0
	49	Pacemaker implantation following TAVI (0, 1)	1
	50	Hospitalization length (days)	0
	51	LV ejection fraction following TAVI (%)	8
	52	Peak transprosthetic pressure gradient (mmHg)	4
	53	Mean transprosthetic pressure gradient (mmHg)	4
	54	Aortic regurgitation following TAVI (0, 1)	9

Supplementary Figure 1. Prediction of 1-year all-cause mortality on external testing. Receiver-operating characteristic curves for prediction of 1-year all-cause mortality following hospital discharge after successful transcatheter aortic valve implantation. The machine learning XGBoost model had a significantly higher area-under-the-curve for all-cause mortality prediction than an established risk scores.



Supplementary Figure 2. Further examples of individual prediction of all-cause mortality with explainable artificial intelligence



Supplementary Figure 3. Feature importance for the machine learning model based on pre-procedural data.