Technische Universität München
TUM School of Medicine and Health

On the impact of artificial intelligence for refined diagnostic and prognostic resolution in patients with severe aortic stenosis undergoing transcatheter aortic valve replacement

Elena Louisa Charlotte Rippen

Bei der folgenden Dissertation wurde die Form der publikationsbasierten Promotion (gemäß TUM Promotionsordnung vom 12. März 2012 §6) gewählt.

Hierfür wurde durch die Doktorandin, Frau Elena Rippen, das Kriterium von mindestens zwei akzeptierten Veröffentlichungen als Co-Erstautorin in englischsprachigen, international verbreiteten Publikationsorganen erfüllt. Zudem haben die publizierten Manuskripte, die dieser Dissertation zugrunde liegen, jeweils ein *peer review*-Verfahren durchlaufen.

Für die Veröffentlichung der Manuskripte im Kontext dieser Dissertation wurde das Einverständnis des jeweiligen Verlages eingeholt.

Die folgende Dissertation ist in Englisch verfasst.

Dedicated to my family

in love and gratitude.

Parts of this dissertation thesis have been published as:

1. Lachmann M*, **Rippen E***, Schuster T, Xhepa E, von Scheidt M, Pellegrini C, Trenkwalder T, Rheude T, Stundl A, Thalmann R, Harmsen G, Yuasa S, Schunkert H, Kastrati A, Laugwitz K-L, Kupatt C, Joner M. Subphenotyping of Patients With Aortic Stenosis by Unsupervised Agglomerative Clustering of Echocardiographic and Hemodynamic Data. JACC Cardiovasc. Interv. 2021;14(19):2127–2140 (* equal contributors).

2. Lachmann M*, **Rippen E***, Rueckert D, Schuster T, Xhepa E, von Scheidt M, Pellegrini C, Trenkwalder T, Rheude T, Stundl A, Thalmann R, Harmsen G, Yuasa S, Schunkert H, Kastrati A, Joner M, Kupatt C, Laugwitz K-L. Harnessing feature extraction capacities from a pre-trained convolutional neural network (VGG-16) for the unsupervised distinction of aortic outflow velocity profiles in patients with severe aortic stenosis. Eur. Heart J. - Digit. Health 2022:ztac004 (* equal contributors).

3. Fortmeier V, Lachmann M, Körber MI, Unterhuber M, von Scheidt M, **Rippen E**, Harmsen G, Gerçek M, Friedrichs KP, Roder F, Rudolph TK, Yuasa S, Joner M, Laugwitz K-L, Baldus S, Pfister R, Lurz P, Rudolph V. Solving the Pulmonary Hypertension Paradox in Patients With Severe Tricuspid Regurgitation by Employing Artificial Intelligence. JACC Cardiovasc. Interv. 2022;15(4):381–394.

4. Lachmann M, **Rippen E**, Schuster T, Xhepa E, von Scheidt M, Trenkwalder T, Pellegrini C, Rheude T, Hesse A, Stundl A, Harmsen G, Yuasa S, Schunkert H, Kastrati A, Laugwitz K-L, Joner M, Kupatt C. Artificial intelligence-enabled phenotyping of patients with severe aortic stenosis: on the recovery of extra-aortic valve cardiac damage after transcatheter aortic valve replacement. Open Heart 2022;9(2):e002068.

5. Trenkwalder T, Lachmann M, Stolz L, Fortmeier V, Covarrubias HAA, **Rippen E**, Schürmann F, Presch A, von Scheidt M, Ruff C, Hesse A, Gerçek M, Mayr NP, Ott I, Schuster T, Harmsen G, Yuasa S, Kufner S, Hoppmann P, Kupatt C, Schunkert H, Kastrati A, Laugwitz K-L, Rudolph V, Joner M, Hausleiter J, Xhepa E. Machine learning identifies pathophysiologically and prognostically informative phenotypes among patients with mitral regurgitation undergoing transcatheter edge-to-edge repair. Eur Heart J - Cardiovasc Imaging. 2023;:jead013.

Abbreviations

AS: aortic stenosis

AVA: aortic valve area

AVG_{mean}: mean aortic valve gradient

LFLG AS: low-flow, low-gradient aortic stenosis

LVEF: left ventricular ejection fraction

LVOT: left ventricular outflow tract

mPAP: mean pulmonary artery pressure

SAVR: surgical aortic valve replacement

SVi: stroke volume index

TAPSE: tricuspid annular plane systolic excursion

TAVR: transcatheter aortic valve replacement

VTI: velocity time integral

Contents

1   Introduction

1.1   Anatomy of the aortic valve.

The aortic valve is located between the left ventricle and the aorta. The normal aortic valve is tricuspid; the cusps have a semilunar shape and are freely mobile (see Figure 1). The cusps are named according to their location with respect to the coronary arteries, i.e. left coronary, right coronary and non-coronary cusp (1). The aortic valve opens during systole, when the left ventricle contracts and the pressure in the left ventricle exceeds that in the aorta. The arrangement of cusps of the aortic valve is designed to allow non-turbulent forward flow of blood into the aorta during systole while also preventing backflow during diastole. Non-turbulent forward blood flow is important, as it results in an even distribution of mechanical stress to the aortic wall. Calcification of the cusps results in altered biomechanical properties of the aortic valve; narrowing of the aortic valve orifice leads to increases in resistance, and hence a greater force of left ventricular contraction is required to eject the same volume of blood.
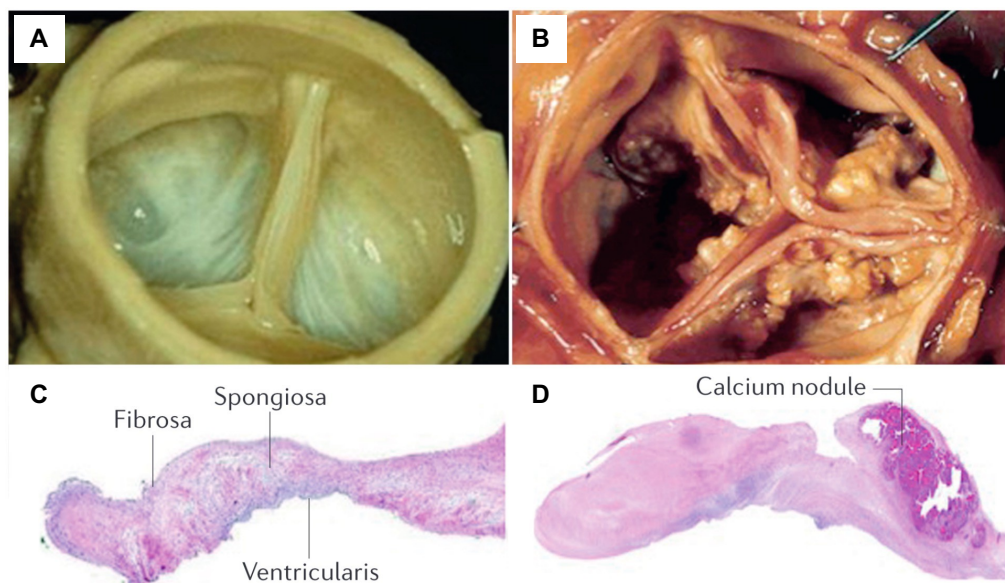
Figure 1: Normal and pathologically calcified macroscopic aortic valve morphology and histological appearance.

Photographs of a healthy tricuspid aortic valve (A) and a severely calcified aortic valve (B). Histopathological sections comparing the trilaminar structure from a healthy aortic valve cusp (C) with pathological alterations such as fibrotic material and a calcified nodule found in a diseased aortic valve cusp (D). After the initial phase, when initial injury occurs, deposition of calcium triggers a vicious circle increasing valve wall stress and tissue injury, reducing leaflet compliance, and enhancing further valve calcification (propagation phase). Adapted with permission from Lindman *et al.* (1).

1.2   Epidemiology of aortic stenosis: becoming the most frequent valvular heart disease.

Calcific aortic stenosis (AS) results from progressive fibrocalcific remodeling of the aortic valve. It represents the most common valvular heart disease in developed countries, and the burden that it poses to health care systems is expected to increase over the next decades owing to ageing populations and the lack of effective prevention strategies to slow down disease progression. Whilst the prevalence of AS among adults between the ages 50 and 59 years ranges at only 0.2%, its prevalence increases almost linearly with age up to 9.8% among octogenarians as a result of its degenerative pathophysiology (2–4).

1.3   Etiology: from streptococcal throat infections to age-related degenerative calcification.

Calcific degeneration, a congenital bicuspid valve and rheumatic heart disease represent the most common causes for severe AS (5). Importantly, by 1970, the main etiology for AS in developed countries shifted from rheumatic fever to age-related degenerative calcification (6). This near complete elimination of rheumatic heart disease in developed countries during the late 20th century can be attributed to improvements of socioeconomic conditions and the widespread use of penicillin G benzathine to treat group A streptococcus throat infections. Nevertheless, rheumatic fever and related rheumatic heart disease persist in the poorest regions of the world (such as Oceania, South Asia, and central sub-Saharan Africa), resulting in 33 million people suffering from rheumatic heart disease and 319,000 deaths due to rheumatic heart disease in 2015 (7). Whilst patients with rheumatic heart disease secondary to untreated streptococcal pharyngitis present during the sixth decade of life, patients with calcific AS typically develop symptoms at the age of ≈80 years (Figure 2) (8). A plethora of anatomical, genetic, and clinical factors all contribute to the pathogenesis of calcific AS. Clinical factors associated with calcific AS include older age, male gender, hypercholesterolemia, hypertension, smoking, and diabetes (9,10). Not surprisingly, coronary artery disease is a widespread comorbidity in patients with calcific AS. Moreover, mechanical stress due to abnormal blood-flow dynamics may also play a key role in the development of calcific AS, as congenitally bicuspid aortic valves, which are found in about 0.5 to 0.8% of the population, disproportionately frequently represent the underlying anatomy in patients undergoing valve replacement for severe AS in developed countries (11).
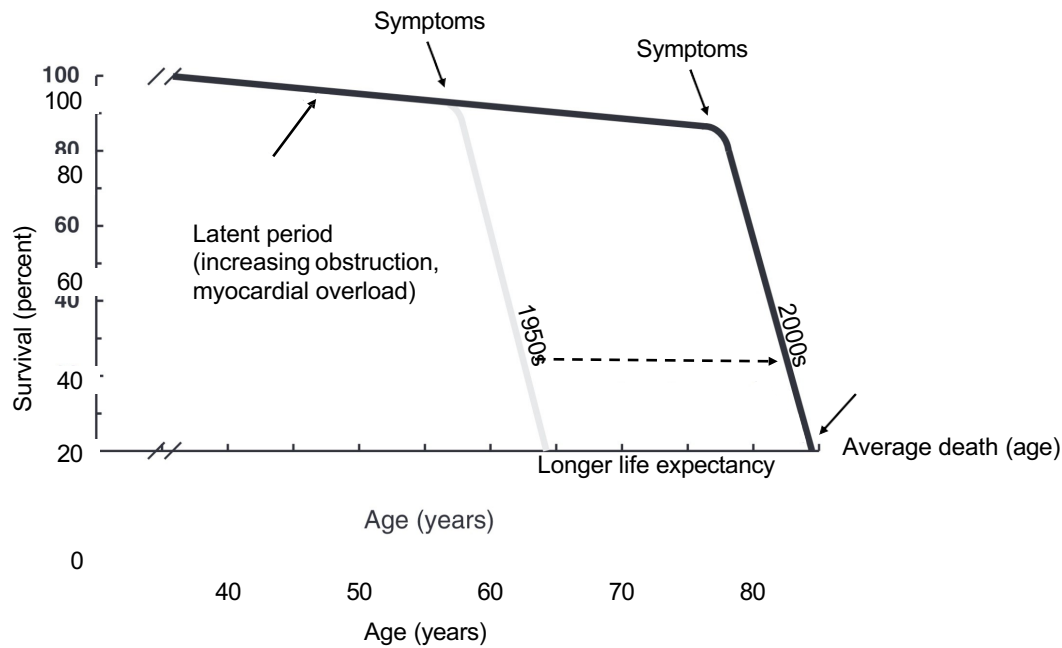
Figure 2: On the shift in age at presentation among patients with severe aortic stenosis.

Onset of symptoms is preceded by a long latent period of progressive narrowing of the aortic valve and consecutive myocardial pressure overload. The onset of symptoms heralds a rapid increase in mortality unless valve replacement is performed. Adapted with permission from Carabello *et al.* (6).

1.4   Clinical presentation: dyspnea, angina, and syncope represent the symptom triad.

The natural course of AS is characterized by a long latency period, in which the left ventricle adapts to progressive aortic valve narrowing with concentric muscle hypertrophy. Eventually, left ventricular remodeling becomes maladaptive and patients begin to experience symptoms. The onset of symptoms in patients with severe AS heralds a rapid extra-aortic valve cardiac disease progression and a high rate of death in untreated patients (see Figure 2) (12). Unless valve replacement is performed, most will die within five years (13). The classic symptom triad of AS include dyspnea, angina, and syncope, and they often only appear upon significant obstruction of the aortic valve. Moreover, it is sometimes difficult in elderly and cardiovascular multimorbid patients to distinguish symptoms caused by severe AS from that of other medical conditions such as coronary artery disease, atrial fibrillation, and deconditioning. Dyspnea on exertion is found as the most common initial symptom, and hemodynamic studies revealed that dyspnea develops particularly in patients with left ventricular diastolic dysfunction and elevated left-sided filling pressures (see Figure 3) (14). Angina as the second most common symptom develops when the myocardial oxygen demand exceeds the coronary oxygen supply, and angina is therefore often described by patients with coexisting coronary artery disease and left ventricular hypertrophy. Those patients experience angina typically during exercise, when the heart rate accelerates and the time available for coronary blood flow decreases (15,16). Approximately 15% of AS patients present with

syncope, which is the least frequent but most urgent component of the classic symptom triad. Patients who present with syncope are typically diagnosed with a smaller left ventricular cavity and a reduced output, ultimately resulting in inadequate cerebral blood flow during exercise or in the context of cardiac arrhythmias (14,17).
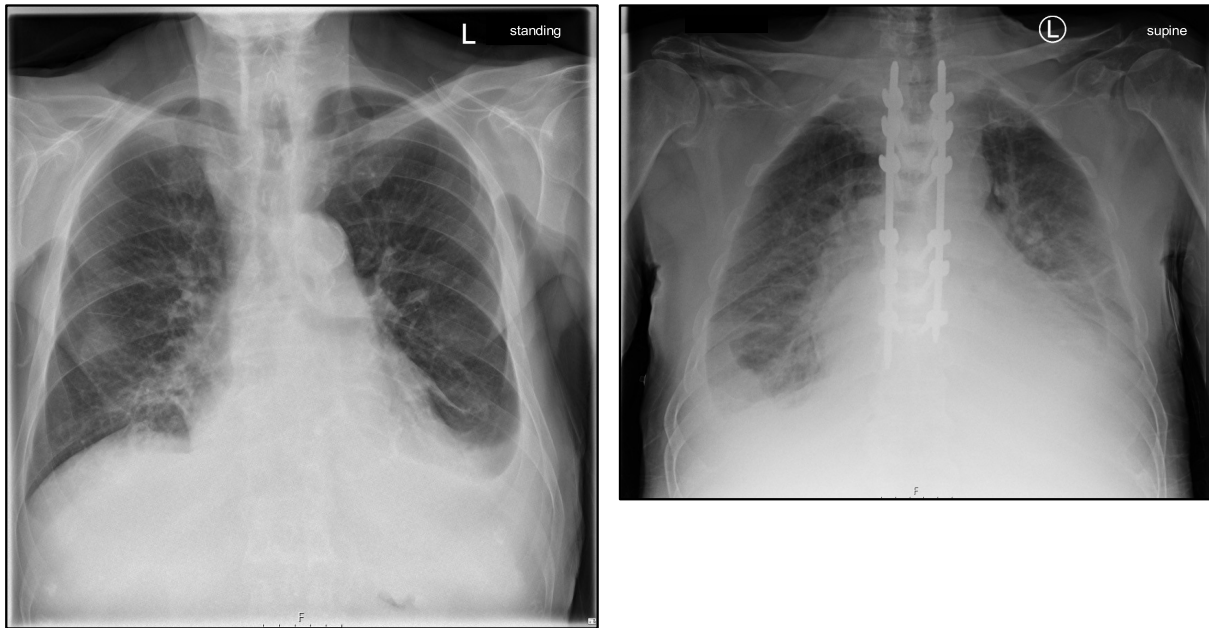
Figure 3: Cardiac decompensation resulting in pleural effusion (chest X-ray on the left side) and pulmonary edema (chest X-ray on the right side) in patients with severe aortic stenosis.

1.5 Diagnostics: universally available echocardiography is the cornerstone, but right heart catheterization might prove helpful in inconclusive cases.

Transthoracic echocardiography represents the primary diagnostic modality for the evaluation of AS. Apart from confirming the diagnosis and assessing the severity of AS, echocardiography also allows to evaluate extra-aortic valve cardiac damages such as left ventricular dysfunction and concurrent valvular insufficiencies, hence providing further prognostic information, and identifying ancillary treatment targets. The severity of AS is typically assessed by continuous-wave Doppler echocardiography measuring the velocity of blood flow through the narrowed aortic valve orifice (see Figure 4). By applying the simplified Bernoulli equation, the pressure gradient across the aortic valve can be further calculated. Notably, the mean aortic valve gradient (AVG_{mean}) is a more robust parameter than the maximum transvalvular jet velocity. Apart from being operator-dependent, the Bernoulli equation is also a gross simplification of valve hemodynamics, as it assumes a steady laminar flow, which is clearly not the case

in severe AS. Whilst mild AS is defined by a maximum transvalvular jet velocity < 3 m/s and/ or an AVG$_{mean}$ < 20 mmHg, severe AS is defined by a maximum transvalvular jet velocity ≥ 4 m/s and /or an AVG$_{mean}$ ≥ 40 mmHg (see Table 1) (18,19).

	Mild	Moderate	Severe
Maximum transvalvular jet velocity (m/s)	2.5-3.0	3.0-4.0	>4.0
Maximum aortic valve gradient (mmHg)	<40	40-65	>65
Mean aortic valve gradient (mmHg)	<20	20-40	>40
Aortic valve area (cm²)	>1.5	1.0-1.5	<1.0
Indexed aortic valve area (cm²/m²)	>0.85	0.60-0.85	<0.60

Table 1: Grading of aortic stenosis.

Moreover, employment of the continuity equation based on the principle of conservation of mass allows for calculation of the functional aortic valve area (AVA), which is particularly useful in cases, where peak transvalvular jet velocity and AVG$_{mean}$ range lower than 4 m/s and 40 mmHg, respectively, but concerns about AS severity persist. According to the continuity equation, the stroke volume ejected through the left ventricular outflow tract (LVOT) must be equal to the stroke volume passing through the stenotic aortic valve orifice. Knowing the respective time velocity integral (VTI) of blood flow through the LVOT and the aortic valve, the functional AVA can then be calculated as follows: AVA = LVOT$_{area}$ x (VTI$_{LVOT}$) / VTI$_{aorta}$ (see Figure 4).

A

Continuity equation: what comes in … comes out.

volume in = volume out

Flow rate (cm^3/s) = area x velocity

area$_x$ x velocity$_x$ = area$_y$ x velocity$_y$

B

velocity$_{LVOT}$
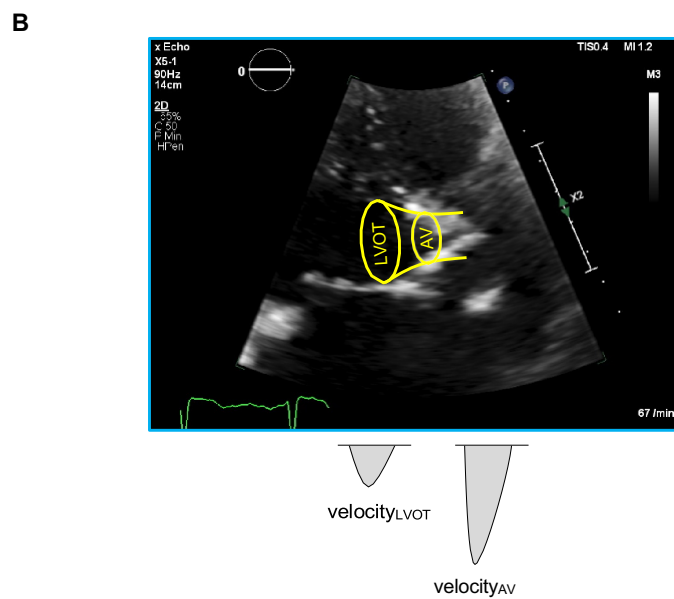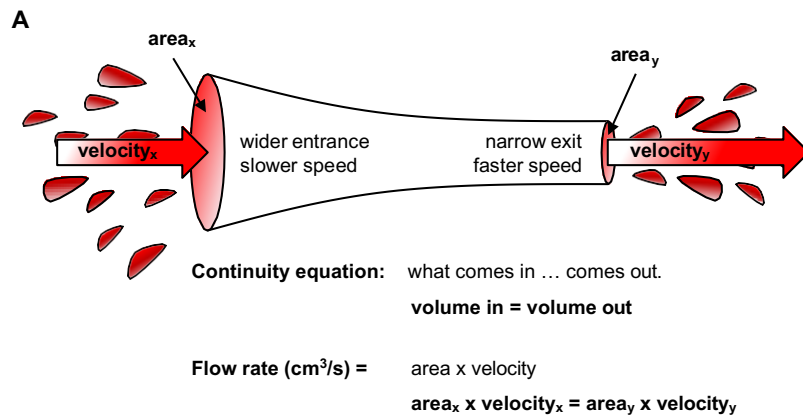
velocity$_{AV}$

area$_{LVOT}$ x velocity$_{LVOT}$ = area$_{AV}$ x velocity$_{AV}$

area$_{AV}$ = (area$_{LVOT}$ x velocity$_{LVOT}$) / velocity$_{AV}$

Figure 4: On the continuity equation to calculate the functional aortic valve area.

AV: aortic valve; LVOT: left ventricular outflow tract.

A high-flow, high-gradient AS is typically met in patients with preserved left ventricular systolic function (see Figure 5).
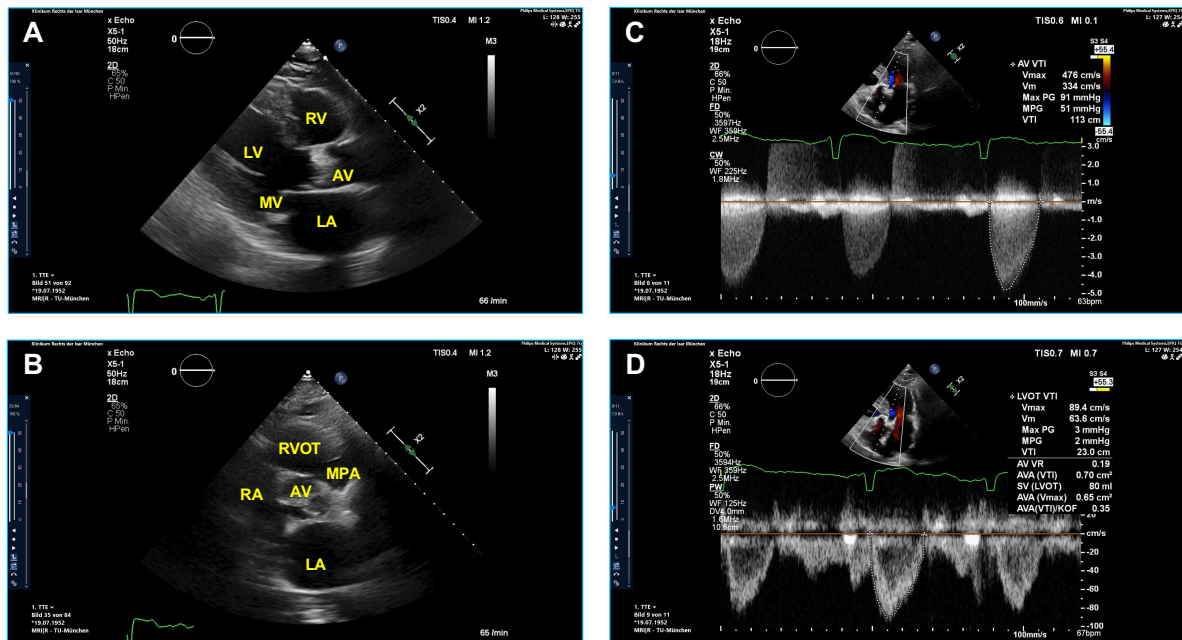
Figure 5: Transthoracic echocardiography in a patient presenting with severe aortic stenosis.

A: Parasternal long-axis view. Please note the thickened right and non-coronary cusp of the aortic valve.

B: Parasternal short-axis view. Please note the irregular calcification and thickening of the aortic valve.

C: Continuous-wave Doppler echocardiography to assess flow velocity in the proximal aorta. Please note that a maximum transvalvular jet velocity of 4.8 m/s and a mean aortic valve gradient of 51 mmHg were measured.

D: Pulsed wave Doppler echocardiography to assess flow velocity in the left ventricular outflow tract. Please note that a functional aortic valve area of 0.70 cm^2 was calculated using the continuity equation, confirming the diagnosis of severe aortic stenosis.

AV: aortic valve; LA: left atrium; LV: left ventricle; MPA: main pulmonary artery; MV: mitral valve; RA: right atrium; RV: right ventricle; RVOT: right ventricular outflow tract.

Criteria to be met for the diagnosis of a classical low-flow, low-gradient (LFLG) AS are as follows: 1) maximum transvalvular jet velocity < 4 m/s and/ or AVG_{mean}< 40 mmHg, 2) AVA ≤ 1 cm^2, and 3) left ventricular ejection fraction (LVEF) < 50% (see Figure 6). Notably, the impaired left ventricular systolic function in classical LFLG AS might theoretically be caused by severe AS itself as a consequence of long-standing afterload mismatch, but more typically LFLG AS develops in multimorbid patients suffering from the combined presence of high afterload due to aortic valve obstruction and intrinsic myocardial

impairment secondary to ischemic heart disease for instance. Another important entity is "paradoxical" LFLG AS as encountered in patients with preserved LVEF. Those patients present with reduced left ventricular stroke volume indices (SVi ≤ 35 mL/m^2), resulting from small left ventricular cavities (typical for elderly, hypertensive women with small body size and concentric hypertrophy), severe diastolic dysfunction (up to one-third of "paradoxical" LFLG AS patients might have concomitant cardiac amyloidosis (20)) or significant mitral regurgitation (21). A normal-flow, low-gradient AS can be encountered in patients with co-existing mitral regurgitation.
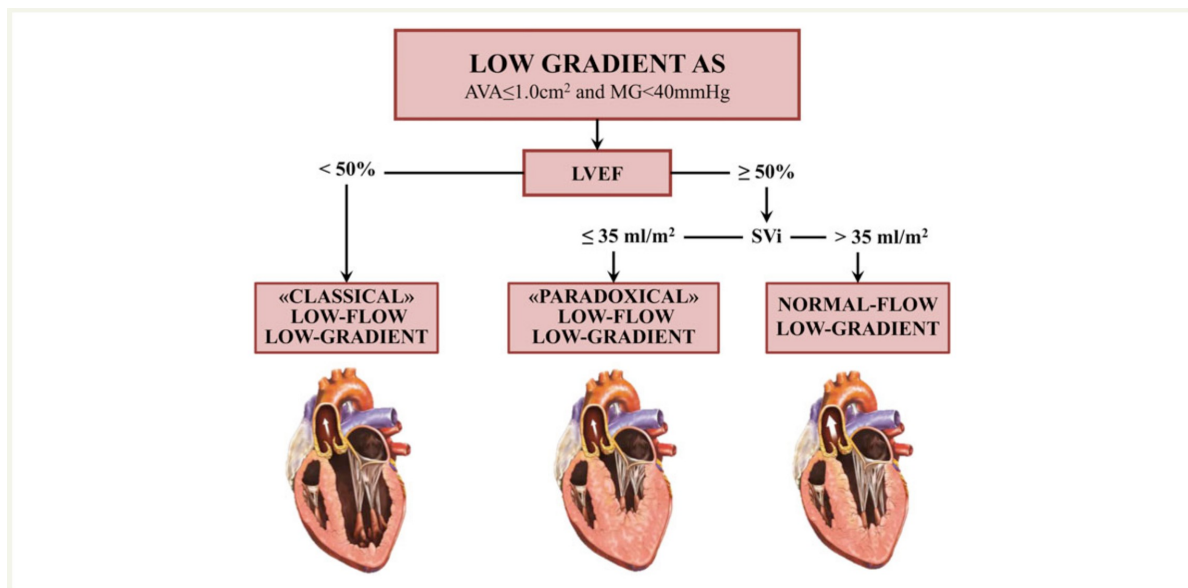
Figure 6: Subtypes of low-gradient aortic stenosis.

AS: aortic stenosis; AVA: aortic valve area; LVEF: left ventricular ejection fraction; MG: mean aortic valve gradient; SVi: stroke volume index. Adapted with permission from Clavel *et al.* (22).

Obviously, measurement accuracy in terms of optimal alignment of the Doppler with the transvalvular jet and adequate recording of the true gradient is crucial to avoid an erroneous diagnosis of apparently moderate AS when severe AS is actually present. Planimetry of the AVA is therefore sometimes performed by transesophageal echocardiography (see Figure 7).
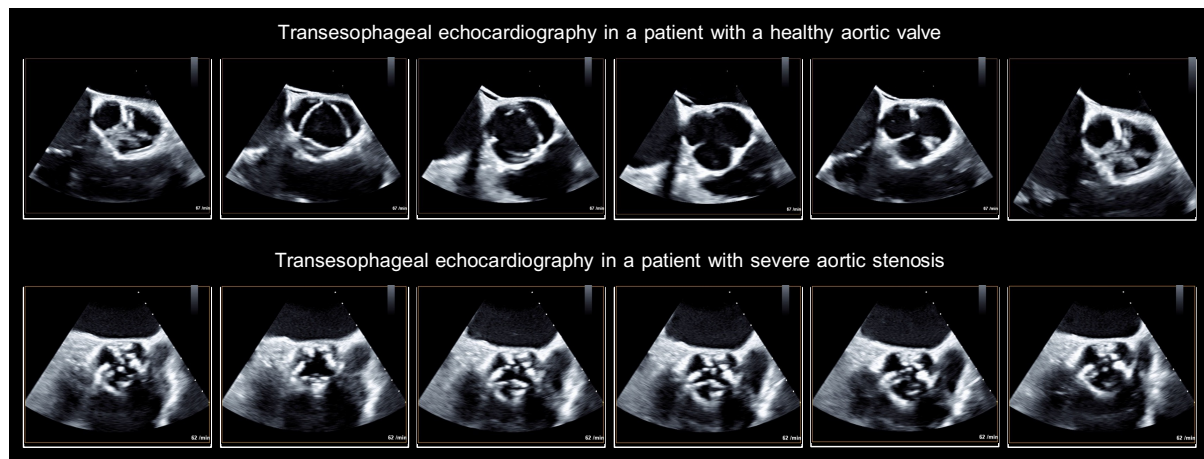
Figure 7: Planimetry of the aortic valve area by transesophageal echocardiography.

In cases with inconclusive clinical and imaging findings, invasive evaluation of AS severity can prove further helpful. Notably, coronary angiography is essential prior to transcatheter aortic valve replacement (TAVR) and surgical aortic valve replacement (SAVR) in order to determine the potential need for concomitant revascularization (18). In inconclusive cases, right heart catheterization can be additionally performed, allowing AVA calculation based on transvalvular volume flow rate and pressure measurements. Moreover, right heart catheterization represents the gold standard to assess and classify pulmonary hypertension (23). Accurate assessment of pulmonary hypertension is also prognostically relevant in patients presenting with severe AS (24). At the same time, echocardiographic assessment of systolic pulmonary artery pressure levels tends to systematic underestimation in patients with severe tricuspid regurgitation (25,26), which can also be found in a considerable proportion of patients with severe AS (27,28).

1.6   Indications for aortic valve replacement: clinical symptoms, aortic stenosis severity, and left ventricular response indicate the need for intervention.

There are no directed medical therapies to slow the progression of severe AS besides generally reducing cardiovascular risk factors and optimizing comorbid conditions. In fact, a large, randomized, prospective clinical trial comparing aggressive lipid lowering therapy (daily intake of 40 mg of simvastatin plus 10 mg of ezetimibe) with placebo could not detect differences with regards to hemodynamic progression or time to valve replacement in adults with mild-to-moderate, asymptomatic AS (29). Patient outcomes therefore crucially depend on appropriate timing and type of aortic valve replacement as the

only effective treatment. Once symptoms occur, the mortality rate increases to > 50% within two years unless aortic valve replacement is performed promptly (30). Aortic valve replacement with a surgical or transcatheter approach is therefore a class I recommendation for symptomatic patients with severe AS, unless comorbidities or limited life expectancy suggest a more appropriate palliative care (18). Since AS is a progressive disease that will inevitably result in obstruction of the left ventricular outflow, even seemingly mild symptoms such as exertional dyspnea or decreased exercise capacity are an indication for aortic valve replacement in patients with severe AS (see Figure 8).

Figure 8: Management of patients with severe aortic stenosis (European Society of Cardiology guideline).

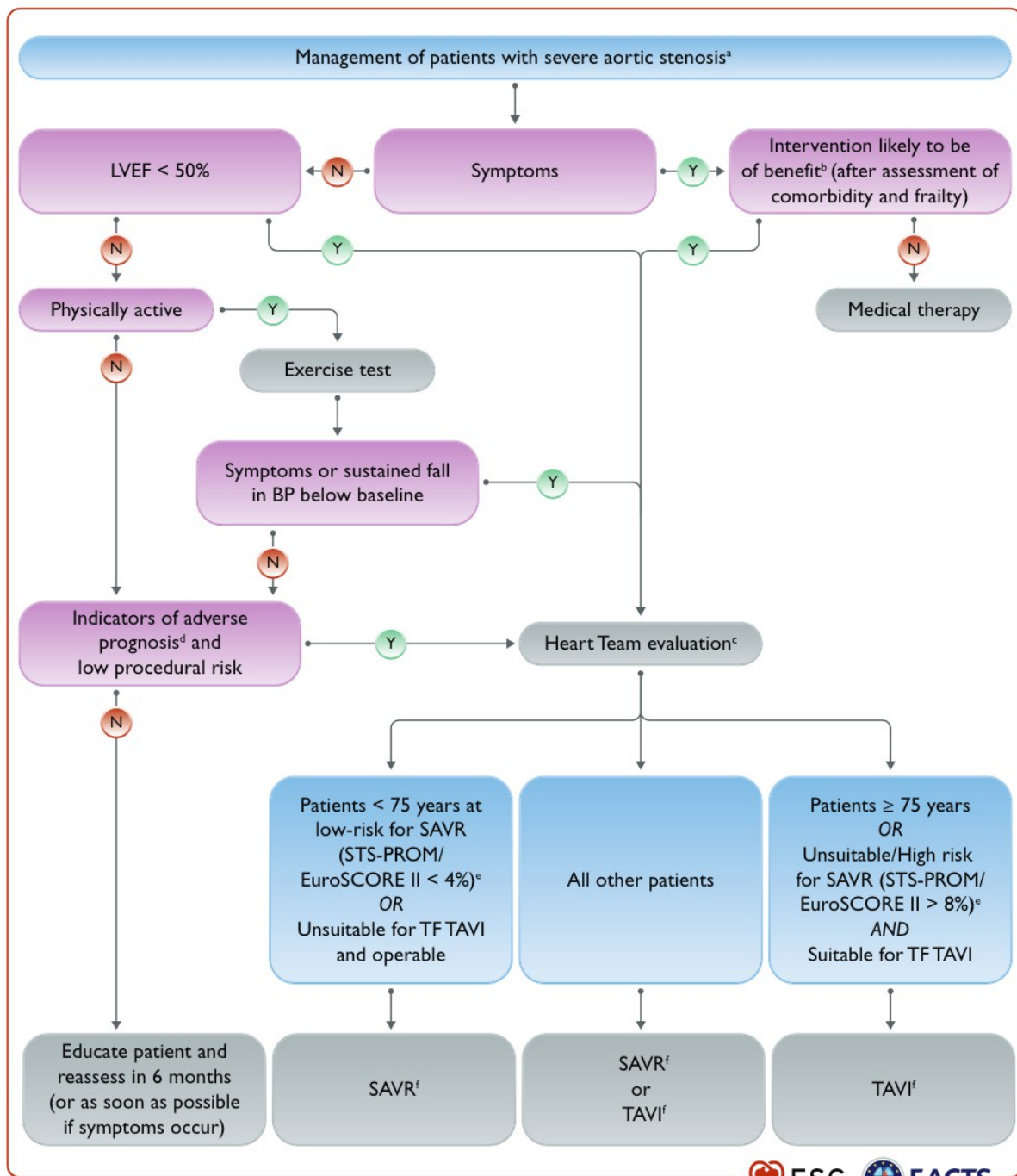BP: blood pressure; EuroSCORE: European System for Cardiac Operative Risk Evaluation; LVEF: left ventricular ejection fraction; SAVR: surgical aortic valve replacement; STS-PROM: Society of Thoracic Surgeons - predicted risk of mortality; TAVI: transcatheter aortic valve implantation; TF: transfemoral. Adapted with permission from Vahanian *et al.* (18).

Whether aortic valve replacement is beneficial in asymptomatic patients with severe AS and normal left ventricular systolic function remains controversially discussed since decades. Whilst it was previously argued that intervention is not needed until symptoms supervene, because the risk of sudden cardiac death would be less than the risk of intervention, innovations in surgical and transcatheter techniques and concomitant reduction of procedural risks argue for early or elective replacement strategy instead of watchful waiting. The Aortic Valve ReplAcemenT versus Conservative Treatment in Asymptomatic SeveRe Aortic Stenosis (AVATAR) trial therefore randomized 157 asymptomatic patients (enrolled between 2015 and 2020) with severe AS and normal left ventricular function to either SAVR or conservative management according to current guidelines. Over a median follow-up of 32 months, patients randomized to early surgery had a significantly lower incidence of the primary composite endpoint of all-cause death, acute myocardial infarction, stroke or unplanned hospitalization for heart failure, as compared to patients allocated to the conservative treatment strategy (hazard ratio: 0.46 [95% confidence interval: 0.23-0.90], *p*-value: 0.021) (31).

Apart from clinical symptoms and severity of aortic valve obstruction, the decision to replace a stenotic aortic valve may also take the left ventricular response to AS-induced pressure overload into account. Considering the complex valvular-ventricular interaction is important since the pathophysiology of adverse outcomes in AS is essentially determined by the extent of imbalance between the increase in left ventricular hemodynamic burden due to obstruction of the aortic valve, on the one hand, and the compensation capacity of the left ventricle to overcome this increase in pressure (and later volume) overload, on the other hand. This means that patients with already impaired left ventricular systolic function due to ischemic heart disease are more vulnerable to any new insult such as rising left ventricular pressure overload. The Transcatheter Aortic Valve Replacement to UNload the Left ventricle in patients with ADvanced heart failure (TAVR UNLOAD) trial therefore aim to investigate whether afterload reduction by means of TAVR on top of optimal heart failure therapy can improve clinical outcomes in patients with moderate AS and reduced LVEF (32).

1.7   Selection of valve replacement strategy: on the emergence of transcatheter aortic valve replacement.

A multidisciplinary group of physicians involving interventional cardiologists, cardiac surgeons, imaging specialists, and anesthetists, referred to as "heart team", typically develops an individualized risk-benefit analysis of the available options for aortic valve replacement considering for instance coexisting coronary artery disease (potentially in need of bypass grafting), severe impediments to surgery such as a heavily calcified, fragile ("porcelain") aorta, and overall life expectancy. In fact, 30% of patients with severe symptomatic AS were denied surgery prior to the emerge of TAVR, and higher age and left ventricular dysfunction were the most striking characteristics of patients deemed inoperable (33).

Similarly to intracoronary artery stenting as a novel, disruptive technology to reduce the disease burden in patients with coronary artery disease and prohibitively high operative risk, TAVR upon its first in human implantation in 2002 (34) has been established as a less invasive treatment option for severe AS in otherwise inoperable patients (35,36). The underlying idea to insert a biological valve inside a large stent and to implant this valve using a transcatheter technique was developed in 1987, when Henning Andersen got inspiration from a lecture by Julio Palmaz, who had previously invented balloon-expandable coronary stents in animals (37). Equipped with wires made of iron and steel from the hardware store and with pig hearts from the slaughterhouse, the initial proof-of-concept was performed in 1989 by an uneventful first implantation of a prothesis prototype in an 80 kg pig. Apart from reducing mortality in inoperable patients with severe AS, the Placement of Aortic Transcatheter Valves (PARTNER) trial has demonstrated that TAVR was an equivalent treatment option to SAVR in high-risk patients with regards to mortality, reduction in symptoms, and improved valve hemodynamics (38). A later randomized trial involving intermediate-risk patients with severe AS (PARTNER 2 trial) has further shown TAVR and SAVR result in similar outcomes regarding death or disabling stroke even in intermediate-risk patients (39), additionally supporting the undeniable efficacy of TAVR. This is remarkable, because TAVR has thus created a paradigm shift in the treatment of patient with severe AS that was not anticipated by cardiac surgeons (37). Technological enhancements and procedural simplification have since then fueled the implementation of TAVR into clinical practice. The PARTNER 3 trial even demonstrates that among patients with severe AS, who were at low risk for death with surgery, the rate of the composite of death, stroke, or rehospitalization at one year was significantly lower with TAVR than with surgical aortic valve replacement (hazard ratio: 0.54 [95% confidence interval: 0.37-0.79], p-value: 0.001) (40). Also considering that patients undergoing TAVR benefit from faster

recovery, shorter hospital stay and rapid return to normal activities (41,42), TAVR at the time of writing is the preferred option in all patients ≥ 75 years of age regardless of the degree of surgical risk (18,43).

1.8 Prognosis: on the impact of extra-aortic valve cardiac damage.

Conventional perioperative risk evaluation and concomitant prognostic assessment for patients with severe AS is based on surgical risk scores mainly relying on valve-related factors, symptoms and co-morbidities, but because those risk models (such as European System for Cardiac Operative Risk Evaluation [EuroScore] score and Society of Thoracic Surgeons [STS] score) were not specifically developed from a patient population undergoing TAVR, their generalized application seems questionable (44). More recently, a bipartite assessment of patients with severe AS grading both AS severity (akin to grading the tumor in oncology) and staging extra-aortic valve cardiac (akin to staging the tumor extension in oncology) damage has been proposed by Généreux et al. (27). This anatomic and functional staging classification based on parameters measured during routine echocardiographic examination distinguishes five stages of disease progression with rising stages as assessed prior to TAVR translating into increased mortality (see Figure 9).

Stages/Criteria	Stage 0	Stage 1	Stage 2	Stage 3	Stage 4
	No Cardiac Damage	LV Damage	LA or Mitral Damage	Pulmonary Vasculature or Tricuspid Damage	RV Damage
Echocardiogram		Increased LV Mass Index >115 g/m² (Male) >95 g/m² (Female)	Indexed left atrial volume >34mL/m²	Systolic Pulmonary hypertension ≥60 mmhg	Moderate-Severe right ventricular dysfunction
		E/e' >14	Moderate-Severe mitral regurgitation	Moderate-Severe tricuspid regurgitation	
		LV Ejection Fraction <50%	Atrial Fibrillation		

Figure 9: Staging classification of aortic stenosis based on the extent of cardiac damage.

LA: left atrial; LV: left ventricular; RV: right ventricular. Adapted with permission from Généreux *et al.* (27).

This concept of staging cardiac repercussions of AS was also successfully validated in patients with moderate AS (45). Yet, this staging classification of cardiac damage assumes a sequential order of accumulated pathologies, hence ignoring the aggravating impact of comorbidities such as ischemic heart disease, chronic obstructive pulmonary disease, and atrial fibrillation. To determine whether the observed damage is caused by the AS *per se* or by a co-developed disease is particularly important in order to identify additional treatment targets beyond the correction of AS. The necessity to accurate capture the patient's cardiac status in its contextual structural and functional complexity is further emphasized by observational studies demonstrating that irreversibility of pulmonary hypertension and/ or right heart dysfunction after TAVR are associated with higher mortality (46–50).

1.9   Artificial intelligence: chances for a new age in cardiology?

In the dawning age of artificial intelligence, machine learning technology is progressively implemented into medical research and clinical decision support, and by performing tasks in a faster and potentially more precise fashion than humans (51–56), machine learning technology could pave the way to patient-tailored (precision) medicine. Importantly, the future of medicine is not about the competition of artificial intelligence against humans, but real-life medical practice will be coined by collaborative setups, where oversight-providing humans receive assistance from artificial intelligence. For instance, Chen *et al.* have developed an augmented reality microscope, where the predictions of an artificial intelligence algorithm on detecting cancer are superimposed on the view of the sample seen by the user through the eyepiece of an optical light microscope in real-time (57).

Artificial intelligence can be simply defined as a field of computer science that enables computers to perform tasks that normally require human intelligence. Whilst traditional artificial intelligence techniques were limited in their ability to process raw data and therefore required cumbersome engineering and considerable domain expertise to design feature extractors, modern machine learning algorithms allow to self-control their performance and make adjustments automatically by using the data itself and hence improving at a task with incremental experience ("learning through iteration") (see Figure 10).

Artificial intelligence
(the computer mimics human intelligence by applying simple rules for decision-making, such as if-then rules and decision trees)

Machine learning
(a subset of artificial intelligence including statistical techniques that enables machines to improve at tasks with incremental experience)

Deep learning
(a subset of machine learning that is based on multilayered "neural" networks and vast amounts of data to train themselves to perform tasks)
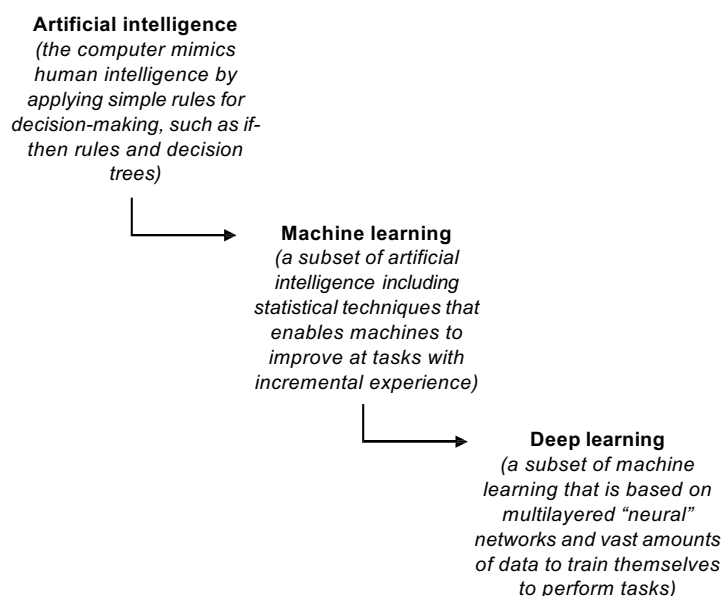
Figure 10: Domains of artificial intelligence.

For example, automation of electrocardiography interpretation was traditionally performed by computers mimicking human intelligence in terms of applying rules-based algorithms. Such rules were explicitly

programmed by humans and intended to determine the heart rate, to detect the underlying rhythm, and to describe any pathognomonic waveform changes (58). Since the input features and subsequent rules to apply on were selected and designed by humans, the resulting diagnosis from the computer was easy to understand. At the same hand, those traditional artificial intelligence techniques were limited to predict human-defined outputs, potentially neglecting the complexity of a disease.

Modern machine learning algorithms, on the other hand, are particularly useful to analyze highly complex data structures without the constraints of applying any *a priori* hypotheses or reducing the data set to a few input features. Mainly two disciplines of machine learning can be distinguished:

1) In unsupervised machine learning, the computer aims at recognizing any consistent patterns within a data set.

2) In supervised machine learning, data need to be labelled and the computer predicts a label of interest.

1.9.1   On the principles of unsupervised machine learning for pattern recognition.

Unsupervised machine learning resembles natural human learning in a way that humans discover the structure of the world by observing and analyzing it (59). Unsupervised machine learning approaches include dimensionality reduction techniques such as principal component analysis and clustering methods such as *k*-means or agglomerative, hierarchical clustering (60). Unsupervised machine learning requires only minimal human interaction, as it is unconstrained by labor-intensive and often variable human labelling. As a consequence, the strength of unsupervised machine learning lies in its ability to identify novel relationships in the data that are potentially unrecognizable to humans. For instance, unsupervised machine learning algorithms have been successfully applied to data sets composed of clinical variables concerning patients with severe AS and automatically identified distinct disease phenotypes (55,61). Importantly, cluster analysis often represents only the first step to group the data, and there exists no a priori guarantee that one unsupervised machine learning technique is superior to all others (62). In order to avoid that the clustering approach results in lumping together patient groups that should be separated, distinct unsupervised machine learning techniques must be tested, and the emerging clusters must be validated with clinical domain knowledge (63). As a result, unsupervised machine learning approaches hold the promise to detect novel phenotypes in complex diseases, which differ in underlying pathophysiology and are associated with different response to

treatment and/ or outcome. By eliciting those formerly hidden associations in data sets, machine learning has the potential to pave the way to precision medicine (64).

1.9.2   On the principles of supervised machine learning as prediction models.

Supervised machine learning is employed to predict an outcome of interest. Based on a training set with labelled instances (requiring considerable human efforts to create), supervised machine learning algorithms autonomously select input features and iteratively weights them to identify the best combination to predict the outcome of interest. After learning and self-constructing the rules for decision-making, the trained model can be applied to never-before-seen data for validation. In contrast to conventional statistics such as linear or logistic regression models, supervised machine learning employing random forests, extreme gradient boosting algorithms or artificial neural networks rely on fewer assumptions, and they have proven particularly useful to decipher complex, non-linear associations between input data and outcome. For instance, supervised machine learning algorithms have been proven to successfully diagnose genetic disorders based on facial phenotypes (65) – data input that is hardly possible to be analyzed by regression analysis, but can be processed by artificial neural networks. Those networks mimic the human cortex: digitized input is processed through multiple layers of artificial neurons until an output is ultimately provided (see Figure 11). Like biological neurons, artificial neurons are interconnected. Each artificial neuron receives incoming signals from multiple neurons, processes the information, and transmits the signal to the next layer (it "fires" similar to the action potential of a biological neuron). By employing rectified linear unit activation functions for neurons in the hidden layers, the strength of their output signal is sensitive to the activation input sum and not subjected to any saturation effects. Importantly, signals are only passed further if the aggregated incoming signals cross a threshold. Weighting the incoming signals is a crucial part of the autodidactic learning aspect of artificial neural networks during training – successive adjustments ultimately enable the artificial neural network to generate outputs that are increasingly similar to the target output (59,66).

Nature

A

B

C

Machine

D

E

F
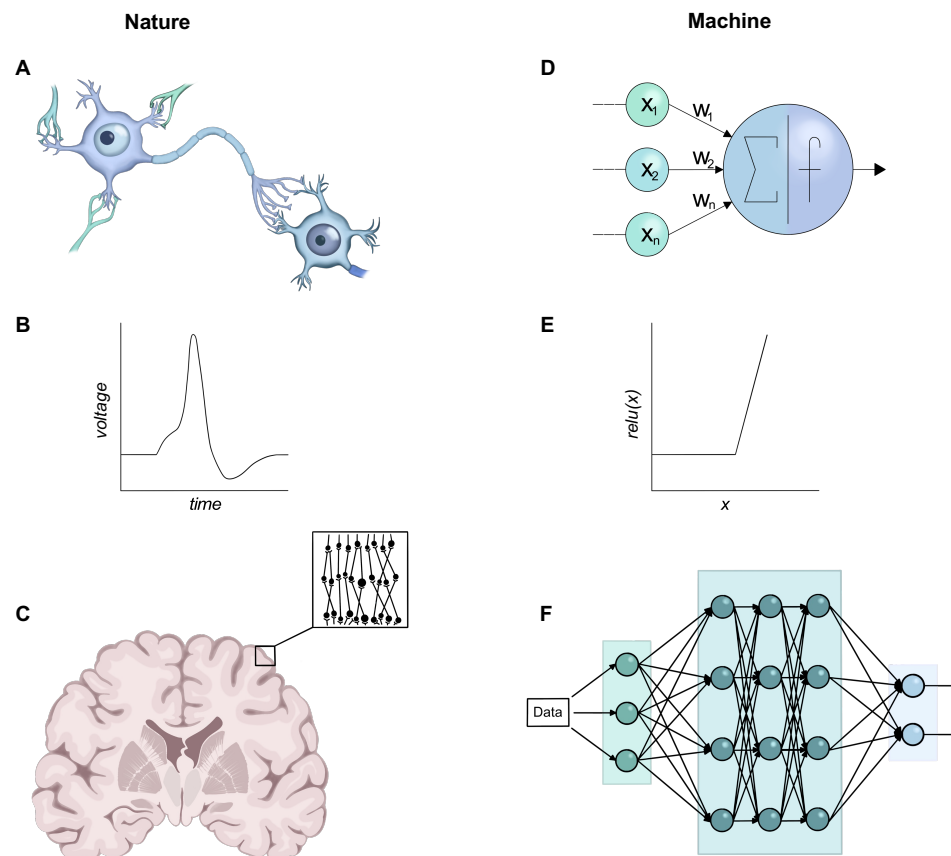
Figure 11: From biological to artificial neurons.

A biological neuron (A) is composed of a cell body containing the nucleus, many branching extensions called dendrites, and one very long axon, which splits off into many branches with synapses – which are in turn connected to the dendrites or cell bodies of other neurons to transmit signals via the release of neurotransmitters. When a neuron receives a sufficient amount of excitatory neurotransmitters within a few milliseconds, the voltage gradient across its membrane changes (B), and the neuron generates a short, all-or-nothing electrical impulse called action potential. This potential travels along the axon and activates synaptic connections as it reaches them. Each neuron being connected to thousands of other neurons, often organized in consecutive layers as in the human cerebral cortex (C), allows for highly complex information processing. The artificial neuron is inspired by the biological neuron: it has one or more inputs and generates one output (D). When the weighted sum of its inputs surpasses a threshold (E), the neuron is activated, and it produces a signal to the next layer. Whilst Mother Nature had chosen to use roughly sigmoid activation functions in neurons (all-or-nothing law), rectified linear unit activation functions perform better in artificial neural networks – mainly because they do not saturate for positive values. Connecting multiple layers of artificial neurons enables information processing similar to the human brain (F). An artificial neural network with a deep stack of hidden layers is referred to as a deep neural network - those are particularly powerful in biosignal, image and natural language processing. Artwork from Amelie Hesse (printed with permission).

1.9.3 Explainable artificial intelligence to fill the gap between predictive power and transparency of machine learning algorithms.

Besides predicting the outcome of interest, supervised machine learning techniques also offer the possibility to explore the relationships between input data and a specific outcome. To explain the determination of output is extremely important, because a systematically flawed "black box" algorithm has tremendous potential for medical malpractice. The European Union's General Data Protection Regulation therefore demands that the output generated by an algorithm must be comprehensible before the algorithm can be implemented into patient care (67). Based on classic game-theoretic Shapley values, Lundberg *et al.* have developed a method that allows global model insights by combining local explanations of model predictions (68). By informing clinicians not only about the probability of an event, but also about the factors that are underlying this prediction, explainable artificial intelligence offers clinicians a target to treat the patient and possibly avert the anticipated event (69).

2 Methods

All statistical and machine learning analyses were performed using R statistical software (R version 3.6.3; R Foundation for Statistical Computing, Vienna, Austria). All computations were performed on a MacBook Pro (macOS Catalina version 10.15.5, Apple Computer) with a 2.3-GHz quad-core Intel Core i7 processor.

3   Objectives

Acknowledging that patients with severe AS present with vast heterogeneity in terms of disease progression and comorbidities, this dissertation thesis employed a combination of unsupervised and supervised machine learning techniques to refine the understanding of phenotypic complexity in patients with severe AS.

1) *"Harnessing feature extraction capacities from a pre-trained convolutional neural network (VGG-16) for the unsupervised distinction of aortic outflow velocity profiles in patients with severe aortic stenosis"*

Since the human brain can process only a finite number of parameters, a lot of information in raw data remain unrecognized. Harnessing feature extraction capacities from an artificial neural network that mimics the human visual cortex, but has sixteen instead of six neural layers, holds the promise to refine interpretation of e.g., aortic outflow velocity profiles. Those profiles represent a crucial part in the echocardiographic process to diagnose a severe AS. Moreover, those profiles contain not only information about obstruction of the aortic valve, but they also provide information about left ventricular contractility. Aortic outflow velocity profiles were therefore presented to an established artificial neural network, which allocated them to two distinct clusters. Cluster-related survival differences after TAVR were hereinafter evaluated, ultimately addressing the question whether a computer can extract relevant information from one single echocardiographic image to stratify patients into high-risk and low-risk cohorts for mortality after TAVR.

2) *"Subphenotyping of patients with aortic stenosis by unsupervised agglomerative clustering of echocardiographic and hemodynamic data"*

Apart from the fact that severe AS can trigger a deleterious cascade of extra-aortic valve damage to the heart and pulmonary circulation, patients with severe AS typically suffer from comorbidities such as coronary artery disease, chronic obstructive pulmonary disease, and atrial fibrillation, which in turn can cause additional impairments to the cardiac structure and function. Unsupervised agglomerative clustering was therefore applied to preprocedural data from echocardiography and right heart catheterization to detect phenotypes among patients with severe AS sharing similar extents of cardiac damage. In comparison to traditional hypothesis-driven staging classifications, this machine learning

approach comes with the advantage not to infer causality between severity of AS and observed cardiac damage, but patients are segregated into clusters based on real-world data. Thus, this approach promises to better capture the patient's cardiac status in its structural and functional complexity. By comparing survival rates after TAVR in accordance with cluster assignment, our phenotyping approach based on unsupervised machine learning was tested to provide clinically relevant information.

4 Results

4.1 Summarized publication #1: *"Harnessing feature extraction capacities from a pre-trained convolutional neural network (VGG-16) for the unsupervised distinction of aortic outflow velocity profiles in patients with severe aortic stenosis"*

Severity of AS is typically assessed by transthoracic echocardiography. The aortic outflow velocity profile contains valuable information about both aortic valve obstruction and left ventricular contractility. For instance, an increasing transvalvular pressure gradient can be observed in the natural course of progressive narrowing of the aortic valve, until acceleration of flow velocity eventually deteriorates due to left ventricular decompensation, resulting in a LFLG AS. Moreover, the shape of the aortic outflow velocity profile changes from a triangular shape with an early peak to a much more rounded form with a later peak, reflecting slow end-systolic opening of the stenotic aortic valve combined with left ventricular dysfunction.

Notably, the human brain can process only a finite number of parameters. Previous attempts to extract information from the aortic outflow velocity profile therefore focused on single characteristics and were hence prone to oversimplification. This study therefore sought to employ unsupervised machine learning techniques to decipher meaningful echocardiographic signatures and related cardiac phenotypes by analyzing aortic outflow velocity profiles from selected patients with severe AS undergoing TAVR. Importantly, the superhuman performances of machine learning algorithms are typically based on a myriad of information, so that the computer can iteratively learn from the data itself. However, bedside research embedded in clinical practice is commonly based on modest-sized patient cohorts. This study therefore applied the concept of transfer learning: a convolutional neural network (VGG-16) was trained on a large data set of publicly available images (ImageNet data set; 14 million images belonging to 1,000 classes), and the convolutional part of the pre-trained VGG-16 model was subsequently employed to analyze aortic outflow velocity profiles from a small, but well-characterized cohort of patients with severe AS. As a matter of fact, the number of enrolled patients was limited to 101 patients to emphasize the ubiquitous problem of data scarcity. After principal component analysis and *k*-means clustering, practice-relevant evidence was assessed by relating cluster assignment with 2-year all-cause mortality after TAVR.

Among 101 patients (mean age: 79.3 ± 6.78 years; 49 [48.5%] women) constituting the study population at hand, 2-year survival following TAVR was 83.0% (95% CI 75.1-91.7%). Patients initially presented with an AVA of 0.804 ± 0.223 cm^2, and predominantly suffered from dyspnoea corresponding to New York Heart Association functional class III (56.4%).

The convolutional part of the pre-trained VGG-16 model in conjunction with principal component analysis and *k*-means clustering of the abstractions of Doppler tracings enabled to distinguish two shapes of aortic outflow velocity profiles. Interestingly, all patients from cluster 2 presented with a AVG_{mean} below 40 mmHg, whilst AVG_{mean} from patients in cluster 1 ranged between 20 and 102 mmHg. Kaplan-Meier analysis revealed that mortality in patients from cluster 2 (*n* = 40 [39.6%]) was significantly increased (hazard ratio for 2-year mortality: 3 [95% confidence interval: 1-8.9]). Besides reduced cardiac output (4.57 ± 1.42 L/min) and signs of pulmonary hypertension (mean pulmonary artery pressure [mPAP]: 31.9 ± 12.2 mmHg), patients from cluster 2 also presented with more severe impairment of right ventricular function (tricuspid annular plane systolic excursion [TAPSE]: 18.1 ± 3.82 mm) and right atrial enlargement (right atrial area: 22.0 ± 8.28 cm^2) in comparison to patients from cluster 1. Contrarily to the initial expectation, patients from cluster 1 with seemingly less extensive cardiac damage were diagnosed with a more severe obstruction of the aortic valve than patients from cluster 2 (AVA: 0.739 ± 0.211 cm^2 vs. 0.903 ± 0.205 cm^2, *p*-value: 0.0001).

In conclusion, this study demonstrates that transfer learning enables sophisticated pattern recognition in a clinical data set of limited size. Relying on good quality Doppler tracings and harnessing the intriguing feature extraction capacity from an established convolutional neural network, agnostic interpretation of aortic flow velocity profiles from patients with severe AS revealed that not so much the actual stenosis of the aortic valve expressed as AVA determines the prognosis after TAVR, but the left ventricular compensation capacity and subsequent development of pulmonary hypertension and right heart failure stratify patients into low-risk and high-risk cohorts.

Contribution from Elena Rippen:

1) Data acquisition, cleaning, and curation.

2) Manual cropping and scaling of aortic outflow velocity profiles.

3) Conception of the study and interpretation of data (together with Mark Lachmann).

4) Drafting of the manuscript (together with Mark Lachmann).

4.2   Summarized publication #2: *"Subphenotyping of patients with aortic stenosis by unsupervised agglomerative clustering of echocardiographic and hemodynamic data"*

Previous staging classifications to stratify patients with AS hypothesized a sequential order of accumulated extra-aortic valve pathologies such as left ventricular dysfunction, mitral regurgitation and left atrial enlargement, pulmonary hypertension, and ultimately right heart failure (27,45). Such an approach infers causality between AS and co-existing pathologies, and it hence ignores the aggravating impact of comorbidities such as ischemic heart disease, atrial fibrillation and chronic obstructive pulmonary disease on cardiac function and structural integrity. This is particularly problematic, because a plethora of contributors to right heart dysfunction, including coronary artery disease and subsequent myocardial ischemia as well as chronic obstructive pulmonary disease and secondary pulmonary hypertension, will persist despite correction of severe AS and hence limit the expected benefit of TAVR. A novel classification system based on unsupervised agglomerative clustering in conjunction with an artificial neural network was therefore established to comprehensively capture the patient's cardiac status in its contextual structural and functional complexity.

Unsupervised agglomerative clustering was applied to pre-procedural data from echocardiography and right heart catheterization from 366 consecutively enrolled patients undergoing TAVR for severe AS at two tertiary centers in Germany between 2014 and 2020.

Cluster analysis revealed four distinct phenotypes, reflecting various pathophysiologies and extents of disease progression, and ultimately differing in prognosis following TAVR.

Patients in cluster 1 (n = 164) presented with preserved cardiac function (LVEF: 57.2 ± 6.36%) and normal pulmonary artery pressures (mPAP 21.2 ± 6.54 mmHg). Kaplan-Meier analysis showed that 2-year survival from patients in cluster 1, hereinafter referred to as reference, ranged high at 90.6%. In contrast, patients in cluster 2 (n = 66) featured elevated pulmonary artery pressures (mPAP: 34.2 ± 7.76 mmHg) but a still preserved right ventricular systolic function (TAPSE: 21.3 ± 3.52 mm) and a low rate of severe tricuspid regurgitation (2 out of 66 patients; 3.03%). Given the isolated postcapillary nature of pulmonary hypertension in cluster 2, which is considered reversible upon resolving the underlying AS, patients in cluster 2 had a 2-year survival rate not statistically different from that of patients in cluster 1 (2-year survival 85.8% [95% confidence interval: 76.9%-95.6%]; hazard ratio for 2-year mortality: 1.5 [95% confidence interval: 0.6-3.6]). Left heart failure (LVEF: 42.4 ± 15.7%), severe pulmonary hypertension (mPAP: 46.9 ± 8.54 mmHg), and right heart dysfunction (TAPSE 16.1 ± 4.57 mm)

characterized patients in cluster 3 (n = 45). A classical LFLG AS was found in 21 out of 45 patients (46.7% vs. 6.10%, 4.55% and 28.6% in clusters 1, 2 and 4, respectively; p-value: 6.1×10^{-13}). Moreover, 24.4% of patients from cluster 3 were diagnosed with chronic obstructive pulmonary disease (in comparison to 8.54%, 15.2% and 15.4% in clusters 1, 2 and 4, respectively; p-value: 0.0376). Subsequently, their 2-year mortality was significantly increased (2-year survival: 77.3% [95% confidence interval: 65.2-91.6%], hazard ratio for 2-year mortality: 2.6 [95% confidence interval: 1.1-6.2]). Interestingly, patients from cluster 4 (n = 91) showed mild postcapillary pulmonary hypertension (mPAP: 27.5 ± 9.15 mmHg), yet dilatation of all heart chambers, biventricular dysfunction (LVEF: $47.3 \pm 12.2\%$; TAPSE: 16.8 ± 4.46 mm), and a high prevalence of both mitral and tricuspid regurgitation (12.1% and 14.3%, respectively). Besides a high prevalence of atrial fibrillation and/ or flutter (75.8%), patients from cluster 4 also presented with highest age at diagnosis (81.3 ± 6.73 years vs. 79.6 ± 5.96, 78.1 ± 6.91 and 79.8 ± 8.73 years in clusters 1, 2 and 3, respectively; p-value: 0.0068). Patients from cluster 4 featured the worst prognosis (2-year survival: 74.9% [95% confidence interval: 65.9-85.2%], hazard ratio for 2-year mortality: 2.8 [95% confidence interval: 1.4-5.5]).

Taken together, artificial intelligence-enabled phenotyping sheds light from a new perspective on the complex, non-linear accumulation of extra-aortic valve cardiac damage in patients with severe AS. In contrast to traditional, hypothesis-driven classification systems, the proposed phenotyping approach does neither hypothesize a linear sequence of accumulated pathologies (thus it incorporates the aggravating impact of comorbidities), nor does it stratify patients into low-risk and high-risk cohorts in accordance with a single variable's dichotomy (thus it reduces the risk of oversimplification). Addressing irreversibility of pulmonary hypertension and persistence of severe tricuspid regurgitation after TAVR should obtain priority in order to improve long-term survival.

Contribution from Elena Rippen:

1) Data acquisition, cleaning, and curation.
2) Conception of the study and interpretation of data (together with Mark Lachmann).
3) Drafting of the manuscript (together with Mark Lachmann).
4) Assisting in revising the manuscript according to the reviewers' comments (together with Mark Lachmann).

5   Discussion

On the recovery of extra-aortic valve cardiac damage after transcatheter aortic valve replacement.

Considering that the extra-aortic valve cardiac damage is not necessarily related to the severity of AS, it remains questionable if the extra-aortic valve cardiac damage can be completely reversed upon TAVR. Contributors to right heart dysfunction such as chronic obstructive pulmonary disease and secondary pulmonary hypertension will persist despite correction of severe AS, and possibly limit the beneficial effect of TAVR. In a follow-up study we have therefore analyzed the recovery of extra-aortic valve cardiac damage in accordance with our artificial intelligence-based phenotyping approach (70). We could show that TAVR elicited favorable effects on left heart hemodynamics and significantly ameliorated pulmonary artery pressures in all patients (Table 2).

	Before (n = 366)	After (n = 247)	p-value
LVEF, mean ± SD [95% CI], %	52.7 ± 11.1 [51.6-53.9]	53.2 ± 9.85 [52.0-54.4]	0.66
sPAP, mean ± SD [95% CI], mmHg	47.2 ± 15.8 [45.4-49.1]	43.3 ± 15.1 [41.2-45.4]	0.0079
Right midventricular diameter, mean ± SD [95% CI], mm	29.5 ± 6.60 [28.8-30.2]	29.9 ± 6.00 [29.1-30.8]	0.088
TAPSE, mean ± SD [95% CI], mm	19.6 ± 5.02 [19.1-20.2]	19.8 ± 5.14 [19.2-20.5]	0.72
LA area, mean ± SD [95% CI], cm²	26.3 ± 8.29 [25.4-27.2]	26.3 ± 7.86 [25.3-27.3]	0.12
RA area, mean ± SD [95% CI], cm²	20.5 ± 7.47 [19.7-21.4]	20.9 ± 7.94 [19.9-22.0]	0.86
MR ≥ III/IV°, No (%)	34 (9.29%)	9 (3.64%)	0.0015
TR ≥ III/IV°, No (%)	33 (9.02%)	25 (10.1%)	0.53

Table 2: Comparison of echocardiographic follow-up data before and after transcatheter aortic valve replacement.

Comparison was calculated by paired-samples Wilcoxon test.

CI: confidence interval; LA area: left atrial area; LVEF: left ventricular ejection fraction; MR: mitral regurgitation; RA area: right atrial area; SD: standard deviation; sPAP: systolic pulmonary artery pressure; TAPSE: tricuspid annular plane systolic excursion; TR: tricuspid regurgitation.

Yet, a detailed look into the recovery of extra-aortic valve cardiac damage per cluster revealed that structural and functional alterations of the right heart persisted in patients assigned to high-risk clusters 3 and 4, which were characterized by a high prevalence of chronic obstructive pulmonary disease and atrial fibrillation and/ or flutter (Figure 12).
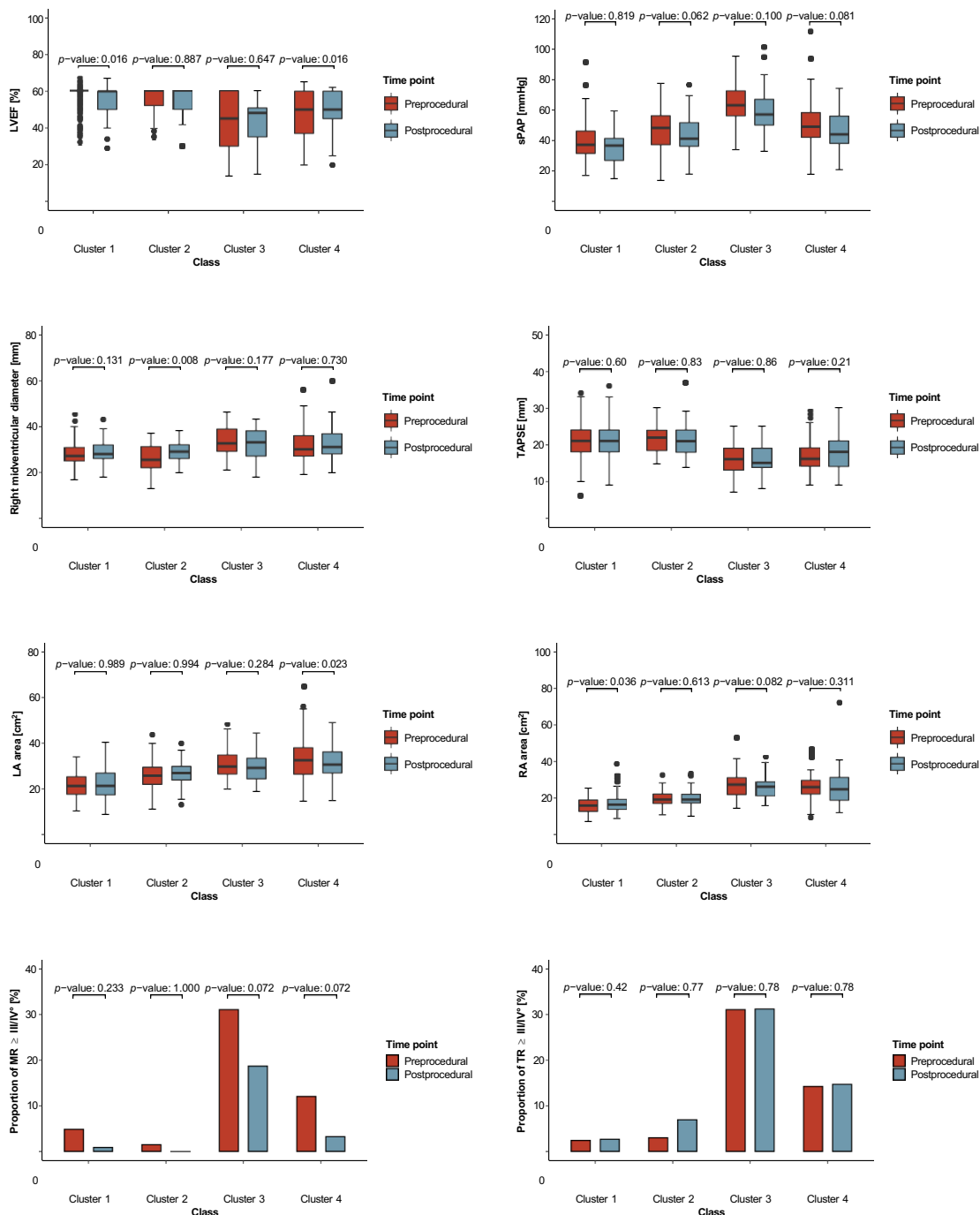
Figure 12: Comparison of echocardiographic parameters before and after transcatheter aortic valve replacement in accordance with cluster assignment.

LA area: left atrial area; LVEF: left ventricular ejection fraction; MR: mitral regurgitation; RA area: right atrial area; sPAP: systolic pulmonary artery pressure; TAPSE: tricuspid annular plane systolic excursion; TR: tricuspid regurgitation.

Whilst previous studies had already described that right heart dysfunction can persist in a number of cases after TAVR (49,71), and that postprocedural decline of right ventricular function and/ or worsening of tricuspid regurgitation are associated with a poor prognosis (50,72), our artificial intelligence-based phenotyping approach predicts the trajectory of cardiopulmonary recovery *ex ante*, and it might thus refine the prognostic assessment. Moreover, since right-sided cardiac damage including severe tricuspid regurgitation persisted after TAVR and ultimately limited prognosis, our follow-up study emphasizes the need for additional treatments such as transcatheter tricuspid valve interventions.

On the generalizability of the proposed artificial intelligence-enabled phenotyping approach to further valvular heart diseases.

Similar to patients with severe AS, patients with severe mitral regurgitation also present with considerable heterogeneity depending on underlying etiology, extent of disease progression and comorbidities. We therefore adapted our artificial intelligence-enabled phenotyping approach to patients with severe mitral regurgitation undergoing transcatheter edge-to-edge repair (73). Among 609 patients undergoing mitral valve transcatheter edge-to-edge repair for both primary and secondary mitral regurgitation, four pathophysiologically novel and prognostically informative phenotypes were unraveled by unsupervised agglomerative clustering (Figure 13):

1) Cluster 1 was constituted by patients with isolated defects of the mitral valve and without significant extra-mitral valve cardiac damage - meaning that they presented with preserved left ventricular function (LVEF: 56.5 ± 7.79%) and regular left ventricular end-systolic diameter (35.2 ± 7.52 mm). Hereinafter serving as a reference, 5-year survival in patients from cluster 1 was 60.9%.

2) Cluster 2 was similar to cluster 1, but patients had already progressed to developing pulmonary hypertension (systolic pulmonary artery pressure levels: 68.4 ± 16.2 mmHg). Even though mitral valve transcatheter edge-to-edge repair resulted in a significant amelioration of systolic pulmonary artery pressure levels in these patients at 1-year follow-up (49.6 ± 18.1 mmHg), their

5-year survival ranged at 43.7% and was significantly lower compared to that from cluster 1 - emphasizing the importance of adequate timing of intervention.

3) Cluster 3 was characterized by left (and right) ventricular failure in terms of chamber dilatation and functional impairment - most of these patients would have been classified as typical secondary mitral regurgitation. Their 5-year survival was reduced to 38.3%.

4) Poorest 5-year survival (23.8%) was observed in patients from cluster 4 presenting with biatrial dilatation. All patients from cluster 4 were diagnosed with atrial fibrillation, possibly sustaining a vicious circle of mitral regurgitation, atrial enlargement, and atrial fibrillation.

Unsupervised machine learning-enabled phenotyping of cardiac damage in patients with mitral regurgitation undergoing mitral valve transcatheter edge-to-edge repair

Cluster 1: *isolated mitral regurgitation*	Cluster 2: *secondary pulmonary hypertension*	Cluster 3: *failing left ventricle*	Cluster 4: *biatrial dilatation*
LVEF: 56.5 ± 7.79%	LVEF: 55.7 ± 7.82%	**LVEF: 31.0 ± 10.4%**	LVEF: 51.5 ± 11.0%
LVESD: 35.2 ± 7.52 mm	LVESD: 34.9 ± 7.68 mm	**LVESD: 53.2 ± 10.9 mm**	LVESD: 41.0 ± 13.4 mm
MV EROA: 0.393 ± 0.153 cm²	**MV EROA: 0.623 ± 0.360 cm²**	MV EROA: 0.287 ± 0.123 cm²	MV EROA: 0.438 ± 0.176 cm²
LA volume: 117 ± 38.3 mL	LA volume: 149 ± 48.6 mL	LA volume: 142 ± 44.5 mL	**LA volume: 312 ± 113 mL**
sPAP: 42.0 ± 11.1 mmHg	**sPAP: 68.4 ± 16.2 mmHg**	sPAP: 48.8 ± 14.6 mmHg	sPAP: 48.9 ± 10.3 mmHg
TAPSE: 20.9 ± 4.20 mm	TAPSE: 17.0 ± 5.27 mm	TAPSE: 14.6 ± 4.11 mm	TAPSE: 17.1 ± 5.09 mm
RA area: 22.7 ± 5.84 cm²	RA area: 28.0 ± 6.85 cm²	RA area: 29.0 ± 10.3 cm²	**RA area: 46.0 ± 8.83 cm²**
5-year survival *60.9%* *(95% CI: 53.3−69.7%)*	*5-year survival* *43.7%* *(95% CI: 33.2−57.6%)*	*5-year survival* *38.3%* *(95% CI: 31.9−46.1%)*	*5-year survival* *23.8%* *(95% CI: 12.8−44.3%)*

Figure 13: Artificial intelligence-enabled phenotyping approach facilitates to capture the complexity of cardiac damage as commonly encountered in patients presenting with mitral regurgitation.

LA volume: left atrial volume; LVEF: left ventricular ejection fraction; LVESD: left ventricular end-systolic diameter; MV EROA: mitral valve effective regurgitant orifice area; RA area: right atrial area; sPAP: systolic pulmonary artery pressure; TAPSE: tricuspid annular plane systolic excursion.

Importantly, it was also confirmed in patients with mitral regurgitation that correction of the valvular defect reduces the hemodynamic burden imposed to the pulmonary circulation expressed as systolic pulmonary artery pressure levels, but right ventricular dysfunction as initially observed in patients from cluster 3 did not recover after mitral valve transcatheter edge-to-edge repair, nor could pathological remodeling obtained as biatrial dilatation in patients from cluster 4 be reversed (Table 3, Figure 14).

	Class				
	Cluster 1 (n = 139)	Cluster 2 (n = 71)	Cluster 3 (n = 150)	Cluster 4 (n = 20)	*p*-value
LVEF, %	54.5 ± 8.70[3]	54.1 ± 8.51[3]	33.7 ± 12.5[1,2,4]	49.8 ± 13.9[3]	<2.2x10^{-16}
LVEDD, mm	51.3 ± 7.23[3]	51.0 ± 6.64[3]	63.5 ± 10.6[1,2,4]	55.4 ± 9.43[3]	<2.2x10^{-16}
LVESD, mm	35.2 ± 7.70[3]	34.0 ± 7.87[3]	54.0 ± 13.3[1,2,4]	40.9 ± 13.4[3]	<2.2x10^{-16}
LVEDV, mL	122 ± 46.5[3]	103 ± 39.4[3]	203 ± 85.1[1,2]	127 ± 15.9	1.0x10^{-10}
LVESV, mL	64.1 ± 34.7[3]	54.2 ± 26.1[3]	149 ± 83.8[1,2]	99.7 ± 62.6	1.2x10^{-12}
MV gradient, mmHg	3.40 ± 1.45	3.51 ± 1.64	2.94 ± 1.47	3.39 ± 1.04	0.0334
sPAP, mmHg	42.2 ± 14.7	49.6 ± 18.1	42.1 ± 12.5	45.8 ± 17.3	0.0668
TAPSE, mm	19.1 ± 3.99[2,3,4]	17.6 ± 5.41[1]	16.2 ± 4.16[1]	15.9 ± 4.24[1]	3.5x10^{-5}
RV diameter (midventricular), mm	29.9 ± 5.73	31.9 ± 5.99	31.5 ± 6.61	36.8 ± 12.4	0.2037
LA volume, mL	109 ± 42.0[2,3,4]	156 ± 71.3[1,4]	130 ± 46.0[1,4]	316 ± 132[1,2,3]	1.4x10^{-7}
RA area, cm²	24.3 ± 7.84[4]	27.3 ± 8.57[4]	25.2 ± 7.29[4]	42.3 ± 9.50[1,2,3]	0.0004
TAPSE/sPAP ratio, mm/mmHg	0.530 ± 0.321[2,3]	0.404 ± 0.214[1]	0.427 ± 0.188[1]	0.398 ± 0.157	0.0127

Table 3: Echocardiographic 1-year follow-up data.

Continuous variables are given as means ± standard deviation. Numbers in parentheses indicate between which clusters significant differences (p-value ≤ 0.05) were detected.

No differences in MV gradient after correction for multiple testing.

LA volume: left atrial volume; LVEDD: left ventricular end-diastolic diameter; LVEDV: left ventricular end-diastolic volume; LVEF: left ventricular ejection fraction; LVESD: left ventricular end-systolic diameter; LVESV: left ventricular end-systolic volume; MV gradient: mitral valve gradient; RA area: right atrial area;

RV diameter (midventricular): right midventricular diameter; sPAP: systolic pulmonary artery pressure; TAPSE: tricuspid annular plane systolic excursion.
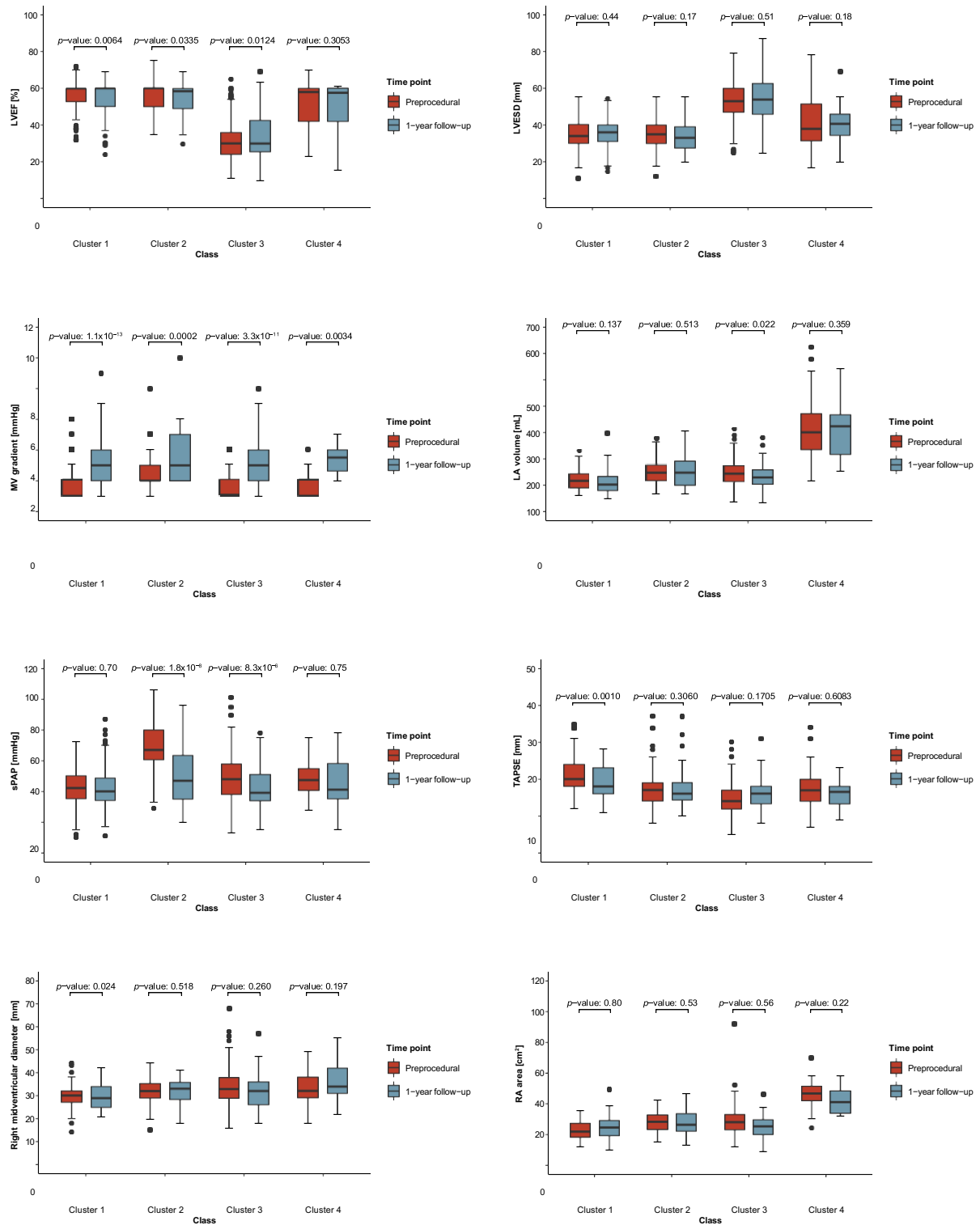
Figure 14: Comparison of echocardiographic parameters before and after mitral valve transcatheter edge-to-edge repair in accordance with cluster assignment (derivation cohort).

LA volume: left atrial volume; LVEF: left ventricular ejection fraction; LVESD: left ventricular end-systolic diameter; MV gradient: mitral valve gradient; RA area: right atrial area; sPAP: systolic pulmonary artery pressure; TAPSE: tricuspid annular plane systolic excursion.

Upon training an artificial neural network to facilitate future patient-to-cluster assignment, the prognostic value of our artificial intelligence-enabled phenotyping approach with regards to 5-year all-cause

mortality was externally confirmed in 817 patients from two independent institutions, which is to be considered as another important milestone before clinical implementation (Figure 15).
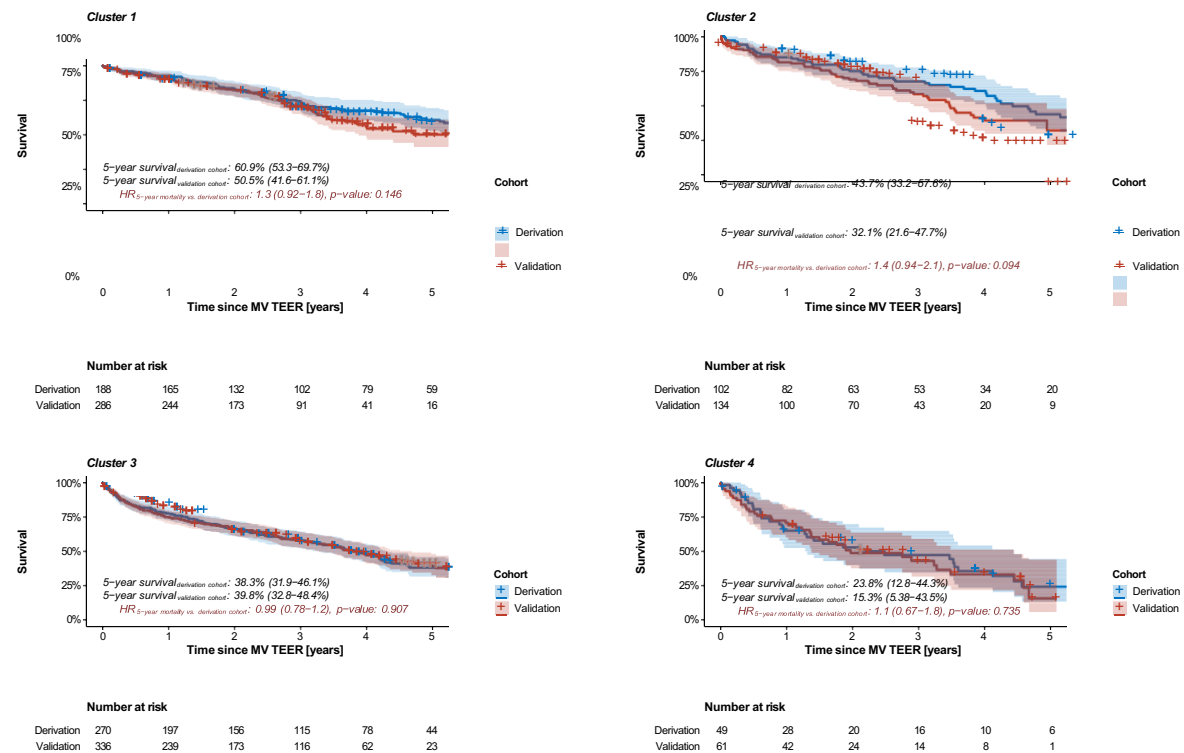

Figure 15: External validation of the proposed machine learning-based phenotyping approach in patients with severe mitral regurgitation undergoing transcatheter edge-to-edge repair.

Kaplan-Meier survival plots comparing cluster-related survival between patients from the derivation and the external validation cohort.

On man-machine interactions for future patient-to-cluster assignment.

The future of medicine is not about the competition of artificial intelligence against humans, but real-life medical practice will be coined by collaborative setups, where oversight-providing humans receive assistance from artificial intelligence (74). Our studies provide evidence on the usefulness of machine learning algorithms in several valvular heart diseases such as AS, mitral regurgitation, and tricuspid regurgitation (26,70,73,75,76). To facilitate their clinical implementation, the underlying codes must be made publicly available. Since our artificial neural network to assign patients with severe AS to one of the four proposed subphenotypes demonstrated an excellent performance to detect patients from high-risk clusters 3 and 4 (100.0% and 85.2% sensitivity, as well as 95.9% and 95.1% specificity, respectively) (76), we decided to export our trained artificial neural network and to embed it into a simplified code that allows future patient-to-cluster assignment (see Figure 16) (70). Currently regarded as a prototype, our

man-machine interaction-based classification model could serve as an online-based decision support

tool in the future. Feeling that we as clinicians should strive to use the best (not the simplest) model available to guide treatment decisions, this would open the avenue for other cardiologists to stratify their patients according to our classification system.
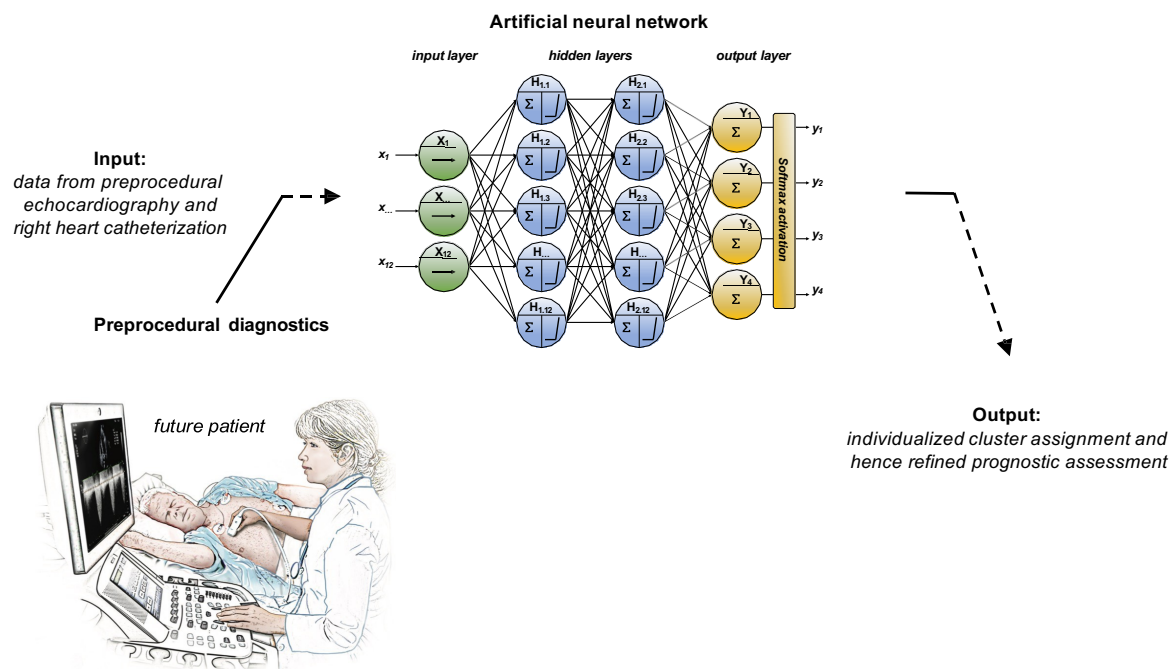

Figure 16: On the future man-machine interaction to stratify patients with severe aortic stenosis according to the proposed artificial intelligence-enabled phenotyping approach.

This schematic illustration demonstrates how we envision the future of cardiology: a future patient with severe aortic stenosis would be examined by echocardiography and right heart catheterization, and the measured parameters would be given to the artificial neural network as input data, which in turn calculates an individualized cluster assignment with concomitant prognosis.

A control group of conservatively treated patients with severe aortic stenosis, assigned to beforehand defined clusters, could have quantified life-extending effects of transcatheter aortic valve replacement in accordance with cluster assignment.

It is inherent to the nature of this retrospective, observational study that no definitive conclusions can be drawn regarding the benefit of TAVR in accordance with cluster assignment, as a (randomized) conservative treatment group as control was missing. Even though 2-year survival rates from patients in high-risk clusters 3 and 4 ranged at low levels (77.3% and 74.9%, respectively), their survival was still better compared to conservatively managed patients featuring 2-year survival rates below 50% (30). Thus, it can be said that the beneficial effect of TAVR in patients from clusters 3 and 4 was limited

compared to patients from clusters 1 and 2, but for sure it was not in vain. For the future, we believe that it would be of interest to additionally apply our man-machine interaction-based phenotyping approach to a cohort of conservatively treated patients in order to ultimately quantify the net survival benefit between treatment groups in accordance with cluster assignment.

6   Conclusion

In this thesis focusing on the employment of artificial intelligence in patients with severe AS, it was shown that smart collaborative setups converging human and machine intelligence can enhance our understanding of often complex clinical presentations. Not so much the obstruction of the aortic valve, which can be resolved by TAVR, but the (irreversible) extra-aortic valve cardiac damage determines prognosis and therefore demands to be properly captured. Our artificial intelligence-enabled phenotyping approach has a clinical consequence as it prompts timely correction before the fixation of irreversible extra-aortic valve damage, and it emphasizes the need for additional treatments such as transcatheter tricuspid valve interventions to improve survival in high-risk clusters.

7 References

1. Lindman BR., Clavel M-A., Mathieu P., et al. Calcific aortic stenosis. Nat Rev Dis Primer 2016;2:16006. Doi: 10.1038/nrdp.2016.6.

2. Eveborn GW., Schirmer H., Heggelund G., Lunde P., Rasmussen K. The evolving epidemiology of valvular aortic stenosis. The Tromsø Study. Heart 2013;99(6):396–400. Doi: 10.1136/heartjnl-2012-302265.

3. Osnabrugge RLJ., Mylotte D., Head SJ., et al. Aortic Stenosis in the Elderly. J Am Coll Cardiol 2013;62(11):1002–12. Doi: 10.1016/j.jacc.2013.05.015.

4. d'Arcy JL., Coffey S., Loudon MA., et al. Large-scale community echocardiographic screening reveals a major burden of undiagnosed valvular heart disease in older people: the OxVALVE Population Cohort Study. Eur Heart J 2016;37(47):3515–22. Doi: 10.1093/eurheartj/ehw229.

5. Otto CM., Prendergast B. Aortic-Valve Stenosis — From Patients at Risk to Severe Valve Obstruction. N Engl J Med 2014;371(8):744–56. Doi: 10.1056/NEJMra1313875.

6. Carabello BA. Introduction to Aortic Stenosis. Circ Res 2013;113(2):179–85. Doi: 10.1161/CIRCRESAHA.113.300156.

7. Watkins DA., Johnson CO., Colquhoun SM., et al. Global, Regional, and National Burden of Rheumatic Heart Disease, 1990–2015. N Engl J Med 2017;377(8):713–22. Doi: 10.1056/NEJMoa1603693.

8. Braunwald E. Aortic Stenosis: Then and Now. Circulation 2018;137(20):2099–100. Doi: 10.1161/CIRCULATIONAHA.118.033408.

9. Stewart BF., Siscovick D., Lind BK., et al. Clinical Factors Associated With Calcific Aortic Valve Disease fn1fn1This study was supported in part by Contracts NO1-HC85079 through HC-850086 from the National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland. J Am Coll Cardiol 1997;29(3):630–4. Doi: 10.1016/S0735-1097(96)00563-3.

10. Katz R., Wong ND., Kronmal R., et al. Features of the Metabolic Syndrome and Diabetes Mellitus as Predictors of Aortic Valve Calcification in the Multi-Ethnic Study of Atherosclerosis. Circulation 2006;113(17):2113–9. Doi: 10.1161/CIRCULATIONAHA.105.598086.

11. Roberts WC., Ko JM. Frequency by decades of unicuspid, bicuspid, and tricuspid aortic valves in adults having isolated aortic valve replacement for aortic stenosis, with or without associated aortic regurgitation. Circulation 2005;111(7):920–5. Doi: 10.1161/01.CIR.0000155623.48408.C5.

12. Ross J., Braunwald E. Aortic Stenosis. Circulation 1968;38(1s5). Doi: 10.1161/01.CIR.38.1S5.V-61.

13. Varadarajan P., Kapoor N., Bansal RC., Pai RG. Clinical profile and natural history of 453 nonsurgically managed patients with severe aortic stenosis. Ann Thorac Surg 2006;82(6):2111–5. Doi: 10.1016/j.athoracsur.2006.07.048.

14. Park S-J., Enriquez-Sarano M., Chang S-A., et al. Hemodynamic patterns for symptomatic

presentations of severe aortic stenosis. JACC Cardiovasc Imaging 2013;6(2):137–46. Doi: 10.1016/j.jcmg.2012.10.013.

15. Rajappan K., Rimoldi OE., Dutka DP., et al. Mechanisms of Coronary Microcirculatory Dysfunction in Patients With Aortic Stenosis and Angiographically Normal Coronary Arteries. Circulation 2002;105(4):470–6. Doi: 10.1161/hc0402.102931.

16. Gould KL., Carabello BA. Why Angina in Aortic Stenosis With Normal Coronary Arteriograms? Circulation 2003;107(25):3121–3. Doi: 10.1161/01.CIR.0000074243.02378.80.

17. Carabello BA. Georg Ohm and the Changing Character of Aortic Stenosis: It's Not Your Grandfather's Oldsmobile. Circulation 2012;125(19):2295–7. Doi: 10.1161/CIRCULATIONAHA.112.105825.

18. Vahanian A., Beyersdorf F., Praz F., et al. 2021 ESC/EACTS Guidelines for the management of valvular heart disease. Eur Heart J 2021:ehab395. Doi: 10.1093/eurheartj/ehab395.

19. Writing Committee Members, Otto CM., Nishimura RA., et al. 2020 ACC/AHA Guideline for the Management of Patients With Valvular Heart Disease: Executive Summary: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. J Am Coll Cardiol 2021;77(4):450–500. Doi: 10.1016/j.jacc.2020.11.035.

20. Ternacle J., Krapf L., Mohty D., et al. Aortic Stenosis and Cardiac Amyloidosis. J Am Coll Cardiol 2019;74(21):2638–51. Doi: 10.1016/j.jacc.2019.09.056.

21. Guzzetti E., Annabi M-S., Pibarot P., Clavel M-A. Multimodality Imaging for Discordant Low-Gradient Aortic Stenosis: Assessing the Valve and the Myocardium. Front Cardiovasc Med 2020;7:570689. Doi: 10.3389/fcvm.2020.570689.

22. Clavel M-A., Magne J., Pibarot P. Low-gradient aortic stenosis. Eur Heart J 2016;37(34):2645–57. Doi: 10.1093/eurheartj/ehw096.

23. Galiè N., Humbert M., Vachiery J-L., et al. 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: The Joint Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS)Endorsed by: Association for European Paediatric and Congenital Cardiology (AEPC), International Society for Heart and Lung Transplantation (ISHLT). Eur Heart J 2016;37(1):67–119. Doi: 10.1093/eurheartj/ehv317.

24. Tang M., Liu X., Lin C., et al. Meta-Analysis of Outcomes and Evolution of Pulmonary Hypertension Before and After Transcatheter Aortic Valve Implantation. Am J Cardiol 2017;119(1):91–9. Doi: 10.1016/j.amjcard.2016.09.015.

25. Lurz P., Orban M., Besler C., et al. Clinical characteristics, diagnosis, and risk stratification of pulmonary hypertension in severe tricuspid regurgitation and implications for transcatheter tricuspid valve repair. Eur Heart J 2020;41(29):2785–95. Doi: 10.1093/eurheartj/ehaa138.

26. Fortmeier V., Lachmann M., Körber MI., et al. Solving the Pulmonary Hypertension Paradox in Patients With Severe Tricuspid Regurgitation by Employing Artificial Intelligence. JACC Cardiovasc Interv 2022;15(4):381–94. Doi: 10.1016/j.jcin.2021.12.043.

27.   Généreux P., Pibarot P., Redfors B., et al. Staging classification of aortic stenosis based on the
       extent of cardiac damage. Eur Heart J 2017;38(45):3351–8. Doi: 10.1093/eurheartj/ehx381.

28.   Tomii D., Okuno T., Praz F., et al. Potential Candidates for Transcatheter Tricuspid Valve
       Intervention After Transcatheter Aortic Valve Replacement. JACC Cardiovasc Interv
       2021:S1936879821014023. Doi: 10.1016/j.jcin.2021.07.030.

29.   Rossebø AB., Pedersen TR., Boman K., et al. Intensive Lipid Lowering with Simvastatin and
       Ezetimibe in Aortic Stenosis. N Engl J Med 2008;359(13):1343–56. Doi:
       10.1056/NEJMoa0804602.

30.   Makkar RR., Fontana GP., Jilaihawi H., et al. Transcatheter Aortic-Valve Replacement for
       Inoperable Severe Aortic Stenosis. N Engl J Med 2012;366(18):1696–704. Doi:
       10.1056/NEJMoa1202277.

31.   Banovic M., Putnik S., Penicka M., et al. Aortic Valve ReplAcemenT versus Conservative
       Treatment in Asymptomatic SeveRe Aortic Stenosis: The AVATAR Trial. Circulation 2021. Doi:
       10.1161/CIRCULATIONAHA.121.057639.

32.   Spitzer E., Van Mieghem NM., Pibarot P., et al. Rationale and design of the Transcatheter Aortic
       Valve Replacement to UNload the Left ventricle in patients with ADvanced heart failure (TAVR
       UNLOAD) trial. Am Heart J 2016;182:80–8. Doi: 10.1016/j.ahj.2016.08.009.

33.   Iung B., Cachier A., Baron G., et al. Decision-making in elderly patients with severe aortic
       stenosis: why are so many denied surgery? Eur Heart J 2005;26(24):2714–20. Doi:
       10.1093/eurheartj/ehi471.

34.   Cribier A., Eltchaninoff H., Bash A., et al. Percutaneous Transcatheter Implantation of an Aortic
       Valve Prosthesis for Calcific Aortic Stenosis: First Human Case Description. Circulation
       2002;106(24):3006–8. Doi: 10.1161/01.CIR.0000047200.36165.B8.

35.   Leon MB., Smith CR., Mack M., et al. Transcatheter Aortic-Valve Implantation for Aortic Stenosis
       in Patients Who Cannot Undergo Surgery. N Engl J Med 2010;363(17):1597–607. Doi:
       10.1056/NEJMoa1008232.

36.   Zahn R., Gerckens U., Grube E., et al. Transcatheter aortic valve implantation: first results from
       a multi-centre real-world registry. Eur Heart J 2011;32(2):198–204. Doi:
       10.1093/eurheartj/ehq339.

37.   Andersen HR. How Transcatheter Aortic Valve Implantation (TAVI) Was Born: The Struggle for a
       New Invention. Front Cardiovasc Med 2021;8:722693. Doi: 10.3389/fcvm.2021.722693.

38.   Kodali SK., Williams MR., Smith CR., et al. Two-Year Outcomes after Transcatheter or Surgical
       Aortic-Valve Replacement. N Engl J Med 2012;366(18):1686–95. Doi: 10.1056/NEJMoa1200384.

39.   Leon MB., Smith CR., Mack MJ., et al. Transcatheter or Surgical Aortic-Valve Replacement in
       Intermediate-Risk Patients. N Engl J Med 2016;374(17):1609–20. Doi: 10.1056/NEJMoa1514616.

40.   Mack MJ., Leon MB., Thourani VH., et al. Transcatheter Aortic-Valve Replacement with a

Balloon-Expandable Valve in Low-Risk Patients. N Engl J Med 2019;380(18):1695–705. Doi: 10.1056/NEJMoa1814052.

41. Tam DY., Hughes A., Wijeysundera HC., Fremes SE. Cost-Effectiveness of Self-Expandable Transcatheter Aortic Valves in Intermediate-Risk Patients. Ann Thorac Surg 2018;106(3):676–83. Doi: 10.1016/j.athoracsur.2018.03.069.

42. Baron SJ., Wang K., House JA., et al. Cost-Effectiveness of Transcatheter Versus Surgical Aortic Valve Replacement in Patients With Severe Aortic Stenosis at Intermediate Risk. Circulation 2019;139(7):877–88. Doi: 10.1161/CIRCULATIONAHA.118.035236.

43. Tang GHL., Verma S., Bhatt DL. Transcatheter Aortic Valve Replacement in Low-Risk Patients: A New Era in the Treatment of Aortic Stenosis. Circulation 2019;140(10):801–3. Doi: 10.1161/CIRCULATIONAHA.119.041111.

44. Beohar N., Whisenant B., Kirtane AJ., et al. The relative performance characteristics of the logistic European System for Cardiac Operative Risk Evaluation score and the Society of Thoracic Surgeons score in the Placement of Aortic Transcatheter Valves trial. J Thorac Cardiovasc Surg 2014;148(6):2830-2837.e1. Doi: 10.1016/j.jtcvs.2014.04.006.

45. Amanullah MR., Pio SM., Ng ACT., et al. Prognostic Implications of Associated Cardiac Abnormalities Detected on Echocardiography in Patients With Moderate Aortic Stenosis. JACC Cardiovasc Imaging 2021;14(9):1724–37. Doi: 10.1016/j.jcmg.2021.04.009.

46. Schewel J., Schmidt T., Kuck K-H., Frerker C., Schewel D. Impact of Pulmonary Hypertension Hemodynamic Status on Long-Term Outcome After Transcatheter Aortic Valve Replacement. JACC Cardiovasc Interv 2019;12(21):2155–68. Doi: 10.1016/j.jcin.2019.08.031.

47. Alushi B., Beckhoff F., Leistner D., et al. Pulmonary Hypertension in Patients With Severe Aortic Stenosis: Prognostic Impact After Transcatheter Aortic Valve Replacement. JACC Cardiovasc Imaging 2019;12(4):591–601. Doi: 10.1016/j.jcmg.2018.02.015.

48. Masri A., Abdelkarim I., Sharbaugh MS., et al. Outcomes of persistent pulmonary hypertension following transcatheter aortic valve replacement. Heart 2018;104(10):821–7. Doi: 10.1136/heartjnl-2017-311978.

49. Cremer PC., Zhang Y., Alu M., et al. The incidence and prognostic implications of worsening right ventricular function after surgical or transcatheter aortic valve replacement: insights from PARTNER IIA. Eur Heart J 2018;39(28):2659–67. Doi: 10.1093/eurheartj/ehy251.

50. Poch F., Thalmann R., Olbrich I., et al. Changes of Right Ventricular Function After Transcatheter Aortic Valve Replacement and Association With Outcomes. J Card Fail 2021:S1071916421001184. Doi: 10.1016/j.cardfail.2021.03.007.

51. Raghunath S., Ulloa Cerna AE., Jing L., et al. Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. Nat Med 2020. Doi: 10.1038/s41591-020-0870-z.

52. Fries JA., Varma P., Chen VS., et al. Weakly supervised classification of aortic valve malformations using unlabeled cardiac MRI sequences. Nat Commun 2019;10(1):3111. Doi:

10.1038/s41467-019-11012-3.

53.   Diller G-P., Kempny A., Babu-Narayan SV., et al. Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease: data from a single tertiary centre including 10 019 patients. Eur Heart J 2019;40(13):1069–77. Doi: 10.1093/eurheartj/ehy915.

54.   Perez MV., Mahaffey KW., Hedlin H., et al. Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation. N Engl J Med 2019;381(20):1909–17. Doi: 10.1056/NEJMoa1901183.

55.   Kwak S., Lee Y., Ko T., et al. Unsupervised Cluster Analysis of Patients With Aortic Stenosis Reveals Distinct Population With Different Phenotypes and Outcomes. Circ Cardiovasc Imaging 2020;13(5). Doi: 10.1161/CIRCIMAGING.119.009707.

56.   Sengupta PP., Shrestha S., Kagiyama N., et al. A Machine-Learning Framework to Identify Distinct Phenotypes of Aortic Stenosis Severity. JACC Cardiovasc Imaging 2021:S1936-878X(21)00286-2. Doi: 10.1016/j.jcmg.2021.03.020.

57.   Chen P-HC., Gadepalli K., MacDonald R., et al. An augmented reality microscope with real-time artificial intelligence integration for cancer diagnosis. Nat Med 2019;25(9):1453–7. Doi: 10.1038/s41591-019-0539-7.

58.   Kossmann CE. Electrocardiographic Analysis by Computer. JAMA J Am Med Assoc 1965;191(11):922. Doi: 10.1001/jama.1965.03080110046011.

59.   LeCun Y., Bengio Y., Hinton G. Deep learning. Nature 2015;521(7553):436–44. Doi: 10.1038/nature14539.

60.   Dey D., Slomka PJ., Leeson P., et al. Artificial Intelligence in Cardiovascular Imaging. J Am Coll Cardiol 2019;73(11):1317–35. Doi: 10.1016/j.jacc.2018.12.054.

61.   Casaclang-Verzosa G., Shrestha S., Khalil MJ., et al. Network Tomography for Understanding Phenotypic Presentations in Aortic Stenosis. JACC Cardiovasc Imaging 2019;12(2):236–48. Doi: 10.1016/j.jcmg.2018.11.025.

62.   Wolpert DH. The Lack of A Priori Distinctions Between Learning Algorithms. Neural Comput 1996;8(7):1341–90. Doi: 10.1162/neco.1996.8.7.1341.

63.   Quer G., Arnaout R., Henne M., Arnaout R. Machine Learning and the Future of Cardiovascular Care. J Am Coll Cardiol 2021;77(3):300–13. Doi: 10.1016/j.jacc.2020.11.030.

64.   Ribeiro JM., Astudillo P., de Backer O., et al. Artificial Intelligence and Transcatheter Interventions for Structural Heart Disease: A glance at the (near) future. Trends Cardiovasc Med 2021:S1050173821000177. Doi: 10.1016/j.tcm.2021.02.002.

65.   Gurovich Y., Hanani Y., Bar O., et al. Identifying facial phenotypes of genetic disorders using deep learning. Nat Med 2019;25(1):60–4. Doi: 10.1038/s41591-018-0279-0.

66.   Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25(1):44–56. Doi: 10.1038/s41591-018-0300-7.

67.   Goodman B., Flaxman S. European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation." AI Mag 2017;38(3):50–7. Doi: 10.1609/aimag.v38i3.2741.

68. Lundberg SM., Erion G., Chen H., et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2020;2(1):56–67. Doi: 10.1038/s42256-019-0138-9.

69. Lauritsen SM., Kristensen M., Olsen MV., et al. Explainable artificial intelligence model to predict acute critical illness from electronic health records. Nat Commun 2020;11(1):3852. Doi: 10.1038/s41467-020-17431-x.

70. Lachmann M., Rippen E., Schuster T., et al. Artificial intelligence-enabled phenotyping of patients with severe aortic stenosis: on the recovery of extra-aortic valve cardiac damage after transcatheter aortic valve replacement. Open Heart 2022;9(2):e002068. Doi: 10.1136/openhrt-2022-002068.

71. Asami M., Stortecky S., Praz F., et al. Prognostic Value of Right Ventricular Dysfunction on Clinical Outcomes After Transcatheter Aortic Valve Replacement. JACC Cardiovasc Imaging 2019;12(4):577–87. Doi: 10.1016/j.jcmg.2017.12.015.

72. Cremer PC., Wang TKM., Rodriguez LL., et al. Incidence and Clinical Significance of Worsening Tricuspid Regurgitation Following Surgical or Transcatheter Aortic Valve Replacement: Analysis From the PARTNER IIA Trial. Circ Cardiovasc Interv 2021;14(8). Doi: 10.1161/CIRCINTERVENTIONS.120.010437.

73. Trenkwalder T., Lachmann M., Stolz L., et al. Machine learning identifies pathophysiologically and prognostically informative phenotypes among patients with mitral regurgitation undergoing transcatheter edge-to-edge repair. Eur Heart J - Cardiovasc Imaging 2023:jead013. Doi: 10.1093/ehjci/jead013.

74. Rajpurkar P., Chen E., Banerjee O., Topol EJ. AI in health and medicine. Nat Med 2022. Doi: 10.1038/s41591-021-01614-0.

75. Lachmann M., Rippen E., Rueckert D., et al. Harnessing feature extraction capacities from a pre-trained convolutional neural network (VGG-16) for the unsupervised distinction of aortic outflow velocity profiles in patients with severe aortic stenosis. Eur Heart J - Digit Health 2022:ztac004. Doi: 10.1093/ehjdh/ztac004.

76. Lachmann M., Rippen E., Schuster T., et al. Subphenotyping of Patients With Aortic Stenosis by Unsupervised Agglomerative Clustering of Echocardiographic and Hemodynamic Data. JACC Cardiovasc Interv 2021;14(19):2127–40. Doi: 10.1016/j.jcin.2021.08.034.

8 Danksagung

Mein größter Dank gebührt meinem Doktorvater Herrn Prof. Karl-Ludwig Laugwitz. Unter seiner schützenden Hand durfte sich unsere junge Arbeitsgruppe zur künstlichen Intelligenz entfalten. Von ihm habe ich gelernt, dass der Fortschritt über die Zeit wichtiger ist als die Perfektion zu Beginn. Für die Zukunft würde ich mir wünschen, dass ich die Faszination für die Synergie aus Klinik und Wissenschaft, wie sie Herr Prof. Laugwitz mir vorgelebt hat, weiterhin verspüre.

Ebenso möchte ich unseren Kooperationspartnern aus dem Deutschen Herzzentrum München danken, insbesondere Herrn Prof. Michael Joner, Herrn PD Erion Xhepa und Frau PD Teresa Trenkwalder. Dank ihrer Hilfe konnten die im Klinikum rechts der Isar entwickelten Ideen an größeren Patientenkollektiven angewendet und schließlich zur Publikation gebracht werden. Herr PD Erion Xhepa fungierte darüberhinaus als Mentor meiner Doktorarbeit und inspirierte uns, unsere Algorithmen für die Patienten mit Aortenklappenstenosen auf seine Patienten mit Mitralklappeninsuffizienzen umzuschreiben.

Mit meinem zweiten Mentor Herrn Dr. Mark Lachmann verband mich eine familiäre Freundschaft, weshalb ich Loyalität als seine wichtigste Qualität beschreiben würde. Seit dem Entstehen unserer Arbeitsgruppe 2019 haben wir im Enthusiasmus für unsere Forschung gefühlt 1000 Projektideen entwickelt, 99% verworfen, und 1% realisiert. Dieses 1% stellt absolut betrachtet meine Promotion und seine Habilitation dar. Es war mir immer eine Freude, ihn zu Kongressen auf der ganzen Welt von San Francisco über Barcelona und Venedig bis nach Tokio begleiten zu dürfen, um unsere aktuellsten Ergebnisse zu präsentieren und neue Patienten zu rekrutieren.

Meiner Familie, meinem Freund und seiner Familie, sowie meinen Freunden möchte ich von ganzem Herzen für die langjährige unermüdliche Unterstützung, sowie Ermutigungen und Motivation während der Schulzeit, des Studiums und natürlich der Dissertation danken. Ich hoffe, sie sind bereit für die nächsten Schritte.

Abschließend möchte ich dem blinden Zufall in Person von Frau Sabine Rössler und Frau Vera Lachmann danken, denn sie haben mich in die Kardiologie an das Klinikum rechts der Isar gelotst und mir meinen Mentor Herrn Dr. Lachmann an die Seite gestellt. Außerdem möchte ich mich bei Frau Dr. Vera Fortmeier (geb. Lachmann) bedanken, welche mich für die Anwendung der künstlichen Intelligenz bei Trikuspidalklappeninsuffizienzen begeisterte und mir eventuell als Vorbild dienen wird, wie man den anspruchsvollen Spagat zwischen Klinik, Forschung und Familie meistern könnte.

9 Appendix

ORIGINAL ARTICLE

Harnessing feature extraction capacities from a pre-trained convolutional neural network (VGG-16) for the unsupervised distinction of aortic outflow velocity profiles in patients with severe aortic stenosis

Mark Lachmann [1†], Elena Rippen [1†], Daniel Rueckert [2,3], Tibor Schuster[4], Erion Xhepa [5,6], Moritz von Scheidt [5,6], Costanza Pellegrini[5], Teresa Trenkwalder [5], Tobias Rheude[5], Anja Stundl [1], Ruth Thalmann[1], Gerhard Harmsen[7], Shinsuke Yuasa[8], Heribert Schunkert [5,6], Adnan Kastrati [5,6], Michael Joner[5,6], Christian Kupatt [1,6], and Karl-Ludwig Laugwitz [1,6]*

[1]First Department of Medicine, Klinikum rechts der Isar, Technical University of Munich, Ismaninger Straße 22, 81675 Munich, Germany; [2]Institute for AI and Informatics in Medicine, Faculty of Informatics and Medicine, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany; [3]Department of Computing, Imperial College London, London, UK ; [4]Department of Family Medicine, McGill University, Montreal, Quebec, Canada; [5]Department of Cardiology, German Heart Centre Munich, Technical University of Munich, Munich, Germany; [6]DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Munich, Germany; [7]Department of Physics, University of Johannesburg, Auckland Park, South Africa; and [8]Department of Cardiology, Keio University School of Medicine, Minato, Tokyo, Japan

Aims	Hypothesizing that aortic outflow velocity profiles contain more valuable information about aortic valve obstruction and left ventricular contractility than can be captured by the human eye, features of the complex geometry of Doppler tracings from patients with severe aortic stenosis (AS) were extracted by a convolutional neural network (CNN).
Methods and results	After pre-training a CNN (VGG-16) on a large data set (ImageNet data set; 14 million images belonging to 1000 classes), the convolutional part was employed to transform Doppler tracings to 1D arrays. Among 366 eligible patients [age: 79.8 ± 6.77 years; 146 (39.9%) women] with pre-procedural echocardiography and right heart catheterization prior to transcatheter aortic valve replacement (TAVR), good quality Doppler tracings from 101 patients were analysed. The convolutional part of the pre-trained VGG-16 model in conjunction with principal component analysis and k-means clustering distinguished two shapes of aortic outflow velocity profiles. Kaplan–Meier analysis revealed that mortality in patients from Cluster 2 ($n = 40$, 39.6%) was significantly increased [hazard ratio (HR) for 2-year mortality: 3; 95% confidence interval (CI): 1–8.9]. Apart from reduced cardiac output and mean aortic valve gradient, patients from Cluster 2 were also characterized by signs of pulmonary hypertension, impaired right ventricular function, and right atrial enlargement. After training an extreme gradient boosting algorithm on these 101 patients, validation on the remaining 265 patients confirmed that patients assigned to Cluster 2 show increased mortality (HR for 2-year mortality: 2.6; 95% CI: 1.4–5.1, P-value: 0.004).

* Corresponding author. Tel: +49 89 4140 2350, Fax: +49 89 4140 4900, Email: KL.Laugwitz@mri.tum.de
†The first two authors contributed equally to the study.

Conclusion	Transfer learning enables sophisticated pattern recognition even in clinical data sets of limited size. Importantly, it is the left ventricular compensation capacity in the face of increased afterload, and not so much the actual obstruc- tion of the aortic valve, that determines fate after TAVR.

Graphical Abstract
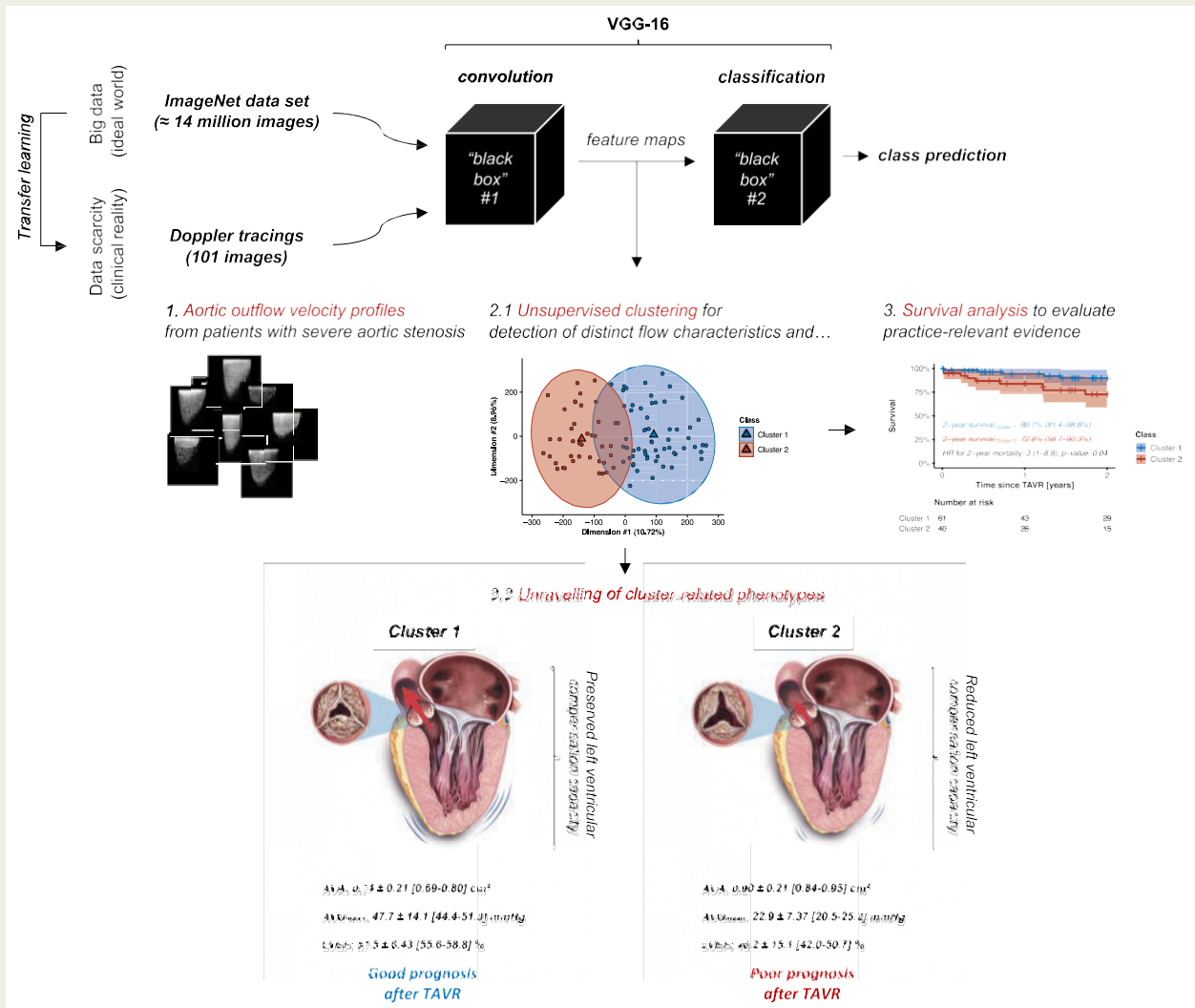
Keywords	Severe aortic stenosis • Transcatheter aortic valve replacement • Aortic outflow velocity profile • Convolutional neural network • Transfer learning

Introduction

In the dawning age of artificial intelligence, machine learning technology is progressively implemented into medical research and clinical decision support.[1-6] Without the constraint of any *a priori* assumption, machine learning algorithms iteratively learn from data, typically requiring a myriad of information. Bedside research embedded in clinical practice, however, is commonly based on modest-sized patient cohorts, e.g. in the setting of rare diseases or novel treatment strategies. Transfer learning holds the promise to

alleviate the bottleneck of insufficient training data by acquiring feature extraction capacity from a large data set and applying it to a related problem in the same domain.

Severe aortic stenosis (AS), which can trigger a deleterious cascade including left heart dysfunction, pulmonary hypertension (PH), and eventually right heart failure,[7] and which is associated with a 2-year mortality of up to more than 50% unless valve replacement is performed promptly,[8] is typically diagnosed by transthoracic echocardiography.[9] Besides an increasing gradient due to progressive narrowing of the aortic valve, the aortic outflow velocity profile in patients with severe AS changes from a triangular shape with an early peak to a much more rounded form with a later peak.[10] Furthermore, acceleration of flow velocity will eventually deteriorate due to left ventricular decompensation, resulting in a 'low flow, low gradient AS'.[11]

Hypothesizing that the aortic outflow velocity profile contains more valuable information about aortic valve obstruction and left ventricular contractility than can be captured by human cognition, this study sought to extract features of the complex geometry of Doppler tracings by employing a convolutional neural network (CNN). VGG-16 is a CNN with state-of-the-art feature extraction capacity, which achieved 92.7% Top 5 test accuracy in a data set of over 14 million regular natural images belonging to 1000 classes (ImageNet data set).[12] Adopting the concept of transfer learning, the convolutional part of the pre-trained VGG-16 model was employed to transform Doppler tracings from a small, but well-characterized cohort of patients with severe AS to 1D arrays. After principal component analysis (PCA) and k-means clustering of those 1D arrays, practice-relevant evidence was assessed by relating cluster assignment with all-cause 2-year mortality after transcatheter aortic valve replacement (TAVR).

Methods

Patient recruitment

This is a *post hoc* analysis of prospectively and systematically collected data from patients undergoing TAVR for severe AS at two tertiary care centres in Munich, Germany, between January 2014 and December 2020. The study was approved by the respective local ethics committees in conformity with the Declaration of Helsinki, and all patients enrolled provided written informed consent. In total, the joint registry listed 2575 patients. Among 366 completely characterized patients with pre-procedural echocardiography and right heart catheterization prior to TAVR, good quality Doppler tracings with sharp, well-defined borders from the intentionally small number of 101 patients were analysed, constituting the derivation cohort (*Figure 1A*). The validation cohort was consequently represented by the remaining 265 patients with complete data from pre-procedural echocardiography and right heart catheterization, yet without good quality Doppler tracings (or no available records at all). As an elderly patient population approaching the end of life was studied, post-procedural 2-year all-cause mortality was defined as a clinically meaningful primary outcome measure. Survival data were obtained from the German Civil Registry in case of patients being registered in Germany ($n = 354$; 96.7%), or from general practitioners, hospitals, and practice cardiologists for patients from foreign countries.

Transthoracic echocardiography

All echocardiographic studies were performed by experienced institutional cardiologists during clinical routine using a commercially available echocardiography system equipped with a 2.5-MHz multifrequency phased-array transducer. The continuous wave Doppler-derived aortic outflow velocity profiles were obtained from the apical four-chamber view (*Figure 2A*). Only 101 aortic outflow velocity profiles were selected for clustering depending on image quality—meaning that Doppler tracings with insufficient contrast or with labelling within the aortic outflow velocity profile were excluded (*Figure 2B*).

Statistical analysis

All statistical analyses were performed using R version 3.6.3 (R Foundation for Statistical Computing, Vienna, Austria; see Supplementary material online, *Table S1* for a complete list of employed R packages). The pre-trained VGG-16 model on the ImageNet data set was loaded from the Keras deep learning library (R package 'keras'), and the classifier part of the VGG-16 model was omitted. After pre-processing (*Figure 2A*), scaled Doppler tracings as input images were thus converted to a feature tensor of shape as the output of the last layer of the convolutional part (R packages 'magick' and 'imager'). After transformation from feature tensor of shape (7, 7, 512) to 1D array with $7 \times 7 \times 512$ values per instance, PCA and k-means clustering were applied (R packages 'FactoMineR', 'factoextra', and 'NbClust'). Notably, we expected two clusters to be segregated. Survival was illustrated using Kaplan–Meier method, and a Cox proportional hazards model was used to estimate hazard ratios (HRs) between identified clusters (R packages 'survival', 'survminer', and 'ggforest'). Missing values among variables that were identified as significant predictors for 2-year mortality in initial univariate analysis were imputed by a random forest algorithm (R package 'missForest')[13] before proceeding with multivariate analysis, but were not used hereinafter, e.g. for cluster comparisons. Because the derivation cohort was unbalanced with regards to cluster assignments, a technique to synthetically over-sample the minority class was applied (synthetic minority over-sampling technique; SMOTE) (R package 'DMwR').[14] After balancing, an extreme gradient boosting algorithm (R package 'xgboost') was selected as the machine learning technique of choice for cluster assignment in future patients, and it was trained on a comprehensive set of functional and structural parameters from echocardiography and right heart catheterization. Again, missing values were imputed by a random forest algorithm. SHAP (SHapley Additive exPlanations) values were calculated to compare the contribution of input variables to the model prediction (R package 'SHAPforxgboost').[15] To sum up, this study was designed as a two-step experiment:

(1) In a first step, we aimed to decipher meaningful echocardiographic signatures and related cardiac phenotypes by analysing aortic outflow velocity profiles (Doppler tracings) from 101 patients with severe AS undergoing TAVR by using a pre-trained CNN in conjunction with PCA and k-means clustering (unsupervised machine learning experiment).

(2) Since the first experiment did not allow to assign future patients to the just defined clusters, we additionally sought to employ an extreme gradient boosting algorithm, which was trained on functional and structural parameters of cardiopulmonary conditions from the 101 patients with good quality Doppler tracings (hereinafter referred to as derivation cohort) to predict cluster assignments as stemming from the first experiment, and which was then validated on the remaining 265 patients (hereinafter referred to as validation
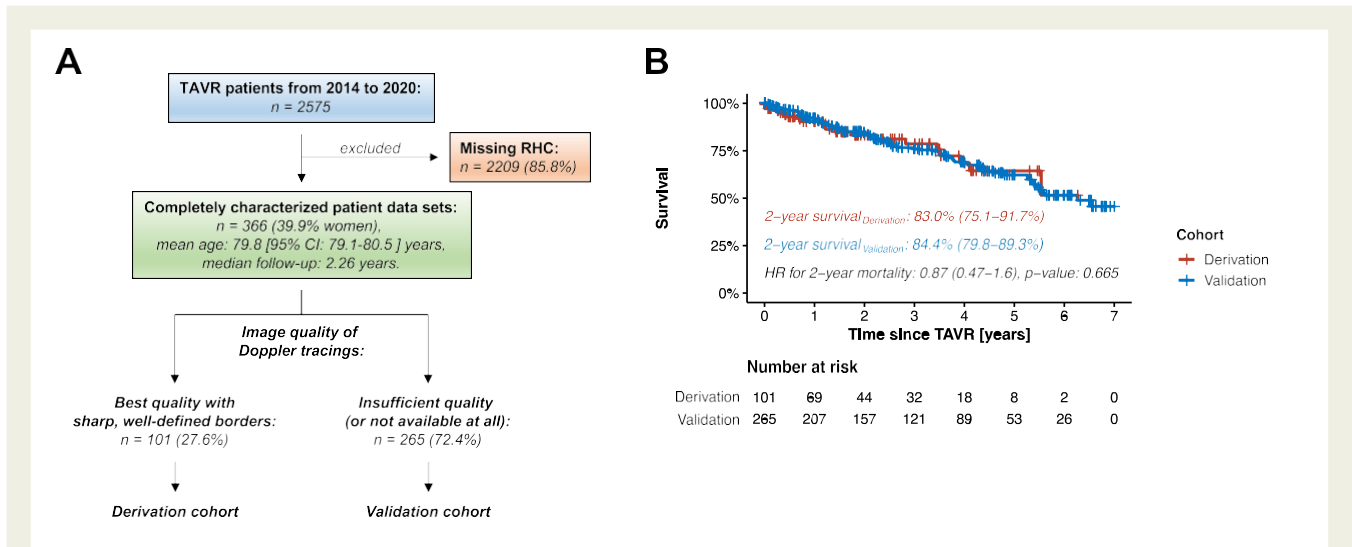
Figure 1 General information about the study population from recruitment to follow-up. (*A*) A flowchart for patient recruitment in order to select 101 patients with best quality Doppler tracings. Notably, 56 out of 366 patients (15.3%) had no Doppler tracings as raw data available. (*B*) Kaplan-Meier survival plot testing for differences in survival between derivation and validation cohorts. RHC, right heart catheterization; TAVR, transcatheter aortic valve replacement.
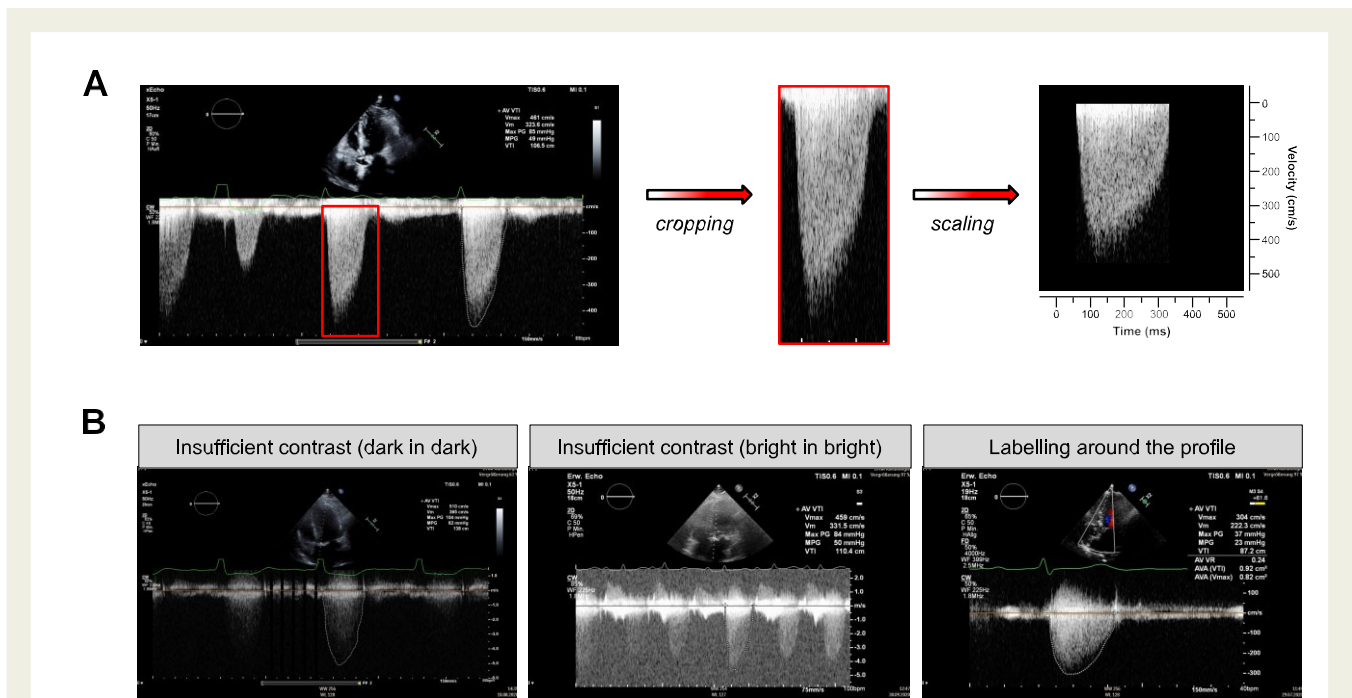
Figure 2 Pre-processing of aortic outflow velocity profiles from patients with severe AS data input for the convolutional neural network. (*A*) Schematic of the image pre-processing pipeline. One representative aortic outflow velocity profile per patient was extracted from records. Cropping of the region of interest, i.e. during systole, was done manually. Since original echocardiographic images were recorded at different scales, but homogeneity of data input had to be provided, Doppler tracings were further manually scaled according to uniform time and velocity axes (see also Supplementary material online, *Figure S1* for a standard operation procedure explaining additional details to create the desired normalized profiles). Neither image normalization nor histogram equalization was applied during pre-processing. Re-sizing to 224 × 224 pixel format as the default input size of the VGG-16 model was already part of the image processing R code after loading the folder with 101 scaled Doppler tracings. (*B*) Representative Doppler tracings that were excluded due to insufficient contrast, or due to labelling within the aortic outflow velocity profile.

cohort) with regards to cluster-related survival after TAVR (supervised machine learning experiment).

Categorical variables are presented as numbers and frequencies (%), whilst continuous variables are given as mean ± standard deviation (SD) and 95% confidence interval (CI). Chi-square or Fisher's exact test was used to evaluate the association between categorical variables, and independent-samples Wilcoxon test was used for comparison of continuous variables. For analysis of collinearity, Pearson's correlation coefficients were calculated. A P-value ≤ 0.05 was considered to indicate statistical significance.

Results

One hundred and one patients with good quality Doppler tracings illustrate the problem of data scarcity in clinical research

Importantly, aortic outflow velocity profiles from only 101 out of 366 patients were initially analysed (*Figure 1A*) to emphasize the problem of data scarcity in clinical research. Therefore, to confirm the representative nature of the small-sized derivation cohort at hand [mean age: 79.3 ± 6.78; 95% CI: 78.0–80.7 years; 49 (48.5%) women], derivation and validation cohorts were initially compared with regards to demographic, clinical, echocardiographic, and haemodynamic characteristics (Supplementary material online, *Tables S2 and S3*). No significant differences were found with regards to age, symptomatic burden expressed as New York Heart Association (NYHA) functional class, obstruction of the aortic valve expressed as aortic valve area (AVA), left ventricular systolic function, mean pulmonary artery pressure (mPAP), and right ventricular dysfunction. In fact, a difference was detected in the proportion of female patients, which were significantly more often represented in the derivation cohort (48.5% vs.

36.6%, *P*-value: 0.0499). Presenting with a mean AVA of 0.804 ± 0.223 (95% CI: 0.760–0.848) cm^2, and predominantly suffering from dyspnoea corresponding to NYHA functional Class III (56.4%) (*Tables 1* and *2*) 2-year survival after TAVR among patients from the derivation cohort ranged at 83.0% (95% CI: 75.1–91.7), which was statistically indifferent compared to patients from the validation cohort (*P*-value: 0.665) (*Figure 1B*).

Two distinct clusters of aortic outflow velocity profiles can be distinguished, reflecting different phenotypes with subsequently differing mortality

The convolutional part of the pre-trained VGG-16 model (*Figure 3A*) in conjunction with PCA and k-means clustering of the abstractions of Doppler tracings enabled to distinguish two shapes of aortic outflow velocity profiles (*Figure 3B*). Interestingly, all patients from Cluster 2 presented with a mean aortic valve gradient (AVG$_{mean}$) below 40 mmHg, whilst AVG$_{mean}$ from patients in Cluster 1 ranged between 20 and 102 mmHg (*Figure 3C*). Kaplan–Meier analysis revealed that mortality in patients from Cluster 2 ($n = 40$, 39.6%) was significantly increased (HR for 2-year mortality: 3; 95% CI: 1–8.9) (*Figure 3D*). Besides reduced cardiac output (4.57 ± 1.42; 95% CI: 4.17–5.04 L/min) and signs of PH (mPAP: 31.9 ± 12.2; 95% CI: 28.5–35.7 mmHg), patients from Cluster 2 also presented with more severe impairment of right ventricular function [tricuspid annular plane systolic excursion (TAPSE): 18.1 ± 3.82; 95% CI: 17.3–19.2 mm] and right atrial enlargement [right atrial (RA) area: 22.0 ± 8.28; 95% CI: 19.5–24.6 cm^2] in comparison to patients from Cluster 1 (*Figure 3E* and *Table 2*). Contrarily to the initial expectation, patients from Cluster 1 with seemingly less extensive cardiac damage were diagnosed with a more severe obstruction of the aortic valve than patients from Cluster 2 (AVA: 0.739 ± 0.211; 95% CI: 0.685–0.793 cm^2 vs. 0.903 ± 0.205; 95% CI: 0.837–0.968 cm^2, *P*-value: 0.0001).

Table 1 Demographic and clinical characteristics in accordance with cluster assignment (derivation cohort)

	All ($n = 101$)	Cluster 1 ($n = 61$)	Cluster 2 ($n = 40$)	*P*-value
		Class		
Age (years), mean ± SD [95% CI]	79.3 ± 6.78 [78.0–80.7]	79.4 ± 5.88 [77.8–80.8]	79.3 ± 8.03 [76.8–81.6]	0.4067
Women, *N* (%)	49 (48.5%)	31 (50.8%)	18 (45.0%)	0.7123
BMI (kg/m^2), mean ± SD [95% CI]	26.8 ± 4.28 [26.0–27.6]	26.9 ± 4.25 [25.8–28.0]	26.7 ± 4.36 [25.5–28.0]	0.8758
Arterial hypertension, *N* (%)	88 (87.1%)	51 (83.6%)	37 (92.5%)	0.3166
Diabetes mellitus, *N* (%)	23 (22.8%)	14 (23.0%)	9 (22.5%)	1
NYHA functional class, mean ± SD [95% CI]	2.61 ± 0.71 [2.47–2.75]	2.54 ± 0.72 [2.36–2.72]	2.72 ± 0.68 [2.53–2.93]	0.1896
NYHA functional Class III	57 (56.4%)	32 (52.5%)	25 (62.5%)	0.4294
NYHA functional Class IV	6 (5.9%)	3 (4.9%)	3 (7.5%)	0.9152
EuroSCORE (%), mean ± SD [95% CI]	17.1 ± 14.3 [14.3–20.0]	13.9 ± 8.94 [11.7–16.3]	22.1 ± 19.0 [16.5–27.6]	0.0694
eGFR (mL/min), mean ± SD [95% CI]	60.7 ± 21.4 [56.4–64.9]	65.4 ± 19.5 [60.1–70.2]	53.3 ± 22.3 [46.8–59.3]	0.0214
CAD, *N* (%)	85 (84.1%)	49 (80.3%)	36 (90.0%)	0.3061
COPD, *N* (%)	12 (11.9%)	7 (11.5%)	5 (12.5%)	1
Atrial fibrillation and/or flutter, *N* (%)	42 (41.6%)	19 (31.1%)	23 (57.5%)	0.0155

BMI, body mass index; CAD, coronary artery disease; CI, confidence interval; COPD, chronic obstructive pulmonary disease; eGFR, estimated glomerular filtration rate; NYHA, New York Heart Association; SD, standard deviation.

Table 2 Comparison of echocardiographic and haemodynamic characteristics in accordance with cluster assignment (derivation cohort)

	All (*n* 5 101)	Cluster 1 (*n* 5 61)	Cluster 2 (*n* 5 40)	*P*-value
		Class		
AVA (cm^2), mean ± SD [95% CI]	0.804 ± 0.223 [0.760–0.848]	0.739 ± 0.211 [0.685–0.793]	0.903 ± 0.205 [0.837–0.968]	0.0001
AVG$_{mean}$ (mmHg), mean ± SD [95% CI]	37.9 ± 17.0 [34.5–41.2]	47.7 ± 14.1 [44.4–51.3]	22.9 ± 7.37 [20.5–25.2]	4.6×10^{-15}
Cardiac output (L/min), mean ± SD [95% CI]	5.08 ± 1.33 [4.82–5.34]	5.41 ± 1.17 [5.11–5.70]	4.57 ± 1.42 [4.17–5.04]	0.0006
LVEF (%), mean ± SD [95% CI]	53.0 ± 12.0 [50.6–55.4]	57.5 ± 6.43 [55.6–58.8]	46.2 ± 15.1 [42.0–50.7]	0.0001
LVEDD (mm), mean ± SD [95% CI]	47.2 ± 9.44 [45.4–49.1]	45.7 ± 7.70 [43.8–47.7]	49.7 ± 11.4 [46.2–52.8]	0.1028
mPCWP (mmHg), mean ± SD [95% CI]	17.8 ± 9.01 [16.1–19.6]	15.5 ± 8.40 [13.3–17.6]	21.4 ± 8.83 [18.6–24.2]	0.0007
mPAP (mmHg), mean ± SD [95% CI]	27.6 ± 11.5 [25.3–29.8]	24.7 ± 10.1 [22.1–27.2]	31.9 ± 12.2 [28.5–35.7]	0.0019
RV-RA gradient (mmHg), mean ± SD [95% CI]	34.0 ± 14.7 [30.9–37.1]	32.2 ± 13.8 [28.5–35.9]	37.0 ± 16.0 [31.3–42.6]	0.1302
PVR (WU), mean ± SD [95% CI]	2.10 ± 1.36 [1.83–2.37]	1.80 ± 0.997 [1.55–2.06]	2.56 ± 1.69 [2.02–3.10]	0.0050
TAPSE (mm), mean ± SD [95% CI]	19.8 ± 4.05 [18.9–20.6]	20.8 ± 3.89 [19.8–21.7]	18.1 ± 3.82 [17.3–19.2]	0.0014
Right midventricular diameter (mm), mean ± SD [95% CI]	29.0 ± 6.46 [27.6–30.3]	27.4 ± 5.82 [25.8–29.0]	31.1 ± 6.77 [28.8–33.3]	0.0088
LA area (cm^2), mean ± SD [95% CI]	25.8 ± 8.21 [24.2–27.5]	24.8 ± 8.11 [22.9–27.3]	27.4 ± 8.21 [24.8–30.1]	0.1017
RA area (cm^2), mean ± SD [95% CI]	19.5 ± 6.89 [18.2–20.9]	17.8 ± 5.17 [16.6–18.9]	22.0 ± 8.28 [19.5–24.6]	0.0133
Low gradient (AVG$_{mean}$ < 40 mmHg), *N* (%)	54 (53.5%)	14 (23.0%)	40 (100%)	1.5×10^{-13}
LV dysfunction (LVEF <_ 45%), *N* (%)	22 (21.8%)	5 (8.2%)	17 (42.5%)	0.0001
PH (mPAP >_ 25 mmHg), *N* (%)	52 (51.5%)	25 (41.0%)	27 (67.5%)	0.0162
RV dysfunction (TAPSE <_ 16 mm), *N* (%)	20 (20.4%)	6 (9.8%)	14 (23.0%)	0.0021
MR >_ III/IV°, *N* (%)	10 (9.90%)	4 (6.56%)	6 (15.0%)	0.1882
TR >_ III/IV°, *N* (%)	7 (6.93%)	2 (3.28%)	5 (12.5%)	0.1101

AVA, aortic valve area; AVG$_{mean}$, mean aortic valve gradient; CI, confidence interval; LA area, left atrial area; LV dysfunction, left ventricular dysfunction; LVEDD, left ventricular end-diastolic diameter; LVEF, left ventricular ejection fraction; mPAP, mean pulmonary artery pressure; mPCWP, mean postcapillary wedge pressure; MR, mitral regurgitation; PH, pulmonary hypertension; PVR, pulmonary vascular resistance; RA area, right atrial area; RV dysfunction, right ventricular dysfunction; SD, standard deviation; TAPSE, tricuspid annular plane systolic excursion; TR, tricuspid regurgitation.

Comorbidities, such as arterial hypertension, coronary artery disease, and chronic obstructive pulmonary disease, were similarly prevalent between Clusters 1 and 2. Yet, patients from Cluster 2 with failing hearts showed a higher prevalence of atrial fibrillation and/or flutter (57.5% vs. 31.1%, *P*-value: 0.0155), and parallelly suffered from reduced renal function (estimated glomerular filtration rate: 53.3 ± 22.3; 95% CI: 46.8–59.3 mL/min vs. 65.4 ± 19.5; 95% CI: 60.1–70.2 mL/min, *P*-value: 0.0214). Notably, no general association between deteriorating cardiac output and worsening of renal function could be described by correlation analysis (*R*: 0.10, *P*-value: 0.3117),

nor did patients with reduced cardiac output generally display impairments of renal function (Supplementary material online, *Figure S2A and B*). An illustration of 20 exemplifying profiles per cluster is provided in *Figure 3F*.

Conventional dichotomization according to AVG$_{mean}$ results in loss of prognostic resolution

In order to compare unsupervised clustering of aortic outflow velocity profiles with a traditional approach of hand-crafted

categorization, the derivation cohort was conventionally dichotomized according to AVG$_{mean}$ (Figure 4A). Survival analysis confirmed that patients with AVG$_{mean}$ <40 mmHg (n = 54, 53.5%) died earlier, but no statistical significance was reached (HR for 2-year mortality: 1.8; 95% CI: 0.59–5.2) (Figure 4B). Apart from identifying well-established predictors for mortality, such as deteriorating renal function, male sex, and EuroSCORE, univariate Cox regression analysis also confirmed the prognostic value of left ventricular ejection fraction, mPAP, and TAPSE. At the same time, no significant association between AVG$_{mean}$ or AVG$_{max}$, on the one hand, and 2-year all-cause mortality, on the other hand, could be detected by regression analysis (Table 3).

An extreme gradient boosting algorithm enables cluster assignment in future patients and confirms that left ventricular compensation capacity rather than the actual obstruction of the aortic valve determines fate after transcatheter aortic valve replacement

To test whether the cluster-related phenotypes as detected by the convolutional part of the pre-trained VGG-16 model in conjunction with PCA and k-means clustering could also be found among the remaining 265 patients with either poor quality or no available Doppler tracings [56 (15.3%) of 366 patients had no Doppler tracings as raw data available], an extreme gradient boosting algorithm was trained on a comprehensive set of functional and structural parameters from pre-procedural echocardiography and right heart catheterization. In total, 12 variables, ideally covering all stages of cardiac and pulmonary circulatory conditions as previously described,[16] served as input data. Moreover, the actual obstruction of the aortic valve expressed as AVA was included as a thirteenth input variable (Supplementary material online, Figure S3 for a complete list of input variables). Since the derivation cohort was predominantly composed of patients assigned to cluster 1 (60.4%), a minority class over-sampling technique (SMOTE) was applied to create a balanced data set (Figure 5). After application of SMOTE, a training and a test set were randomly defined using a 0.75:0.25 split ratio, meaning that 120 'patients' were assigned to the training set and 40 'patients' were assigned to the test set. As a holdout data set, this test set was designated to finally assess the extreme gradient boosting algorithm's performance, before eventually using the trained algorithm for patient-to-cluster assignment in the validation cohort. The purpose of the validation cohort was to evaluate cluster-related survival differences as they were observed for the clusters that have been segregated during the first, unsupervised machine learning experiment among the derivation cohort. In total, 2.44% of the 1313 data points related to 101 patients from the derivation cohort had missing values for those 13 variables (Supplementary material online, Figure S3A), and the largest proportion of missing values was found for measurements of right midventricular diameter (12.9% of values missing) (Supplementary material online, Figure 3B). After imputing missing values, initially observed and later imputed values for right
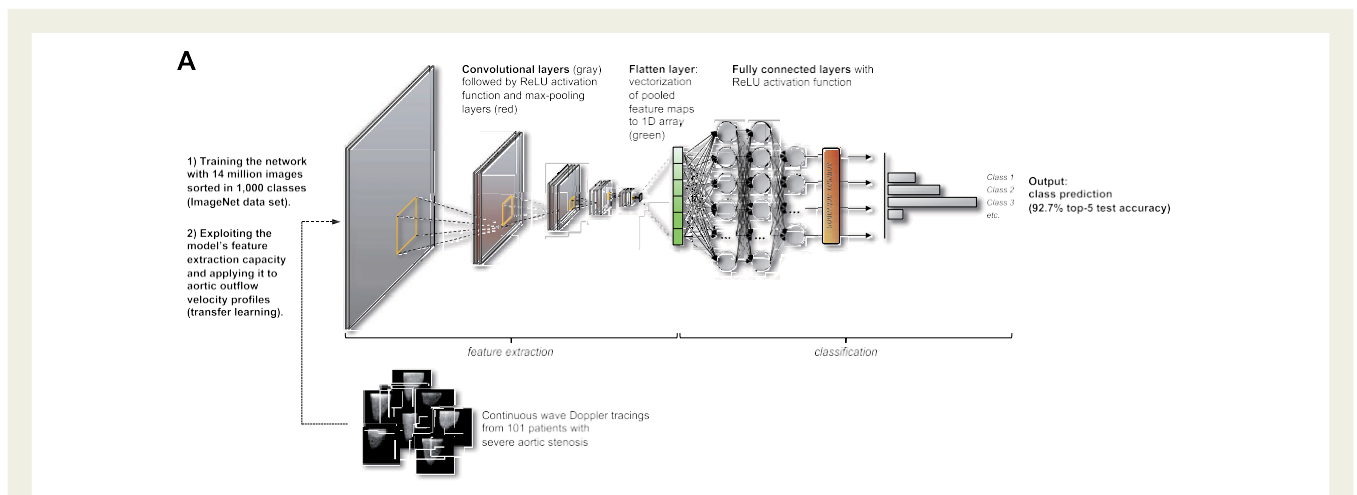
Figure 3 A convolutional neural network followed by PCA and unsupervised k-means clustering provides the proof-of-principle that two subgroups of patients with severe AS can be distinguished according to the aortic outflow velocity profile. (A) VGG-16 network architecture (schematic). The VGG-16 network can be split into two parts: 13 convolutional layers constitute the first part, through which each image is passed through for feature extraction. The convolutional layers are followed by three fully connected layers for classification, and the last layer uses a softmax activation function for final class prediction. Since the aortic outflow velocity profiles were no established class within the ImageNet data set, the classification part of VGG-16 was omitted after pre-training, and hence only the model's feature extraction capacity was exploited in order to transform aortic outflow velocity profiles to 1D arrays (flatten layer), which were subsequently used for unsupervised clustering. (B) PCA of 1D arrays from 101 aortic outflow velocity profiles. (C) Scatter plot including 95% confidence ellipse in order to illustrate cardiac output and mean aortic valve gradient in accordance with cluster assignment. (D) Kaplan–Meier survival analysis in accordance with cluster assignment. (E) Bee swarm plots for comparison of baseline echocardiographic and haemodynamic data. (F) Representative aortic outflow velocity profiles in accordance with cluster assignment. AVA, aortic valve area; AVG$_{mean}$, mean aortic valve gradient; LA area, left atrial area; LVEDD, left ventricular end-diastolic diameter; mPAP, mean pulmonary artery pressure; RA area, right atrial area; ReLU, Rectified Linear Unit; TAPSE, tricuspid annular plane systolic excursion.

Continued
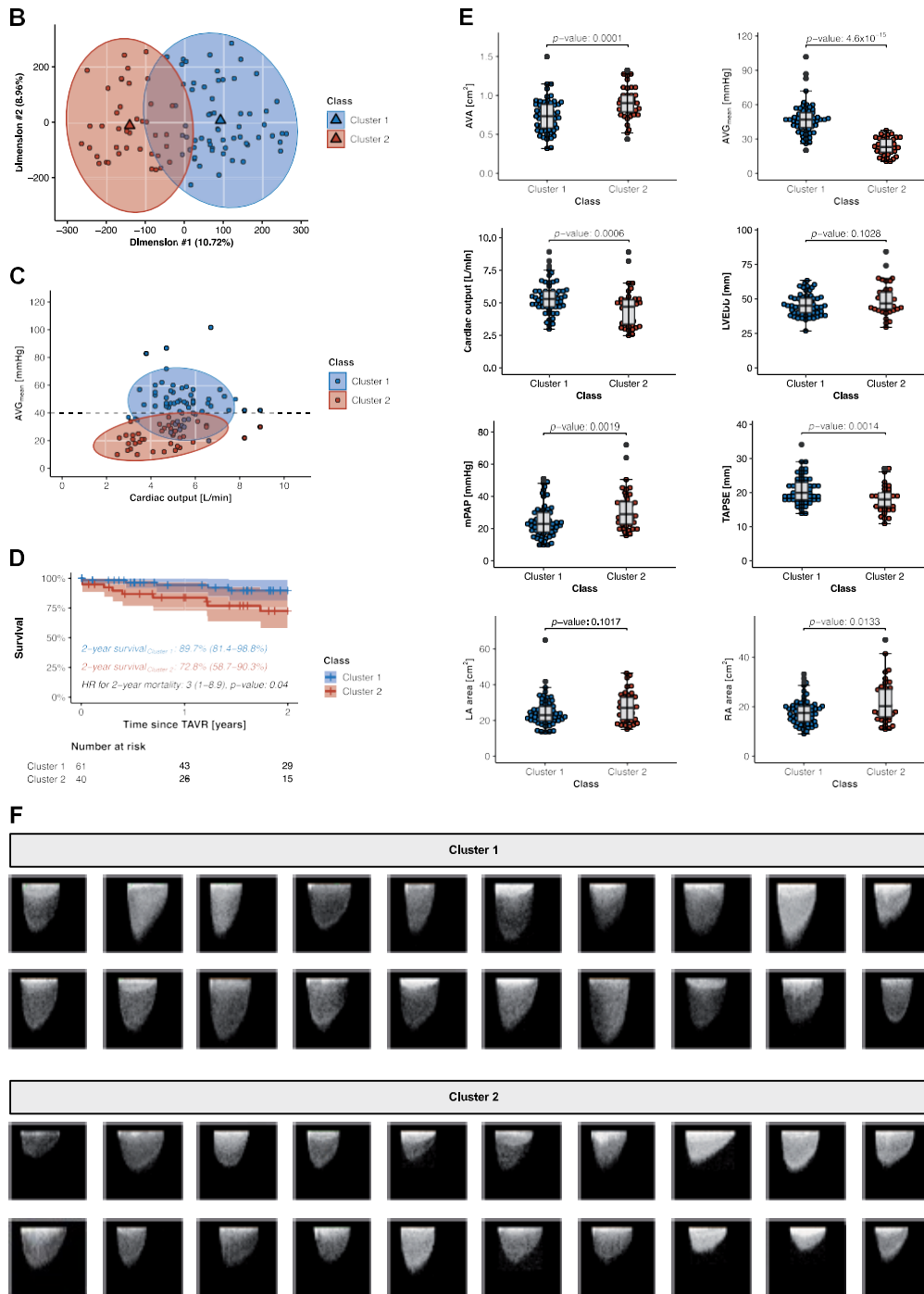
Figure 3 (*Continued*)

midventricular diameter displayed a similar distribution (29.0 ± 6.46; 95% CI: 27.6-30.3 mm vs. 27.6 ± 2.45; 95% CI: 26.1-29.1 mm, *P*-value: 0.5256) (Supplementary material online, *Figure S3C* and *D*). Importantly, the main characteristics of Clusters 1 and 2 in terms of cardiac output, AVG$_{mean}$ and AVA were preserved after over-sampling (*Figure 6A*). An extreme gradient boosting algorithm for cluster assignment was hereinafter trained on 58 instances for 'Cluster 1' and on 62 instances 'Cluster 2', respectively, and it reached an accuracy of 97.5%, significantly outperforming the no information

rate (*P*-value: 1.4×10^{-9}), as demonstrated in the test set of 40 'patients' (*Figure 6B*). Notably, AVG$_{mean}$ showed by far the highest global feature importance for cluster prediction as determined by SHAP values (*Figure 6C*). Applying the trained extreme gradient boosting algorithm to the validation cohort of 265 patients (*Figure 5*) enabled identification of patients belonging to high-risk Cluster 2. Again, those patients were characterized by a functionally and structurally failing left heart in conjunction with PH and right heart impairment (*Table 4*). Compared to patients from Cluster 1, survival was

Figure 4 Conventional dichotomization of the study population in accordance with elevation in mean aortic valve gradient. (A) Scatter plot illustrating cardiac output and mean aortic valve gradient after dichotomization in accordance with elevation in mean aortic valve gradient. (B) Kaplan-Meier survival analysis in accordance with elevation in mean aortic valve gradient. AVG_{mean}, mean aortic valve gradient.

subsequently reduced (*Figure 6D*), and the hazard ratio for 2-year mortality after TAVR was significantly increased (2.6, 95% CI: 1.4-5.1, P-value: 0.004). Importantly, a less severe obstruction of the aortic valve was found again in patients assigned to high-risk cluster 2 (AVA: 0.839 ± 0.219, 0.789-0.889 in Cluster 2 vs. 0.742 ± 0.186, 0.715-0.769 in Cluster 1, P-value: 0.0007) (*Table 4*), confirming the initially surprising finding from the derivation cohort (*Figure 6E*).

Discussion

Transfer learning exploiting big data could be key to overcome the obstacle of data scarcity as commonly encountered in clinical reality, and learning from a related problem aids in gaining novel insights into phenotypic presentations of patients with severe aortic stenosis

Identifying patients at risk is a core element in the practice of medicine, but risk stratification for patients with severe AS in contemporary clinical practice is often limited by hypothesis-driven selection of a few factors typically regarded in isolation, by suggesting a model of orderly progression of accumulated pathologies upstream of the causative AS, or by the assumption of a parametric linear relationship between predictor variable and outcome. This study demonstrates that prognostic resolution of survival in patients with severe AS undergoing TAVR can be refined by harnessing the intriguing feature extraction capacity from an established CNN pre-trained on big data in order to subsequently recognize complex geometries in aortic outflow velocity profiles, which integrate crucial information about left ventricular contractility and aortic valve obstruction. Thus, two major phenotypes with important clinical implications

could be unravelled. The main messages from our study are therefore as follows (*Graphical Abstract*):

(1) Transfer learning has the potential to unearth hidden gems even in clinical data sets of limited size.
(2) Not so much the actual stenosis of the aortic valve expressed as AVA determines the prognosis after TAVR, but the left ventricular compensation capacity and subsequent development of PH and right heart failure stratify patients into low-risk and high-risk cohorts.

On the drawbacks of traditional methods for risk assessment—and how machine learning technology can pave the way to personalized risk stratification prior to transcatheter aortic valve replacement

In order to illustrate the almost ubiquitous problem of data scarcity in medical research on the one hand, and the vast potential of transfer learning, on the other hand, the number of aortic outflow velocity profiles to be analysed was intentionally kept small. Possibly, differences in survival after dichotomization according to AVG_{mean} would have become statistically significant, if more patients were included. Nonetheless, dichotomization of continuous variables is prone to reducing statistical power without notable benefit (oversimplification), and physicians in a real-world scenario therefore rarely rely on a single variable's dichotomy for decision-making or prognostic assessment, but rather prefer context-specific interpretation of extensive (raw) data. Aiming to detect predictors of mortality among a similar cohort of patients with severe AS undergoing TAVR, Weber et al.[17] analysed a set of echocardiographic and haemodynamic data, and identified presence of combined pre- and post-capillary PH, and a lower AVG_{mean} as independent predictors by using multivariate Cox regression analysis. However, traditional regression models assume a parametric linear function relating the predictor variables with the response. This assumption might not hold true in the natural
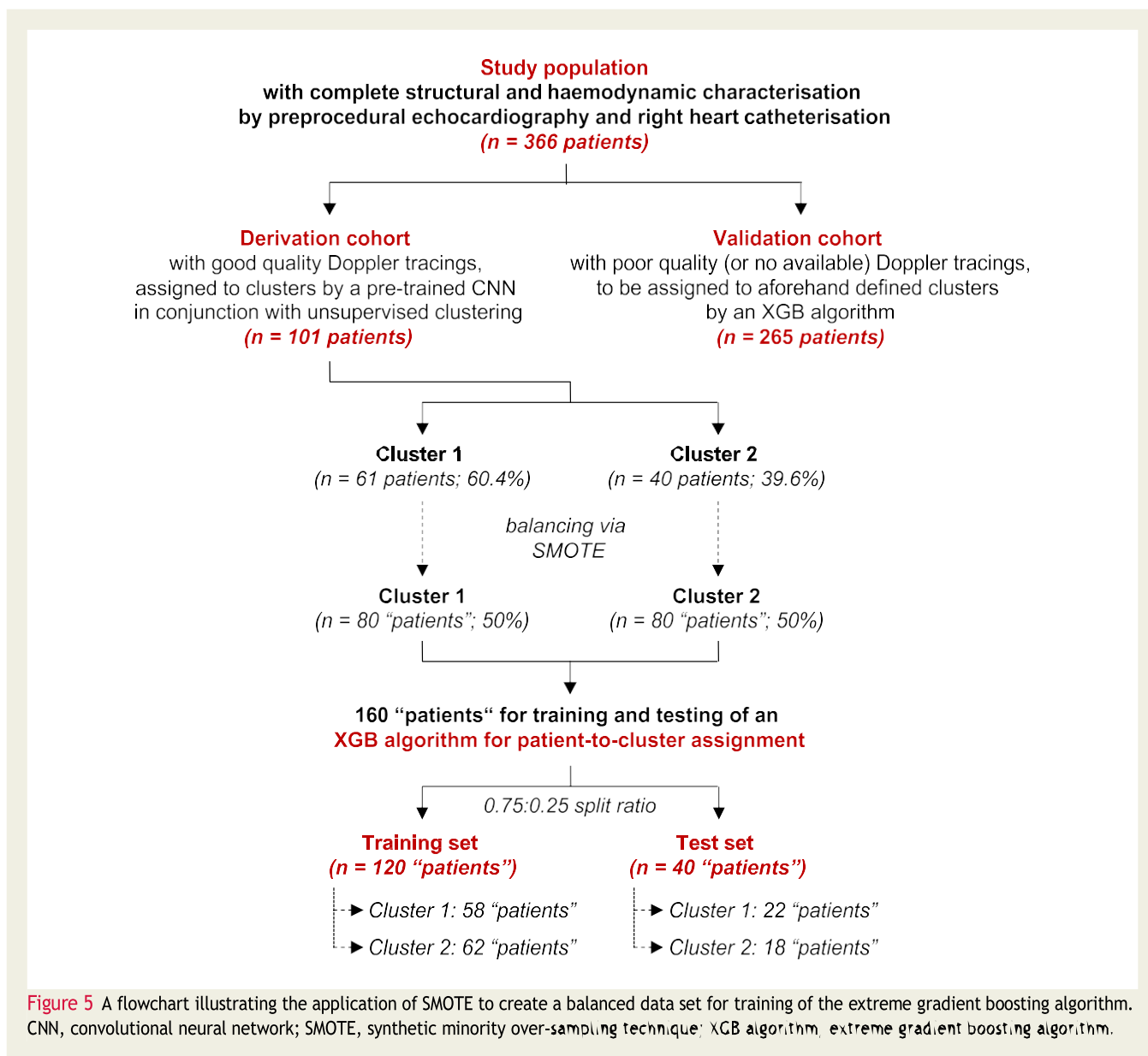
Table 3 Univariate and multivariate cox regression analysis with 2-year mortality as a dependent variable (derivation cohort)

	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value
Age	0.96 (0.89–1) per year	0.21		
Sex (female)	0.26 (0.074–0.95)	0.042	0.20 (0.04–1.02)	0.0531
BMI	0.93 (0.81–1.1) per kg/m^2	0.26		
Arterial hypertension	1.9 (0.24–14)	0.55		
Smoking	1 (0.33–3)	1		
Diabetes mellitus	1.9 (0.62–5.5)	0.27		
NYHA functional class	1.4 (0.57–3.2) per class	0.49		
EuroSCORE	1 (1–1.1) per %	0.0017	1.03 (0.99–1.08) per %	0.1374
eGFR	0.97 (0.95–1) per mL/min	0.024	0.97 (0.94–1.00) per mL/min	0.0831
Hb	1 (0.75–1.4) per g/dL	0.95		
CAD	0.91 (0.2–4.1)	0.91		
COPD	2.1 (0.6–7.7)	0.24		
Atrial fibrillation and/or flutter	3.7 (1.2–12)	0.026	1.71 (0.35–8.50)	0.5099
AVA	1.4 (0.16–13) per cm^2	0.75		
AVG$_{max}$	0.98 (0.96–1) per mmHg	0.13		
AVG$_{mean}$	0.97 (0.94–1) per mmHg	0.11		
Cardiac output	0.66 (0.42–1) per L/min	0.064		
LVEF	0.94 (0.9–0.97) per %	0.0004	0.96 (0.84–1.09) per %	0.4913
LVEDD	1.1 (1–1.1) per mm	0.0056	0.95 (0.87–1.05) per mm	0.3149
mPAP	1 (1–1.1) per mmHg	0.04	1.00 (0.94–1.08) per mmHg	0.9269
mPCWP	1 (1–1.1) per mmHg	0.073		
PVR	1.3 (1–1.6) per WU	0.045	0.81 (0.47–1.40) per WU	0.4539
TAPSE	0.85 (0.73–1) per mm	0.045	0.85 (0.67–1.09) per mm	0.1967
Right midventricular diameter	1.1 (1–1.2) per mm	0.02	1.06 (0.96–1.18) per mm	0.2344
LA area	1.1 (1–1.1) per cm^2	0.011	1.03 (0.93–1.14) per cm^2	0.5449
RA area	1.1 (1.1–1.2) per cm^2	7.7×10^{-5}	1.04 (0.94–1.16) per cm^2	0.4110
Low gradient (AVG$_{mean}$ < 40 mmHg)	1.8 (0.59–5.2)	0.314		
LV dysfunction (LVEF ≤ 45%)	4.7 (1.6–14)	0.0052	0.44 (0.03–7.33)	0.5663
PH (mPAP ≥ 25 mmHg)	1.6 (0.53–4.7)	0.42		
RV dysfunction (TAPSE ≤ 16 mm)	2.3 (0.77–6.8)	0.14		
Assignment to Cluster 2	3 (1–8.9)	0.04	1.12 (0.29–4.37)	0.8676

AVA, aortic valve area; AVG$_{max}$, maximum aortic valve gradient; AVG$_{mean}$, mean aortic valve gradient; BMI, body mass index; CAD, coronary artery disease; CI, confidence interval; COPD, chronic obstructive pulmonary disease; GFR, glomerular filtration rate; HR, hazard ratio; IVS, interventricular septum thickness; LA area, left atrial area; LVEDD, left ventricular end-diastolic diameter; LVEF, left ventricular ejection fraction; LVESD, left ventricular end-systolic diameter; mean RV pressure, mean right ventricular pressure; mPAP, mean pulmonary artery pressure; mPCWP, mean postcapillary wedge pressure; NYHA, New York Heart Association; PVR, pulmonary vascular resistance; PW, posterior wall thickness; RA area, right atrial area; RA pressure, right atrial pressure; TAPSE, tricuspid annular plane systolic excursion.

course of AS, since the AVG$_{mean}$ initially increases with progressive narrowing of the AVA, but later decreases as the left ventricle decompensates ('low flow, low gradient AS'). If patients on the transition from moderate to severe AS were analysed by logistic regression analysis, an increasing AVG$_{mean}$ would clearly be interpreted as an indicator for disease progression, and hence serve as a marker for worsened prognosis.[18] The beauty of the hereby established approach lays in the improvement to identify and segregate patients with similar characteristics firstly without applying any *a priori* assumption and secondly without restricting the analysis to human-selected patient characteristics as data features. At the same time, many Doppler tracings were not suitable to be analysed by a CNN due to poor echocardiographic acquisition or due to suboptimal alignment

with the jet and hence inadequate recording of the true transvalvular gradient. So, how can our assignment to distinct clusters and its clinical implication be generalized to the majority of patients? Ideally, this study is not only perceived as a proof-of-concept valid for selected patients, but as yet another step along the road to implementation of artificial intelligence in clinical decision-making. We have therefore decided to additionally train an extreme gradient boosting algorithm on functional and structural data from preprocedural echocardiography and right heart catheterization, thus opening the avenue for other cardiologists to stratify their patients according to our beforehand created classification generated by transfer learning. Upon loading the trained extreme gradient boosting algorithm and adding the requested input data into

Study population
with complete structural and haemodynamic characterisation
by preprocedural echocardiography and right heart catheterisation
(n = 366 patients)

Derivation cohort
with good quality Doppler tracings,
assigned to clusters by a pre-trained CNN
in conjunction with unsupervised clustering
(n = 101 patients)

Validation cohort
with poor quality (or no available) Doppler tracings,
to be assigned to aforehand defined clusters
by an XGB algorithm
(n = 265 patients)

Cluster 1
(n = 61 patients; 60.4%)

Cluster 2
(n = 40 patients; 39.6%)

balancing via
SMOTE

Cluster 1
(n = 80 "patients"; 50%)

Cluster 2
(n = 80 "patients"; 50%)

160 "patients" for training and testing of an
XGB algorithm for patient-to-cluster assignment

0.75:0.25 split ratio

Training set
(n = 120 "patients")

› *Cluster 1: 58 "patients"*

› *Cluster 2: 62 "patients"*

Test set
(n = 40 "patients")

› *Cluster 1: 22 "patients"*

› *Cluster 2: 18 "patients"*

Figure 5 A flowchart illustrating the application of SMOTE to create a balanced data set for training of the extreme gradient boosting algorithm. CNN, convolutional neural network; SMOTE, synthetic minority over-sampling technique; XGB algorithm, extreme gradient boosting algorithm.

the corresponding R code (both available from the corresponding author; see Supplementary material online, *Figure S4* for a preview of the R code), future patients can be assigned to either Cluster 1 (good prognosis) or Cluster 2 (poor prognosis).

The extent of cardiac damage is already mirrored in the aortic outflow velocity profile, and it is the left ventricular response to the increased afterload that determines fate in patients with severe aortic stenosis

Capturing the complexity of cardiac damage subsequent to severe AS is key to sophisticated risk stratification prior to TAVR. This is particularly true, as PH and right ventricular dysfunction can persist in a substantial number of cases after TAVR, and persistence

translates into distressing mortality.[19-21] G´en´ereux *et al.*[7] therefore established a staging classification, which considers disease progression beyond the compensation capacity of the left ventricle. This divisive, top-down staging classification is driven by the hypothesis that extravalvular damages to the heart and pulmonary circulation subsequent to severe AS occur in a sequential order of left heart failure, PH, and right heart dysfunction. Despite its simplicity, this staging classification cannot be easily implemented into clinical practice, as clinicians commonly encounter disparities between AS-induced haemodynamic burden and extravalvular damages (possibly influenced by comorbidities, such as atrial fibrillation and chronic obstructive pulmonary disease, or by genetic predisposition).[16,22,23] Failure of left ventricular compensation capacity and subsequent backwards transmission of elevated left-sided filling pressures was more frequently observed in Cluster 2 than in Cluster 1 (mean post-capillary wedge pressure: 21.4 ± 8.83, 95% CI:

18.6–24.2 mmHg vs. 15.5 ± 8.40, 95% CI: 13.3–17.6 mmHg, *P*-value: 0.0007), whilst patients in Cluster 1 presented with a more severe obstruction of the aortic valve, indicating a longer disease progression, yet resulting in less cardiopulmonary impairments. This insight into disease progression in patients with severe AS emphasizes the importance of the complex pathophysiologic valvular-ventricular interactions, which obviously vary among individuals. In the contemporary 'one-size-fits-all' practice of medicine, the timing of intervention mainly focuses on the aortic valve. Our study may alter the perception of ideal timing of intervention as well as it may facilitate the development of individualized treatment, as an earlier intervention in patients from Cluster 2 might have had prevented further aggravation of left heart decompensation, PH, and right heart dysfunction. Addressing a similar issue, a study investigating the benefit of early intervention in patients with moderate AS and impaired left ventricular function has already been initialized [TAVR UNLOAD (Transcatheter Aortic Valve Replacement to Unload the Left Ventricle in Patients with Advanced Heart Failure) trial].[24] Moreover, it will be interesting to analyse future echocardiographic follow-up studies in accordance with cluster assignment, since recovery from cardiopulmonary damages that cannot be totally attributed to the obstruction of the aortic valve seems questionable. Thus, suspected persistence of PH and right heart dysfunction despite correction of severe AS by TAVR could emerge as an unmodifiable (?) driver for increased mortality in patients from Cluster 2.

Unsupervised clustering could reveal diversity of aortic stenosis phenotypes with unprecedented precision, but extensive quality control is mandatory before unleashing machine learning algorithms in clinical practice

Extending this proof-of-principle study based on good quality Doppler tracings from 101 patients to a larger cohort could reveal even more diversity in aortic outflow velocity profiles by unravelling additional clusters. Even nowadays, AS with discordant markers of severity, such as severely reduced AVA and low AVG$_{mean}$, but preserved left ventricular ejection fraction ('paradoxical low-gradient AS')[25] remains a conundrum in diagnosis and treatment.[26] It will be interesting to see whether contemporary classifications of AS phenotypes will be mirrored by unsupervised clustering, or if distinct clinical presentations will emerge. Admittedly, involvement of artificial intelligence in clinical decision-making is still frowned upon due to the 'black box' nature, and the potential for a flawed machine learning algorithm to induce iatrogenic harm is vast. The opaqueness in the
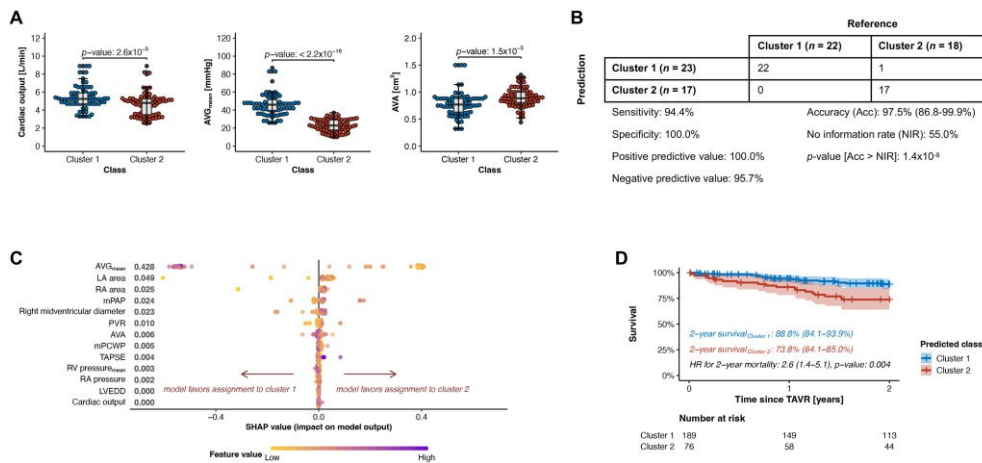
Figure 6 An extreme gradient boosting algorithm opens the perspective to assign patients to beforehand defined clusters by a comprehensive set of functional and structural parameters of cardiac and pulmonary circulatory conditions. (*A*) Bee swarm plots for comparison of key characteristics between clusters after over-sampling (SMOTE). (*B*) Confusion matrix (test set). (*C*) Shedding light on the black box of extreme gradient boosting algorithm-mediated cluster assignment by calculating SHAP (SHapley Additive exPlanations) values for its input variables. The *y*-axis represents the in- put variables in descending order of global feature importance, whilst the *x*-axis indicates the adjustment to the predicted cluster. Moreover, each dot in this sina plot represents an observation, i.e. a patient from the derivation cohort, and the gradient colour denotes the value of the respective in- put variable. Therefore, if the dots on one side of the central line are increasingly yellow or purple, that suggests that increasing values or decreasing values, respectively, move the predicted cluster in the respective direction (left: Cluster 1; right: Cluster 2). For instance, higher values of AVG$_{mean}$ (purple dots) are associated with assignment to Cluster 1. (*D*) Kaplan–Meier survival analysis in accordance with extreme gradient boosting-algorithm-mediated cluster assignment (validation cohort). (*E*) Comparison of clusters as defined by the CNN in conjunction with PCA and k-means clustering (derivation cohort; red) or as determined by the trained extreme gradient boosting algorithm (validation cohort; blue). The central line in each box plot denotes the median value, while the box contains all values ranging between the 25th and 75th percentiles of the data set. The black whiskers mark the 5th and 95th percentiles, and values falling beyond these upper and lower bounds are considered outliers, plotted as black dots. AVA, aortic valve area; AVG$_{mean}$, mean aortic valve gradient; LA area, left atrial area; LVEDD, left ventricular end-diastolic diameter; mPAP, mean pulmonary artery pressure; mPCWP, mean postcapillary wedge pressure; PVR, pulmonary vascular resistance; RA area, right atrial area; RA pressure, right atrial pressure; RV pressure$_{mean}$, mean right ventricular pressure; TAPSE, tricuspid annular plane systolic excursion.
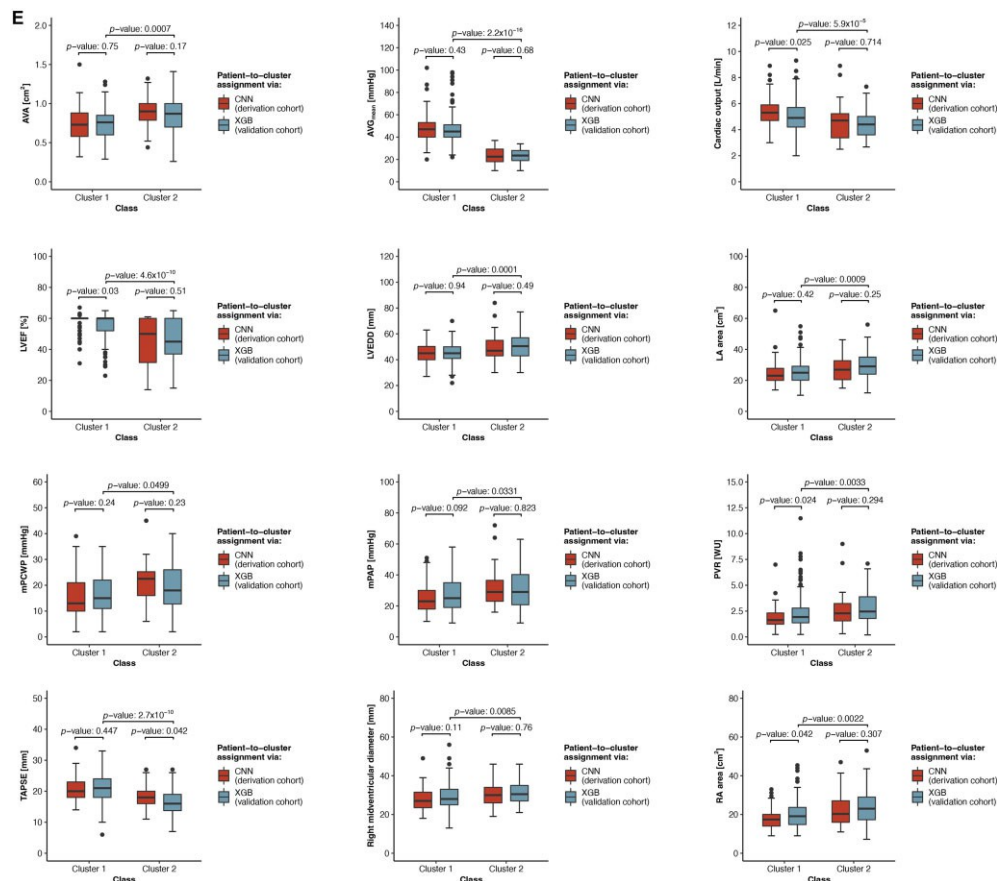
Figure 6 (*Continued*)

determination of output has therefore fuelled demands for explainability as expressed in the European Union's General Data Protection Regulation.[27] Gradient-weighted class activation mapping (GRAD-CAM) visualizations from the tool box of explainable artificial intelligence are typically applied in order to inspect images and to get insights into CNN decisions.[28] Yet, GRAD-CAM visualizations connecting the raw image to the decision of a classifier were not used in this study for two reasons:

(1) Patient-to-cluster assignment was based on PCA and k-means clustering of Doppler tracings, which represents a form of unsupervised learning, and which is different from assignment by means of a trained classifier, which would have represented a form of supervised learning.

(2) Training a classifier in terms of fully connected layers following the convolutional part of the pre-trained VGG-16 network would have required thousands of Doppler tracings, which could have only been collected in a labour-intensive, multicentric effort.

It, therefore, remains enigmatic which characteristics of the aortic outflow velocity profile would result in assignment to either Cluster 1 or 2. As shown by the PCA (*Figure 3B*), there cannot be a single 'most important' feature that defines the echocardiographic signature of patients assigned to Cluster 1 or 2, as the first two dimensions of the PCA explain only 10.72% and 8.96% of the variation among all

transformed aortic outflow velocity profiles, respectively. This gap in mechanistic inference can be perceived as a limitation to this study, but it also demonstrates the strengths of neural networks, which enable to identify novel relationships in complex and finely nuanced data sets and which therefore go beyond (simplified) stratification in accordance with human-selected features.[29] To explain at least partially which feature within the aortic outflow velocity profiles drive the differences between Clusters 1 and 2, the 20 most distant Doppler tracings (hereinafter referred to as 'top 10' and 'bottom 10' Doppler tracings) along PCA dimension #1 were identified (Supplementary material online, *Figure S5A–C*) and related echocardiographic and haemodynamic characteristics were compared (Supplementary material online, *Figure S5D* and *Table S4*): among the studied characteristics, the strongest difference in terms of statistical significance expressed as the respective P-value level was found for AVG_{mean} (59.5±15.6; 95% CI: 48.3–70.7 mmHg among Top 10 Doppler tracings vs. 17.9±7.62; 95% CI: 12.4–23.4 mmHg among bottom 10 Doppler tracings, P-value: 0.0002). This finding was confirmed by direct comparison of top 10 and bottom 10 Doppler tracings (Supplementary material online, *Figure S5E*).

Scrutinizing the generalizability of our findings as generated on 101 selected patients, we decided to test if cluster-related phenotypes could also be detected among the initially excluded 265 patients due

Table 4 Comparison of echocardiographic and haemodynamic characteristics in accordance with cluster assignment (validation cohort)

		Class		
	All (*n* 5 265)	Cluster 1 (*n* 5 189)	Cluster 2 (*n* 5 76)	*P*-value
AVA (cm^2), mean ± SD [95% CI]	0.770 ± 0.201 [0.746–0.794]	0.742 ± 0.186 [0.715–0.769]	0.839 ± 0.219 [0.789–0.889]	0.0007
AVG$_{mean}$ (mmHg), mean ± SD [95% CI]	40.4 ± 15.3 [38.5–42.2]	47.1 ± 12.3 [45.3–48.9]	23.5 ± 6.00 [22.1–24.9]	<2.2 × 10^{-16}
Cardiac output (L/min), mean ± SD [95% CI]	4.86 ± 1.19 [4.72–5.01]	5.04 ± 1.23 [4.86–5.22]	4.42 ± 0.970 [4.19–4.64]	5.9 × 10^{-5}
LVEF (%), mean ± SD [95% CI]	52.6 ± 10.8 [51.3–53.9]	55.5 ± 8.05 [54.3–56.7]	45.4 ± 13.1 [42.4–48.4]	4.6 × 10^{-10}
LVEDD (mm), mean ± SD [95% CI]	46.9 ± 8.16 [45.8–47.9]	45.4 ± 7.41 [44.3–46.5]	50.4 ± 8.86 [48.2–52.5]	0.0001
LA area (cm^2), mean ± SD [95% CI]	26.5 ± 8.33 [25.4–27.6]	25.3 ± 7.85 [24.1–26.5]	29.3 ± 8.82 [27.2–31.4]	0.0009
mPCWP (mmHg), mean ± SD [95% CI]	17.3 ± 8.28 [16.3–18.3]	16.6 ± 7.56 [15.5–17.6]	19.2 ± 9.66 [17.0–21.4]	0.0499
mPAP (mmHg), mean ± SD [95% CI]	28.5 ± 11.4 [27.1–29.9]	27.5 ± 10.8 [25.9–29.0]	31.1 ± 12.5 [28.2–34.0]	0.0331
PVR (WU), mean ± SD [95% CI]	2.48 ± 1.59 [2.29–2.67]	2.34 ± 1.58 [2.11–2.56]	2.83 ± 1.57 [2.47–3.19]	0.0033
TAPSE (mm), mean ± SD [95% CI]	19.6 ± 5.36 [18.9–20.3]	21.0 ± 5.10 [20.2–21.7]	16.3 ± 4.53 [15.3–17.4]	2.7 × 10^{-10}
Right midventricular diameter (mm), mean ± SD [95% CI]	29.7 ± 6.65 [28.8–30.5]	29.0 ± 6.89 [28.0–30.1]	31.3 ± 5.81 [29.9–32.6]	0.0085
RA area (cm^2), mean ± SD [95% CI]	21.0 ± 7.68 [20.0–22.0]	19.9 ± 6.95 [18.9–21.0]	23.5 ± 8.75 [21.4–25.6]	0.0022

AVA, aortic valve area; AVG$_{mean}$, mean aortic valve gradient; CI, confidence interval; LA area, left atrial area; LVEDD, left ventricular end-diastolic diameter; LVEF, left ventricular ejection fraction; mPAP, mean pulmonary artery pressure; mPCWP, mean postcapillary wedge pressure; PVR, pulmonary vascular resistance; RA area, right atrial area; RV dysfunction, right ventricular dysfunction; SD, standard deviation; TAPSE, tricuspid annular plane systolic excursion.

to poor or missing Doppler tracings. The cluster-related clinical implications could be confirmed by an extreme gradient boosting algorithm, and calculation of SHAP values as a state-of-the-art metric to quantify the contribution of input variables to model prediction highlighted the importance of transvalvular gradients, incorporating information about both aortic valve obstruction and left ventricular contractility. Notably, continuous wave Doppler echocardiography in combination with the Bernoulli equation to assess transvalvular pressure gradients is based on oversimplification of human haemodynamics, as for instance a column of flow with uniform velocity distribution is assumed, which is clearly not the case in patients with severe AS.[30] The analysis of the spatio-temporal pattern of the ejection jet, e.g. by three-dimensional cardiovascular magnetic resonance imaging, could therefore reveal novel insights into AS phenotypes.

Limitations: on the prohibitive costs of poor image quality, and why you should not trust artificial intelligence implicitly

Machine learning algorithms *per se* learn from data, meaning that insufficient data quality or systematic bias during data collection would hamper the algorithm to identify any consistent and generalizable patterns. The accuracy of a CNN therefore strictly relies on the input data quality. Physicians in a real-world scenario yet commonly encounter difficulties in examining patients with severe AS, as they typically present dyspnoeic and are hence less suited for optimal positioning for echocardiography. It was, therefore, important to demonstrate that the subset of 101 patients with good quality Doppler tracings was representative of the entire study population of 366 patients (Supplementary material online, *Tables S2* and *S3*). Moreover, we had to ensure by cumbersome manual cropping that the Doppler tracings serving as input images contain no other information than the aortic outflow velocity profile of interest (*Figure 2*). An example of a seemingly high-performance machine learning algorithm flawed by shortcuts in the training set is a model that is supposed to distinguish a wolf from a husky by animal characteristics but eventually reveals to derive its performance from the simple, but undesired identification of patches of snow on the photograph.[31] Moreover, the unscrutinized synthesis of training data from separate data sets of COVID-19-negative and COVID-19-positive images was demonstrated to introduce near worst-case confounding and thus abundant opportunity for machine learning algorithms to learn shortcuts due to variations in image acquisition and radiographic projection.[32]

Claiming to have found a reasonable echocardiographic signature among patients presenting with severe AS, it was therefore of paramount importance to us to validate (and finally confirm) the clinical implications of related phenotypes in a second cohort by yet another machine learning algorithm. Obviously, we cannot guarantee that the algorithms employed in this study would outperform all other algorithms in clustering patients (unsupervised learning experiment) and in assigning patients to clusters (supervised learning experiment). In the commonly accepted absence of any *a priori* guarantee that one machine learning technique is superior to all others,[33] the only way to determine which algorithms works best for the given data structure is to evaluate them all. However, this is practically impossible, innovative and more powerful algorithms might emerge in the future, and it ultimately also relies on the programmer's ability to tune the hyperparameter of respective models to perfection. As a matter of fact, we applied hierarchical agglomerative clustering to the transformed aortic outflow velocity profiles, hence testing yet another popular clustering algorithm equivalent to k-means clustering. Hierarchical agglomerative clustering also facilitated to identify a cluster with significantly reduced AGV_{mean}; however, the two segregated clusters were vastly overlapping as demonstrated by the first two dimensions of a PCA as well as by a correlation plot depicting AGV_{mean} and cardiac output, and subsequently, 2-year survival differences did not reach statistical significance (Supplementary material online, *Figure S6*). Importantly, AS represents a progressive disease with a continuous transition of stages of disease severity. Unlike clustering of e.g. bone marrow cells of distinct haematopoietic lineages (where you would expect clearly defined clusters of e.g. erythrocytes and lymphatic cells based on their gene expression profiles), it is practically impossible to distinguish any clearly separated clusters among patients with severe AS and their respective aortic outflow velocity profiles. This is reflected by the silhouette diagram (Supplementary material online, *Figure S7*) revealing a mean silhouette coefficient of only 0.0689 ± 0.0459 among the clusters as defined by k-means clustering. Moreover, we acknowledge that the accuracy of the extreme gradient boosting algorithm was evaluated on a test set that was at least partially composed of synthetic data. We have therefore added an alternative experimental design with a test set containing only real and unseen patients (explicitly no synthetic data) (Supplementary material online, *Figure S8A*). Thus, we could confirm the satisfying accuracy of the extreme gradient boosting algorithm for patient-to-cluster assignment based on 13 variables from pre-procedural echocardiography and right heart catheterization (accuracy: 92.0%; 95% CI: 74.0–99.9%) (Supplementary material online, *Figure S8B*). Applying the algorithm trained under the alternative experimental design to the validation cohort of patients with poor quality or no available Doppler tracings also confirmed the increased risk of mortality for patients assigned to Cluster 2 in comparison to Cluster 1 (HR for 2-year mortality: 2.1; 95% CI: 1.1–4.1, *P*-value: 0.022) (Supplementary material online, *Figure S8C*). Again, patients assigned to Cluster 2 were characterized by a relatively larger AVA (0.824 ± 0.214; 95% CI: 0.779–0.870 cm^2) and also by a reduced left ventricular function (left ventricular ejection fraction: 47.5 ± 13.3; 95% CI: 44.7–50.4%) and by a lower AVG_{mean}

(25.6 ± 8.14, 95% CI: 23.9–27.4 mmHg) (Supplementary material online, *Figure S8D* and *Table S5*).

Conclusion

In summary, this is the first study to demonstrate the usefulness of transfer learning for unsupervised clustering of aortic outflow velocity profiles in patients with severe AS. Since the perception of patients presenting with severe AS is in a state of flux from a valve-centred perspective to a personalized comprehensive view covering all aspects of co-developed cardiopulmonary impairments, the unravelled phenotypes in this study hold the promise to better stratify patients into low-risk and high-risk cohorts. Importantly, it is the left ventricular response to the increased afterload, not so much the actual obstruction of the aortic valve, that determines fate after TAVR. As a new arrow in the quiver from interventional cardiologists to refine prognostic assessment prior to TAVR, the trained extreme gradient boosting algorithm for individual cluster assignment in future patients can be requested from the corresponding author.

Supplementary material

Supplementary material is available at *European Heart Journal - Digital Health* online.

Data availability

The pre-processed (deidentified) aortic outflow velocity profiles from those 101 patients with good quality Doppler tracings are attached as Supplementary material online. Moreover, data concerning baseline echocardiographic and haemodynamic characteristics (which were used for training of the extreme gradient boosting algorithm) as well as the respective survival status are given for those 101 patients (including a code book). Furthermore, the complete R code and the trained extreme gradient boosting algorithm for patient-to-cluster assignment can be requested from the corresponding author.

References

1. Raghunath S, Ulloa Cerna AE, Jing L, et al. Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. *Nat Med* 2020;26: 886–891.
2. Fries JA, Varma P, Chen VS, et al. Weakly supervised classification of aortic valve malformations using unlabeled cardiac MRI sequences. *Nat Commun* 2019;10: 3111.
3. Diller G-P, Kempny A, Babu-Narayan SV, et al. Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease: data from a single tertiary centre including 10 019 patients. *Eur Heart J* 2019;40:1069–1077.
4. Perez MV, Mahaffey KW, Hedlin H, et al.; Apple Heart Study Investigators. Large-scale assessment of a smartwatch to identify atrial fibrillation. *N Engl J Med* 2019;381:1909–1917.

5. Kwak S, Lee Y, Ko T, et al. Unsupervised cluster analysis of patients with aortic stenosis reveals distinct population with different phenotypes and outcomes. *Circ Cardiovasc Imaging* 2020;13.

6. Sengupta PP, Shrestha S, Kagiyama N, et al.; Artificial Intelligence for Aortic Stenosis at Risk International Consortium. A machine-learning framework to identify distinct phenotypes of aortic stenosis severity. *JACC Cardiovasc Imaging* 2021;14:1707–1720.

7. G´en´ereux P, Pibarot P, Redfors B, et al. Staging classification of aortic stenosis based on the extent of cardiac damage. *Eur Heart J* 2017;38:3351-3358.

8. Makkar RR, Fontana GP, Jilaihawi H, et al. Transcatheter aortic-valve replacement for inoperable severe aortic stenosis. *N Engl J Med* 2012;366:1696-1704.

9. Baumgartner H, Falk V, Bax JJ, et al.; ESC Scientific Document Group. 2017 ESC/EACTS Guidelines for the management of valvular heart disease. *Eur Heart J* 2017;38:2739-2791.

10. Bing R, Gu H, Chin C, et al. Determinants and prognostic value of echocardiographic first-phase ejection fraction in aortic stenosis. *Heart* 2020;106:1236-1243.

11. Chambers J. Low 'gradient', low flow aortic stenosis. *Heart* 2006;92:554-558.

12. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *ArXiv14091556 Cs* 2015.

13. Stekhoven DJ, Buhlmann P. MissForest-non-parametric missing value imputation for mixed-type data. *Bioinformatics* 2012;28:112-118.

14. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 2002;16:321-357.

15. Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell* 2020;2:56-67.

16. Lachmann M, Rippen E, Schuster T, et al. Subphenotyping of patients with aortic stenosis by unsupervised agglomerative clustering of echocardiographic and hemodynamic data. *JACC Cardiovasc Interv* 2021;14:2127-2140.

17. Weber L, Rickli H, Haager PK, et al. Haemodynamic mechanisms and long-term prognostic impact of pulmonary hypertension in patients with severe aortic stenosis undergoing valve replacement: Impact of PH in severe aortic stenosis. *Eur J Heart Fail* 2019;21:172-181.

18. Slimani A, Roy C, de Meester C, et al. Structural and functional correlates of gradient-area patterns in severe aortic stenosis and normal ejection fraction. *JACC Cardiovasc Imaging* 2021;14:525-536.

19. Masri A, Abdelkarim I, Sharbaugh MS, et al. Outcomes of persistent pulmonary hypertension following transcatheter aortic valve replacement. *Heart* 2018;104:821-827.

20. Cremer PC, Zhang Y, Alu M, et al. The incidence and prognostic implications of worsening right ventricular function after surgical or transcatheter aortic valve replacement: insights from PARTNER IIA. *Eur Heart J* 2018;39:2659-2667.

21. Asami M, Stortecky S, Praz F, et al. Prognostic value of right ventricular dysfunction on clinical outcomes after transcatheter aortic valve replacement. *JACC Cardiovasc Imaging* 2019;12:577-587.

22. Guzzetti E, Annabi M-S, Pibarot P, Clavel M-A. Multimodality imaging for discordant low-gradient aortic stenosis: assessing the valve and the myocardium. *Front Cardiovasc Med* 2020;7:570689.

23. Little SH, O'Gara PT. Considering the hazards of aortic valve stenosis. *JACC Cardiovasc Imaging* 2021;14:1738-1741.

24. Spitzer E, Van Mieghem NM, Pibarot P et al. Rationale and design of the Transcatheter Aortic Valve Replacement to Unload the Left ventricle in patients with Advanced heart failure (TAVR UNLOAD) trial. *Am Heart J* 2016;182:80-88.

25. Hachicha Z, Dumesnil JG, Bogaty P, Pibarot P. Paradoxical low-flow, low-gradient severe aortic stenosis despite preserved ejection fraction is associated with higher afterload and reduced survival. *Circulation* 2007;115:2856-2864.

26. Clavel M-A, Magne J, Pibarot P. Low-gradient aortic stenosis. *Eur Heart J* 2016;37:2645-2657.

27. Goodman B, Flaxman S. European union regulations on algorithmic decision-making and a "right to explanation". *AI Mag* 2017;38:50-57.

28. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. *Int J Comput Vis* 2020;128:336-359.

29. Altes A, Thellier N, Bohbot Y, et al. Relationship between the ratio of acceleration time/ejection time and mortality in patients with high-gradient severe aortic stenosis. *J Am Heart Assoc* 2021;10:e021873.

30. Donati F, Myerson S, Bissell MM, et al. Beyond Bernoulli: improving the accuracy and precision of noninvasive estimation of peak pressure drops. *Circ Cardiovasc Imaging* 2017;10.

31. Ribeiro MT, Singh S, Guestrin C. 'Why should I trust you?': explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* San Francisco, CA, USA: ACM; 2016. p.1135-1144.

32. DeGrave AJ, Janizek JD, Lee S-I. AI for radiographic COVID-19 detection selects shortcuts over signal. *Nat Mach Intell* 2021;3:610-619.

33. Wolpert DH. The lack of a priori distinctions between learning algorithms. *Neural Comput* 1996;8:1341-1390.

Subphenotyping of Patients With Aortic Stenosis by Unsupervised Agglomerative Clustering of Echocardiographic and Hemodynamic Data

Mark Lachmann, MD,[a],* Elena Rippen,[a],* Tibor Schuster, PhD,[b] Erion Xhepa, MD, PhD,[c,d] Moritz von Scheidt, MD,[c,d] Costanza Pellegrini, MD,[c] Teresa Trenkwalder, MD,[c] Tobias Rheude, MD,[c] Anja Stundl, MD,[a] Ruth Thalmann, MD,[a] Gerhard Harmsen, PhD,[e] Shinsuke Yuasa, MD, PhD,[f] Heribert Schunkert, MD,[c,d] Adnan Kastrati, MD,[c,d] Karl-Ludwig Laugwitz, MD,[a,d] Christian Kupatt, MD,[a,d] Michael Joner, MD[c,d]

ABSTRACT

OBJECTIVES The aim of this retrospective analysis was to categorize patients with severe aortic stenosis (AS) according to clinical presentation by applying unsupervised machine learning.

BACKGROUND Patients with severe AS present with heterogeneous clinical phenotypes, depending on disease progression and comorbidities.

METHODS Unsupervised agglomerative clustering was applied to preprocedural data from echocardiography and right heart catheterization from 366 consecutively enrolled patients undergoing transcatheter aortic valve replacement for severe AS.

RESULTS Cluster analysis revealed 4 distinct phenotypes. Patients in cluster 1 (n ¼ 164 [44.8%]), serving as a reference, presented with regular cardiac function and without pulmonary hypertension (PH). Accordingly, estimated 2-year survival was 90.6% (95% CI: 85.8%-95.6%). Clusters 2 (n ¼ 66 [18.0%]) and 4 (n ¼ 91 [24.9%]) both comprised patients with postcapillary PH. Yet patients in cluster 2 with preserved left and right ventricular structure and function showed a similar survival as those in cluster 1 (2-year survival 85.8%; 95% CI: 76.9%-95.6%), whereas patients in cluster 4 with dilatation of all heart chambers and a high prevalence of mitral and tricuspid regurgitation (12.5% and 14.8%, respectively) died more often (2-year survival 74.9% [95% CI: 65.9%-85.2%]; HR for 2-year mortality: 2.8 [95% CI: 1.4-5.5]). Patients in cluster 3, the smallest (n ¼ 45 [12.3%]), displayed the most extensive disease characteristics (ie, left and right heart dysfunction together with combined pre- and postcapillary PH), and 2-year survival was accordingly reduced (77.3% [95% CI: 65.2%-91.6%]; HR for 2-year mortality: 2.6 [95% CI: 1.1-6.2]).

CONCLUSIONS Unsupervised machine learning aids in capturing complex clinical presentations as observed in patients with severe AS. Importantly, structural alterations in left and right heart morphology, possibly due to genetic predisposition, constitute an equally sensitive indicator of poor prognosis compared with high-grade PH.
(J Am Coll Cardiol Intv 2021;14:2127–2140) © 2021 by the American College of Cardiology Foundation.

Calcific aortic stenosis (AS) represents one of the most frequent cardiovascular diseases after coronary artery disease and systemic arterial hypertension in developed countries, and the prevalence of severe AS among the elderly population (aged $75 years) is estimated to be 3.4% (1). As a consequence of backward transmission of left-sided filling pressures caused by severe AS, about 50% of patients present with pulmonary hypertension (PH) (2). Furthermore, as right heart performance is closely coupled with pulmonary circulation, initial compensatory right ventricular remodeling processes will eventually fail and lead to right heart dilatation with impairment of right ventricular function and progressing tricuspid regurgitation (3,4).

Patients with severe AS are traditionally categorized into subgroups according to the presence of symptoms and aortic valve gradient (AVG), flow, and left ventricular ejection fraction (LVEF) (5). This conventional categorization of patients represents a form of hypothesis-driven and divisive clustering, incorporating only a limited set of characteristics. Moreover, in a real-world scenario, physicians commonly encounter a disparity between AS-induced hemodynamic burden and extravalvular damage to the heart and pulmonary circulation, which might illustrate the consequences both of chronicity of AS and of comorbidities such as atrial fibrillation, coronary artery disease and chronic obstructive pulmonary disease. As unsupervised agglomerative clustering works on a theoretically endless set of variables and without the constraint of any a priori assumption, it holds the promise to improve identification and segregation of patients with similar characteristics. Unsupervised clustering approaches have proved extremely successful in analyzing high-throughput experimental data, for example, in the field of genomics or cancer biology (6). Adopting machine learning to cardiovascular medicine could aid in advancing the field of personalized medicine, especially when it concerns subphenotyping of heterogeneous diseases (7-10).

We therefore hypothesized that unsupervised agglomerative clustering of a complementary set of echocardiographic and hemodynamic parameters from right heart catheterization (RHC) could capture the complexity of clinical presentation in patients with severe AS by revealing distinct phenotypes, which would subsequently differ in mortality depending on the extent of left ventricular dysfunction, PH, and right heart failure.

METHODS

PATIENT RECRUITMENT. This was a retrospective cohort study drawing on prospectively and systematically collected echocardiographic and hemodynamic data from patients with severe AS. Enrolled patients underwent transcatheter aortic valve replacement (TAVR) for severe AS at 2 tertiary centers in Munich, Germany, between January 2014 and December 2020. Patients were included in the registry only after written informed consent was received. As the aim of this study was to analyze the extent of extravalvular damage subsequent or parallel to severe AS in depth, only patients with both echocardiography and RHC, obtained preprocedurally prior to TAVR, were included in this study. Baseline demographic and clinical characteristics were obtained from registry data or clinical records as appropriate. The primary outcome measure was all-cause mortality within 2 years after TAVR. Survival data were obtained from the German Civil Registry in case of patients being registered in Germany (96.7%), or from general practitioners, hospitals, and practice cardiologists for patients from foreign countries. Planned and conducted in conformity with the Declaration of Helsinki, this study was approved by the local ethics committee.

TRANSTHORACIC ECHOCARDIOGRAPHY. All echocardiographic studies were performed by experienced institutional cardiologists during clinical routine using a commercially available echocardiographic system equipped with a 2.5-MHz multifrequency phased-array transducer.

INVASIVE ASSESSMENT OF HEMODYNAMIC STATUS. A 7-F Swan-Ganz catheter was routinely used for preprocedural RHC via femoral access. Systolic and diastolic pulmonary artery pressures were directly recorded. Mean pulmonary artery pressure (mPAP) was calculated as: diastolic pulmonary artery pressure + 1/3 × (systolic pulmonary artery pressure — diastolic pulmonary artery pressure). Mean postcapillary wedge pressure was calculated over the entire cardiac cycle. Cardiac output was determined using the indirect Fick method. Pulmonary vascular resistance (PVR) was defined as: (mPAP — mean postcapillary wedge pressure)/cardiac output.

STATISTICAL ANALYSIS AND VARIABLE SELECTION CRITERIA. All statistical analyses were performed
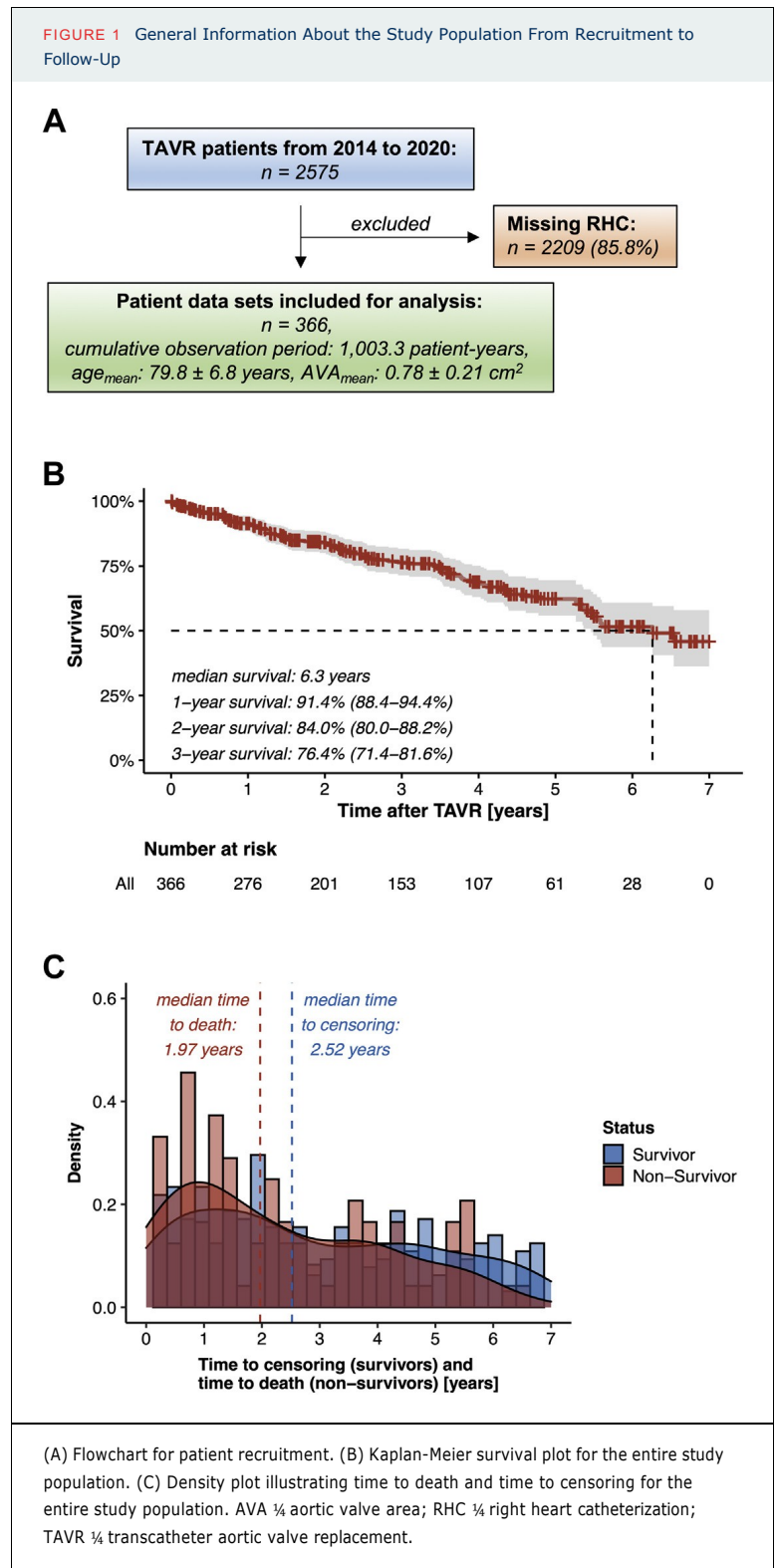
using R version 3.6.3 (R Foundation for Statistical Computing) (see Supplemental Table 1 for a complete list of R packages used). All computations were per- formed on a MacBook Pro (macOS Catalina version 10.15.5, Apple Computer) with a 2.3-GHz quad-core Intel Core i7 processor. The complete R code is available from the corresponding author upon reasonable request.

Categorical variables are presented as numbers and/or frequencies and continuous variables as mean ! SD. The chi-square or Fisher exact test was used to evaluate associations between categorical variables, and the Kruskal-Wallis test in combination with the pairwise Wilcoxon test with correction for multiple testing (Benjamini-Hochberg method) was used for comparisons of continuous variables, as appropriate. Pairwise comparisons of clusters as determined by unsupervised agglomerative clustering or predicted by an artificial neural network (ANN) were also calcu- lated using an independent-samples Wilcoxon test.

For analysis of collinearity, Pearson correlation coefficients were calculated, and the correlation ma- trix was hereinafter visualized (R package corrplot). The following criteria were defined for the selection of variables to be used for hierarchical clustering: 1) routine measurements in preprocedural work flow of echocardiography and RHC before TAVR; 2) good quality of representation on the first 5 dimensions of a principal component analysis, defined by squared cosine; and 3) no collinearity with other variables, as determined by the Pearson correlation coefficient.

After variable selection, a distance matrix was calculated using the Euclidean distance metric, the ideal number of clusters was determined using the elbow method, and agglomerative hierarchical clus- tering was finally performed, applying Ward's mini- mum variance method (R packages FactoMineR, factoextra, and NbClust). It is important to note that missing values were imputed using a random forest algorithm before calculating the distance matrix for the principal component analysis and hierarchical clustering (R package missForest) but were not used thereafter (eg, for cluster comparisons). A heatmap including dendrograms was generated for visualization of clustering results (R package ComplexHeatmap).

Survival was illustrated using the Kaplan-Meier method, and a Cox proportional hazards model was used to estimate HRs between identified clusters (R packages survival, survminer, and ggforest). A further univariate Cox proportional hazards model was used to shed light on additional contributing factors to 2-year mortality. For parameter estimates of interest, 95% CIs were reported.

FIGURE 1 General Information About the Study Population From Recruitment to Follow-Up

A

TAVR patients from 2014 to 2020: n = 2575

excluded → Missing RHC: n = 2209 (85.8%)

Patient data sets included for analysis: n = 366, cumulative observation period: 1,003.3 patient-years, age_{mean}: 79.8 ± 6.8 years, AVA_{mean}: 0.78 ± 0.21 cm^2

B

median survival: 6.3 years
1–year survival: 91.4% (88.4–94.4%)
2–year survival: 84.0% (80.0–88.2%)
3–year survival: 76.4% (71.4–81.6%)

Number at risk

All	366	276	201	153	107	61	28	0

C

median time to death: 1.97 years

median time to censoring: 2.52 years

Status
Survivor
Non–Survivor

(A) Flowchart for patient recruitment. (B) Kaplan-Meier survival plot for the entire study population. (C) Density plot illustrating time to death and time to censoring for the entire study population. AVA ¼ aortic valve area; RHC ¼ right heart catheterization; TAVR ¼ transcatheter aortic valve replacement.
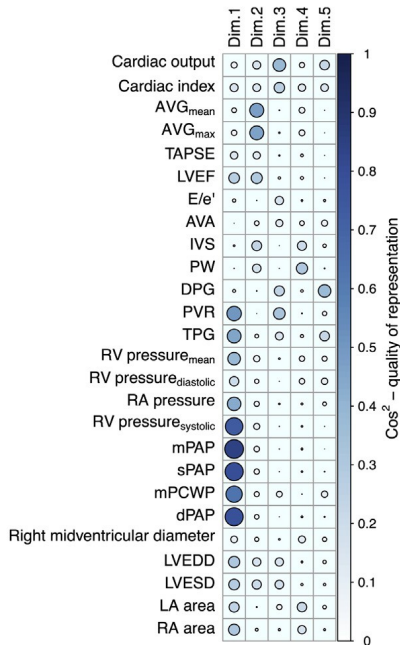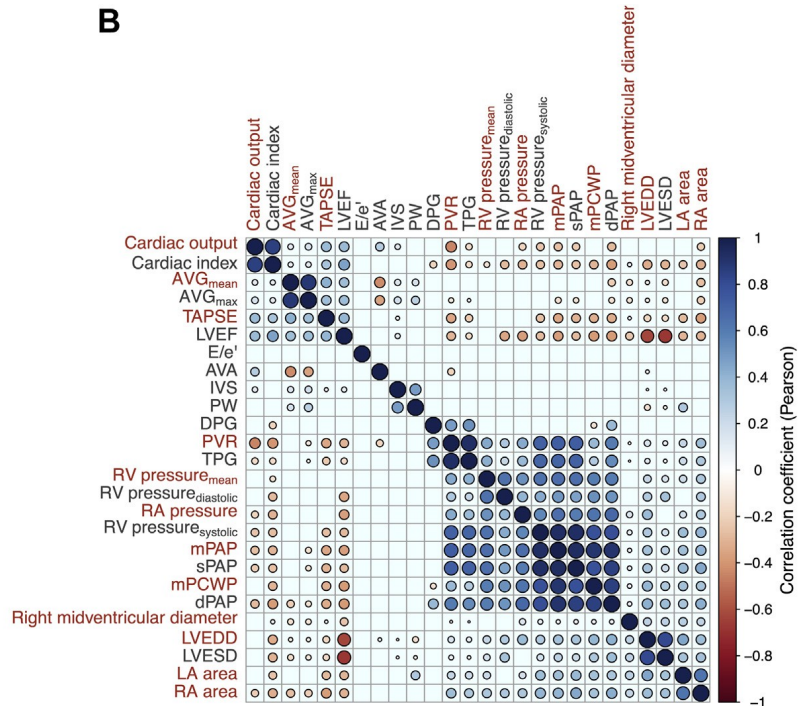
An ANN for cluster prediction was programmed using Keras for R with TensorFlow as a backend en- gine (R packages keras and tensorflow). The ideal ANN architecture (ie, the number of hidden layers

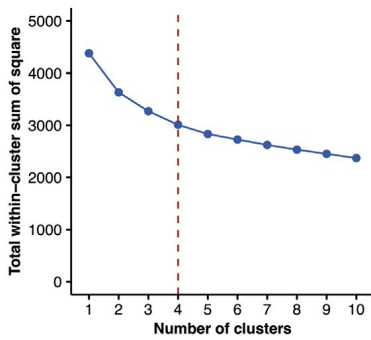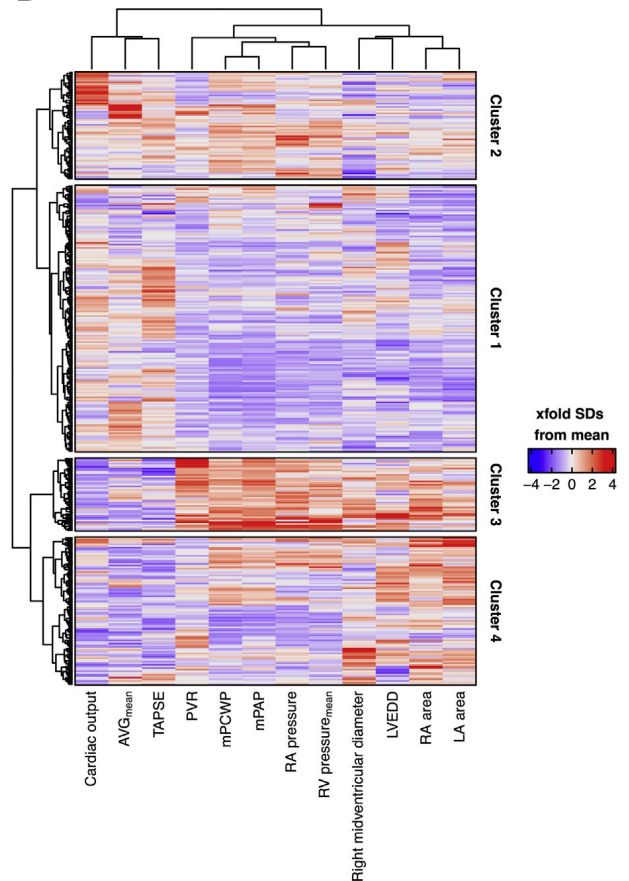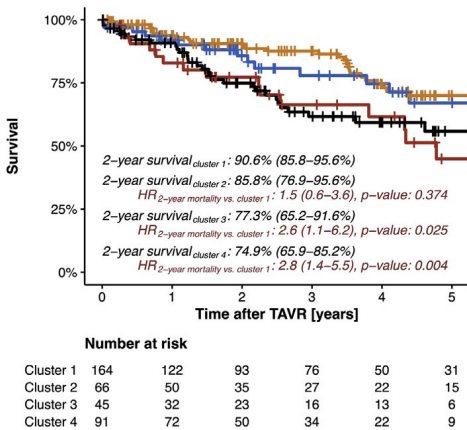**FIGURE 2** Variable Selection and Resulting Cluster Analysis

TABLE 1 Demographic and Clinical Characteristics in Accordance With Cluster Assignment

	Cluster 1 (n = 164)	Cluster 2 (n = 66)	Cluster 3 (n = 45)	Cluster 4 (n = 91)	*P* Value
Age (y)	79.6 ± 6.0 (4)	78.1 ± 6.9 (4)	79.8 ± 8.7	81.3 ± 6.7 (1, 2)	0.0068
Female	43.9	42.4	42.2	29.7	0.1483
BMI (kg/m^2)	26.3 ± 4.4	27.6 ± 4.0	26.7 ± 4.4	27.1 ± 4.8	0.1539
Arterial hypertension	86.5	93.9	100.0 (1)	94.5	0.0095
Smoking	19.0	28.8	28.9	31.9	0.1625
Diabetes mellitus	25.2	24.2	24.4	36.3	0.2127
NYHA functional class	2.48 ± 0.71 (3, 4)	2.52 ± 0.66 (3, 4)	3.02 ± 0.72 (1, 2)	2.87 ± 0.78 (1, 2)	5.4×10^{-7}
EuroSCORE (%)	13.7 ± 10.1 (3, 4)	13.9 ± 11.2 (3, 4)	29.1 ± 20.9 (1, 2)	22.6 ± 14.6 (1, 2)	8.6×10^{-10}
eGFR (mL/min)	63.0 ± 20.5 (3, 4)	64.6 ± 21.6 (3)	52.8 ± 21.1 (1, 2)	56.0 ± 19.9 (1)	0.0046
CAD	84.0	87.9	80.0	87.9	0.5635
COPD	8.6 (3)	15.2	24.4 (1)	15.4	0.0376
Atrial fibrillation and/or flutter	20.1 (2, 3, 4)	37.9 (1, 3, 4)	75.6 (1, 2)	79.1 (1, 2)	0.0005

Values are mean ± SD or %. The chi-square or Fisher exact test was used to evaluate associations between categorical variables, and the Kruskal-Wallis test in combination with the pairwise Wilcoxon test with correction for multiple testing (Benjamini-Hochberg method) was used for comparisons of continuous variables, as appropriate. Numbers in parentheses indicate between which clusters significant differences (*P* < 0.05) were detected. Pairwise comparison could not detect significant differences among clusters for the prevalence of arterial hypertension.

BMI = body mass index; CAD = coronary artery disease; COPD = chronic obstructive pulmonary disease; eGFR = estimated glomerular filtration rate; EuroSCORE = European System for Cardiac Operative Risk Evaluation; NYHA = New York Heart Association.

and the respective number of nodes per layer) was found by plain trial and error. As the ANN had to solve a multiclass classification task (ie, to assign each pa- tient to exactly 1 cluster), the final output layer con- sisted of 1 node per cluster using the softmax activation function. Accuracy and loss were chosen as metrics to observe the model's performance during training and testing. To prevent overfitting, a callback function was programmed to automatically interrupt training when the loss in the test set did not further decrease after 20 epochs. A *P* value <0.05 was considered to indicate statistical significance.

RESULTS

THREE HUNDRED SIXTY-SIX PATIENTS WITH RHC BEFORE TAVR FOR SEVERE AS CONSTITUTED THE STUDY POPULATION. In total, the registry contained 2,575 patients undergoing TAVR for severe AS be- tween 2014 and 2020. Among them, 2,209 patients were excluded because of missing preprocedural RHC, so that the detailed analysis was performed on a dataset consisting of 366 patients. The mean age of the study population was 79.8 ± 6.8 years, the mean aortic valve area was 0.78 ± 0.21 cm², and median survival was 6.3 years (Figures 1A and 1B). Fifty percent of deaths occurred within 1.97 years after TAVR (Figure 1C).

TWELVE VARIABLES FROM ECHOCARDIOGRAPHY AND RHC WERE SELECTED FOR CLUSTERING. Initially, 26 routinely measured candidate variables from echocardiography and RHC were assessed regarding their significance in explaining the variance among patients. The first 5 dimensions of a principal component analysis explained 62.9% of the variance within the dataset. Variables with low quality of representation on those first 5 dimensions, such as E/ e⁰ ratio and posterior wall thickness, were regarded as not meaningful enough to be included in cluster analysis. Notably, aortic valve area also had no

FIGURE 2 Continued

(A) Illustration of candidate variables from echocardiography and right heart catheterization (RHC) with regard to their quality of representation (expressed as squared cosine) on the first 5 dimensions of a principal component analysis. (B) Correlation matrix of variables from echocardiography and RHC. Color intensity and circle size are proportional to the correlation coefficients (if *P* < 0.05), while correlation coefficients with *P* values > 0.05 are left blank. Variables selected for clustering are shown in red, while variables with no implementation into the clustering process are shown in gray. (C) Illustration of elbow method to determine the ideal number of clusters. (D) Heatmap and dendrogram for visualization of clustering results. Data are given as color-coded SDs from mean after centering and scaling. (E) Kaplan-Meier survival plot in accordance with cluster assignment. AVA = aortic valve area; AVG = aortic valve gradient; dPAP = diastolic pulmonary artery pressure; DPG = diastolic pressure gradient; IVS = interventricular septal thickness; LA = left atrial; LVEDD = left ventricular end-diastolic diameter; LVEF = left ventricular ejection fraction; LVESD = left ventricular end-systolic diameter; mPAP = mean pulmonary artery pressure; mPCWP = mean postcapillary wedge pressure; PVR = pulmonary vascular resistance; PW = posterior wall thickness; RA = right atrial; RV = right ventricular; sPAP = systolic pulmonary artery pressure; TAPSE = tricuspid annular plane systolic excursion; TPG = transpulmonary pressure gradient.

TABLE 2 Comparison of Echocardiographic and Hemodynamic Data From RHC in Accordance With Cluster Assignment

	Cluster 1 (n ¼ 164)	Cluster 2 (n ¼ 66)	Cluster 3 (n ¼ 45)	Cluster 4 (n ¼ 91)	P Value
AVA (cm^2)	0.79 ! 0.19	0.78 ! 0.23	0.74 ! 0.24	0.79 ! 0.20	0.6768
AVG$_{mean}$ (mm Hg)	41.6 ! 12.7 (2, 3, 4)	50.3 ! 18.0 (1, 3, 4)	33.5 ! 13.7 (1, 2)	31.9 ! 14.7 (1, 2)	7.5 × 10^{-13}
AVG$_{max}$ (mm Hg)	64.3 ! 21.4 (2, 3, 4)	75.4 ! 24.0 (1, 3, 4)	50.0 ! 21.1 (1, 2)	49.8 ! 21.0 (1, 2)	3.6 × 10^{-12}
Cardiac output (L/min)	5.12 ! 0.88 (2, 3, 4)	5.79 ! 1.55 (1, 3, 4)	3.82 ! 0.87 (1, 2, 4)	4.47 ! 1.12 (1, 2, 3)	<2 × 10^{-16}
LVEF (%)	57.2 ! 6.4 (2, 3, 4)	55.9 ! 6.5 (1, 3, 4)	42.4 ! 15.7 (1, 2)	47.3 ! 12.2 (1, 2)	4.2 × 10^{-16}
LVESD (mm)	30.3 ! 8.1 (3, 4)	32.7 ! 8.3	39.6 ! 14.7 (1)	36.5 ! 11.4 (1)	9.1 × 10^{-5}
LVEDD (mm)	43.9 ! 7.2 (2, 3, 4)	46.8 ! 5.7 (1, 3, 4)	52.7 ! 10.3 (1, 2)	49.5 ! 9.2 (1, 2)	6.8 × 10^{-9}
IVS (mm)	13.6 ! 2.4	14.4 ! 2.6 (3)	12.8 ! 2.4 (2)	13.6 ! 2.8	0.0117
PW (mm)	11.9 ! 2.3	12.3 ! 2.4	12.1 ! 3.4	12.5 ! 2.6	0.3775
RV-RA gradient (mm Hg)	30.3 ! 11.5 (2, 3, 4)	36.6 ! 14.0 (2, 3)	48.7 ! 14.0 (1, 2, 4)	37.5 ! 14.4 (1, 3)	3.0 × 10^{-12}
sPAP (mm Hg)	34.3 ! 10.1 (2, 3, 4)	51.6 ! 11.2 (1, 3, 4)	69.8 ! 14.0 (1, 2, 4)	44.1 ! 13.2 (1, 2, 3)	<2 × 10^{-16}
mPAP (mm Hg)	21.2 ! 6.5 (2, 3, 4)	34.2 ! 7.8 (1, 3, 4)	46.9 ! 8.5 (1, 2, 4)	27.5 ! 9.2 (1, 2, 3)	<2 × 10^{-16}
mPCWP (mm Hg)	12.7 ! 5.3 (2, 3, 4)	22.2 ! 6.2 (1, 3, 4)	28.5 ! 6.6 (1, 2, 4)	17.1 ! 8.6 (1, 2, 3)	<2 × 10^{-16}
PVR (Wood units)	1.68 ! 0.84 (2, 3, 4)	2.26 ! 1.16 (1, 3)	4.96 ! 1.95 (1, 2, 4)	2.43 ! 1.12 (1, 3)	<2 × 10^{-16}
Mean RV pressure (mm Hg)	7.5 ! 4.2 (2, 3)	12.2 ! 4.1 (1, 3, 4)	16.1 ! 7.5 (1, 2, 4)	8.9 ! 5.5 (2, 3)	<2 × 10^{-16}
RA pressure (mm Hg)	5.4 ! 2.8 (2, 3, 4)	9.6 ! 4.8 (1, 3, 4)	14.2 ! 4.8 (1, 2, 4)	7.0 ! 4.6 (1, 2, 3)	<2 × 10^{-16}
TAPSE (mm)	21.5 ! 4.7 (3, 4)	21.3 ! 3.5 (3, 4)	16.1 ! 4.6 (1, 2)	16.8 ! 4.5 (1, 2)	<2 × 10^{-16}
Right midventricular diameter (mm)	28.0 ! 5.1 (2, 3, 4)	26.0 ! 6.4 (1, 3, 4)	34.1 ! 6.3 (1, 2, 4)	32.1 ! 7.1 (1, 2, 3)	7.8 × 10^{-11}
LA area (cm^2)	21.5 ! 5.2 (2, 3, 4)	26.3 ! 6.2 (1, 3, 4)	30.4 ! 6.9 (1, 2)	32.9 ! 9.0 (1, 2)	<2 × 10^{-16}
RA area (cm^2)	15.8 ! 3.8 (2, 3, 4)	19.6 ! 4.6 (1, 3, 4)	27.8 ! 8.1 (1, 2)	26.1 ! 7.2 (1, 2)	<2 × 10^{-16}
Mitral regurgitation grade $ III/IV	5.0 (3, 4)	1.6 (3, 4)	31.1 (1, 2, 4)	12.5 (1, 2, 3)	2.1 × 10^{-6}
Tricuspid regurgitation grade $III/IV	2.5 (3, 4)	3.2 (3, 4)	31.1 (1, 2, 4)	14.8 (1, 2, 3)	5.1 × 10^{-8}

Values are mean ! SD or %. The Kruskal-Wallis test in combination with the pairwise Wilcoxon test with correction for multiple testing (Benjamini-Hochberg method) was used for comparisons among clusters. Numbers in parentheses indicate between which clusters significant differences (*P* # 0.05) were detected. Variables from echocardiography: LVEF, LVESD, LVEDD, IVS, PW, RV-RA gradient, TAPSE, LA area, and RA area. Variables from right heart catheterization: sPAP, mPAP, mPCWP, PVR, mean RV pressure, and RA pressure. Measurements of AVA, AVG$_{mean,}$ and AVG$_{max}$ were preferentially taken from right heart catheterization, if available.

AVA ¼ aortic valve area; AVG ¼ aortic valve gradient; IVS ¼ interventricular septal thickness; LA ¼ left atrial; LVEDD ¼ left ventricular end-diastolic diameter; LVEF ¼ left ventricular ejection fraction; LVESD ¼ left ventricular end-systolic diameter; RV ¼ right ventricular; mPAP ¼ mean pulmonary artery pressure; mPCWP ¼ mean postcapillary wedge pressure; PVR ¼ pulmonary vascular resistance; PW ¼ posterior wall thickness; RA ¼ right atrial; sPAP ¼ systolic pulmonary artery pressure; TAPSE ¼ tricuspid annular plane systolic excursion.

relevant impact on explaining the observed variance among patients and did therefore not contribute as a variable for clustering (Figure 2A). For cases in which variables measure similar characteristics and there-fore show close correlation, such as PVR and trans- pulmonary pressure gradient or maximum and mean AVG, only 1 variable was implemented in the cluster analysis. In total, the 26 candidate variables were reduced to 12 final variables for clustering, which covered all stages of potential disease progression from AVG over left heart function and pulmonary circulation to right heart structural and functional parameters (Figure 2B). In total, 5.2% of 4,392 data points had missing values for those 12 variables (Supplemental Figure 1A), and the largest proportion of missing values was found for measurements of right ventricular mean pressure (11.7% of values missing) (Supplemental Figure 1B). After imputing missing values, initially observed and later imputed values for right ventricular mean pressure displayed a
similar distribution (9.8 ! 5.8 mm Hg vs 9.2 ! 3.5 mm Hg) (Supplemental Figures 1C and 1D). The

potential number of clusters was further determined using the elbow method, which demonstrated that 4 clusters would be ideal (Figure 2C). Subsequently, patients were categorized into 4 clusters using un- supervised agglomerative clustering (Figure 2D).

FOUR DISTINCT CLINICAL PHENOTYPES COULD BE DISTINGUISHED, REFLECTING VARIOUS EXTENTS OF DISEASE SEVERITY AND HENCE DIFFERING IN MORTALITY. Cluster 1 (n ¼ 164 [44.8%]), constituting the majority of patients and serving as a reference cluster, comprised patients who presented with an intermediate symptomatic burden expressed as a mean New York Heart Association functional class of 2.48 ! 0.71 (Table 1). Patients in this cluster demonstrated preserved cardiac function (cardiac output 5.12 ! 0.88 L/min, LVEF 57.2% ! 6.4%, tricuspid annular plane systolic excursion [TAPSE] 21.5 ! 4.7 mm) and a high gradient across the aortic valve (AVG$_{max}$ 64.3 ! 21.4 mm Hg) (Table 2). Furthermore, PH was absent in cluster 1 (mPAP 21.2 ! 6.5 mm Hg). In agreement with the regular size of the left and right atria in patients in cluster 1, the prevalence of atrial

fibrillation and/or flutter was lowest among all clusters (20.1%). Taking the intermediate risk of comorbidities into account (European System for Cardiac Operative Risk Evaluation [EuroSCORE] 13.7% ! 10.1%), Kaplan-Meier analysis showed an excellent 2- year survival rate of 90.6% (95% CI: 85.8%-95.6%) for patients in cluster 1 (Figure 2E).

In contrast, patients in cluster 2 (n ¼ 66 [18.0%]) were characterized by postcapillary PH, revealing elevations in mPAP (34.2 ! 7.8 mm Hg) and mean postcapillary wedge pressure (22.2 ! 6.2 mm Hg), indicating disease progression beyond the compen- satory capacity of the left ventricle (Table 2). Notably, the integrity of left and right ventricular systolic function was maintained (LVEF 55.9% ! 6.5%, TAPSE 21.3 ! 3.5 mm), and the prevalence of severe mitral and tricuspid regurgitation (1.6% and 3.2%, respectively) was comparable with that in cluster 1. Given the isolated postcapillary nature of PH in cluster 2, which is considered reversible upon resolving the underlying AS, patients in cluster 2 had a 2-year sur- vival rate not statistically different from that of pa- tients in cluster 1 (2-year survival 85.8% [95% CI: 76.9%-95.6%]; HR for 2-year mortality: 1.5 [95% CI: 0.6-3.6]) (Figure 2E).

The most extensive disease characteristics subse- quent to severe AS, in contrast, were found in pa- tients in cluster 3 (n ¼ 45 [12.3%]), who presented with a mean New York Heart Association functional class of 3.02 ! 0.72. Patients in this cluster presented with the most severely impaired cardiac function, expressed as cardiac output of merely 3.82 ! 0.87 L/min (LVEF 42.4% ! 15.7%, TAPSE 16.1 ! 4.6 mm). At the same time, patients in cluster 3 presented with the highest mPAP (46.9 ! 8.5 mm Hg) and the highest PVR (4.96 ! 1.95 Wood units). Together with elevated mean postcapillary wedge pressure (28.5 ! 6.6 mm Hg), these parameters indicate combined pre- and postcapillary PH (Table 2). Concomitant with left and right heart enlargement, the prevalence of severe mitral and tricuspid regurgitation was dramatically increased (31.1% and 31.1%, respectively). Compared with cluster 1, 2-year survival in patients in cluster 3 was significantly decreased (77.3% [95% CI: 65.2%-91.6%]; HR for 2-year mortality: 2.6 [95% CI: 1.1-6.2]) (Figure 2E).

Finally, patients in cluster 4 (n ¼ 91 [24.9%]) dis- played dilatation of all cardiac chambers and impairment of left and right ventricular function (LVEF 47.3% ! 12.2%, TAPSE 16.8 ! 4.5 mm) (Table 2). Concomitant with the structural changes of the left and right heart, the prevalence of severe mitral and tricuspid regurgitation was increased (12.5% and 14.8%, respectively), although less impressively than

| TABLE 3 Univariate Cox Regression Analysis Using Death Within 2 Years After TAVR as a Dependent Variable		
	HR (95% CI)	*P* Value
Age	0.99 (0.95-1) per year	0.78
Female	0.74 (0.41-1.3)	0.32
BMI	0.95 (0.89-1) per kg/m^2	0.1
Arterial hypertension	1.3 (0.41-4.2)	0.65
Smoking	1.5 (0.91-2.6)	0.11
Diabetes mellitus	1.3 (0.69-2.3)	0.45
NYHA functional class	1.9 (1.3-3) per class	0.0019
EuroSCORE	1 (1-1) per %	4.3×10^{-6}
eGFR	0.98 (0.97-0.99) per mL/min	0.0063
CAD	1.1 (0.48-2.7)	0.78
COPD	2.3 (1.2-4.3)	0.01
Atrial fibrillation and/or flutter	1.7 (1-2.8)	0.048
AVA	1.3 (0.34-4.7) per cm^2	0.74
AVG$_{mean}$	0.97 (0.94-0.99) per mm Hg	0.0011
AVG$_{max}$	0.98 (0.97-0.99) per mm Hg	0.0039
Cardiac output	0.73 (0.57-0.93) per L/min	0.013
LVEF	0.96 (0.94-0.98) per %	5.9×10^{-5}
LVESD	1 (1-1.1) per mm	0.0007
LVEDD	1 (0.99-1.1) per mm	0.11
IVS	0.88 (0.78-0.99) per mm	0.036
PW	0.96 (0.86-1.1) per mm	0.53
mPAP	1 (1-1) per mm Hg	0.035
mPCWP	1 (0.99-1.1) per mm Hg	0.23
PVR	1.3 (1.1-1.5) per Wood unit	0.0006
Mean RV pressure	1 (0.97-1.1) per mm Hg	0.44
RA pressure	1 (0.98-1.1) per mm Hg	0.17
TAPSE	0.93 (0.87-0.98) per mm	0.0099
Right midventricular diameter	1.1 (1-1.1) per mm	0.014
LA area	1 (0.99-1.1) per cm^2	0.2
RA area	1 (1-1.1) per cm^2	0.0061
Mitral regurgitation grade $III/ IV	2.4 (1.2-5)	0.016
Tricuspid regurgitation grade $III/ IV	0.86 (0.31-2.4)	0.78

Abbreviations as in Tables 1 and 2.

in cluster 3. In the context of seemingly milder PH than in clusters 2 and 3, how would these severe structural changes translate into 2-year outcomes? The alterations in cardiac morphology and function are well reflected in survival: 2-year survival in pa- tients in cluster 4 was significantly reduced compared with those in cluster 1 (74.9% [95% CI: 65.9%-85.2%];

HR for 2-year mortality: 2.8 [95% CI: 1.4-5.5]) (Figure 2E).

UNIVARIATE COX REGRESSION ANALYSIS SUGGESTED ADDITIONAL NONCARDIAC CONTRIBUTORS TO MORTALITY AFTER TAVR. Various variables from demography, echocardiography, and RHC were stud- ied using univariate Cox regression analysis to

FIGURE 3 Use of an ANN for Cluster Assignment
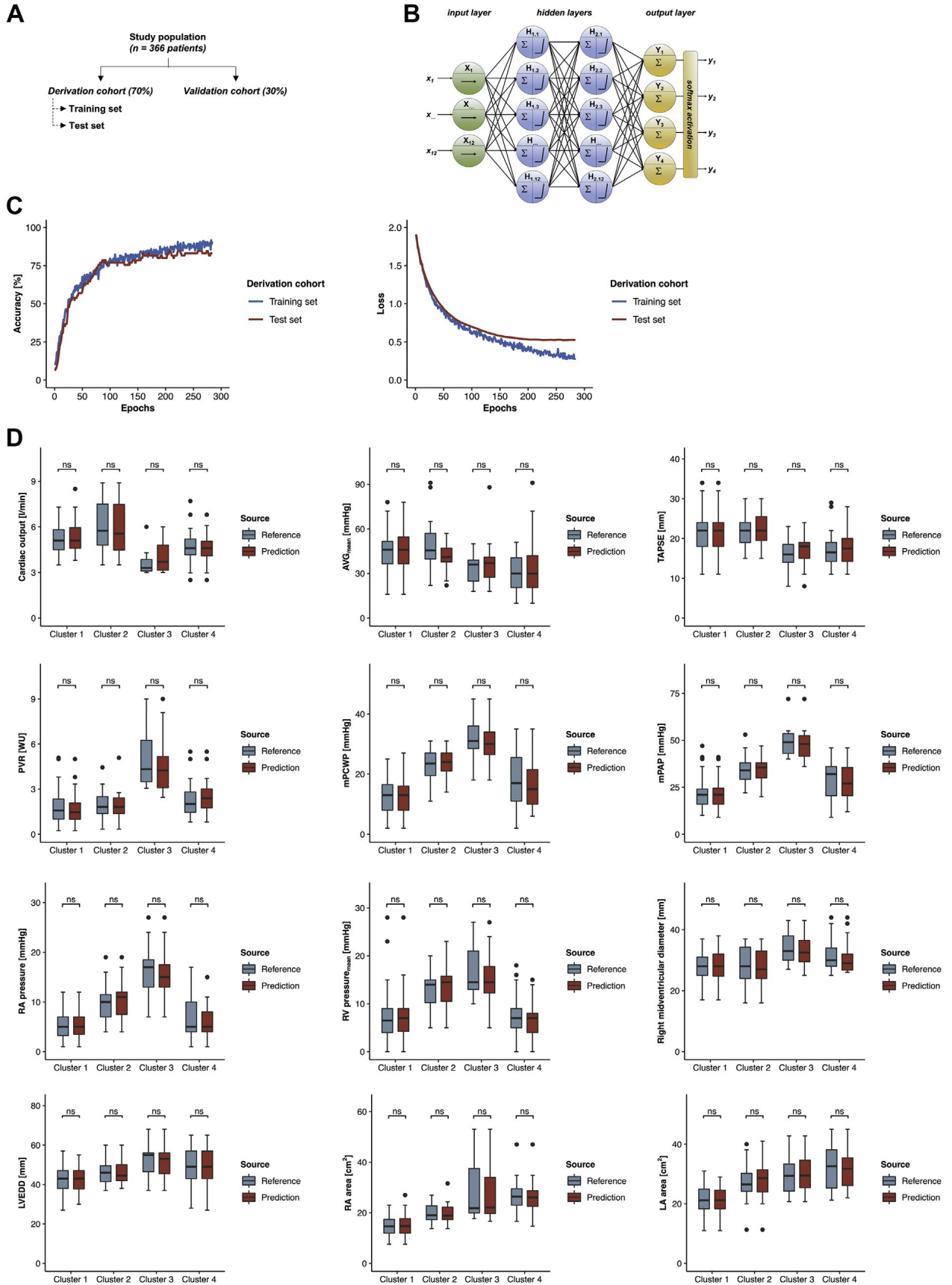
identify additional contributing factors for 2-year mortality (Table 3). Among well-established noncardiac predictors such as deteriorating estimated glomerular filtration rate and presence of chronic obstructive pulmonary disease, univariate Cox regression analysis also suggested a linear association between AVG_{mean} and mortality after TAVR.

AN ANN COULD ACCURATELY ASSIGN PATIENTS TO PREDEFINED CLUSTERS.

To open an avenue to prospectively assign patients to the predefined clusters, the study population, containing a cluster assignment from unsupervised agglomerative clustering for each patient, was randomly divided into derivation and validation cohorts, and an ANN with 2 hidden layers was subsequently trained (see Figures 3A and 3B). After training for 284 epochs, when the loss in the test set had not further decreased for 20 epochs, the ANN reached accuracy of 90.6% in the training set and loss of 0.526 in the test set (Figure 3C). Applying the trained ANN to the validation cohort demonstrated accuracy of 83.5% (95% CI: 75.2%-89.9%), which significantly outperformed the no-information rate ($P = 2.3 \times 10^{-15}$). Patients in high-risk clusters 3 and 4 were detected with high sensitivity (100.0% and 85.2%, respectively) and specificity (95.9% and 95.1%, respectively) (Table 4). Moreover, comparison of clusters within the validation cohort as determined by unsupervised agglomerative clustering (reference) or predicted by the ANN (prediction) demonstrated a strong resemblance in terms of echocardiographic and hemodynamic characteristics (Supplemental Table 2, Figure 3D).

DISCUSSION

EXPANDING THE ANALYTICAL ARMAMENTARIUM WITH MACHINE LEARNING TECHNOLOGY OFFERS NOVEL INSIGHTS INTO CLINICAL PHENOTYPES PRESENTED BY PATIENTS WITH SEVERE AS.

Innovative algorithms and improved computing power enable machine learning to extract information from complex datasets at a scale that exceeds the capacity of conventional methods. Unsupervised agglomerative clustering holds the promise to pro- vide finer resolution by grouping observations on the basis of similarities instead of differences and hence potentially challenges the traditional perspective, especially in the field of medicine. To our knowledge, this is the first detailed study focusing on sub- phenotyping patients with severe AS according to their cardiopulmonary profile, as assessed by pre- procedural echocardiography and RHC, by applying unsupervised agglomerative clustering. Implement- ing 12 carefully chosen variables, which cover all stages of cardiac and pulmonary circulatory condi- tion, into the calculation for a distance matrix, basi- cally 4 distinct clinical phenotypes of patients with severe AS could be distinguished, reflecting various extents of disease severity and hence differing in mortality. Finally, our study supports the theory that AS should not be regarded as an isolated valvular disease, but also upstream damages such as left ventricular dysfunction, PH, and right heart failure must be taken into consideration. Distinct from pre- vious studies, we provide a comprehensive structural and hemodynamic evaluation based on unsupervised clustering. We hence aimed to refine diagnostic res- olution, as we neither classify patients in accordance with a hypothesis-driven staging system of sequential damages nor infer causality (eg, decreased left ven- tricular systolic function in cluster 3 could be caused by afterload mismatch due to severe AS or due to intrinsic myocardial impairment secondary to ischemic heart disease, myocardial fibrosis, or arte- rial hypertension).

PREPROCEDURAL RHC AS A PREREQUISITE FOR INCLUSION RESTRICTED THE STUDY POPULATION TO 366 PATIENTS (14.2%), YET COMPARISON WITH PUBLISHED RESEARCH ARGUES FOR ITS REPRESENTATIVE CHARACTER.

As a major limita- tion to the generalization of our findings, we acknowledge that only a minority of patients from the

---

FIGURE 3 Continued

(A) Study scheme to train and validate the artificial neural network (ANN). The derivation cohort was randomly divided into a training and test set in a 0.75:0.25 split ratio. (B) ANN architecture. The network consisted of 12 input nodes for each variable beforehand implemented in agglomerative hierarchical clustering. Furthermore, 2 fully connected hidden layers with 12 nodes each were constructed. Rectified linear units were chosen as activation functions to propagate the signal. The output layer consisted of 4 nodes, using the softmax activation function for final cluster prediction. (C) Learning curves with accuracy and loss as metrics for ANN performance. (D) Comparison of clusters (pairwise comparisons calculated using Wilcoxon test) within the validation cohort as determined by unsupervised agglomerative clustering (reference) or predicted by the ANN (prediction). The central line in each box plot denotes the median value, while the box contains all values ranging between the 25th and 75th percentiles of the dataset. The black whiskers mark the 5th and 95th percentiles, and values falling beyond these upper and lower bounds are considered outliers, plotted as black dots. Abbreviations as in Figure 2.

TABLE 4 Confusion Matrix of Cluster Assignments Within the Validation Cohort (n ¼ 109 [30% of the Study Population]) as Determined by Unsupervised Agglomerative Clustering (Reference) and as Predicted by the Artificial Neural Network (Prediction)

Prediction	Reference			
	Cluster 1 (n ¼ 51)	Cluster 2 (n ¼ 20)	Cluster 3 (n ¼ 11)	Cluster 4 (n ¼ 27)
Cluster 1 (n ¼ 51)	45	5	0	1
Cluster 2 (n ¼ 16)	3	12	0	1
Cluster 3 (n ¼ 15)	0	2	11	2
Cluster 4 (n ¼ 27)	3	1	0	23
Sensitivity (%)	88.2	60.0	100.0	85.2
Specificity (%)	89.7	95.5	95.9	95.1
Positive predictive value (%)	88.2	75.0	73.3	85.2
Negative predictive value (%)	89.7	91.4	100.0	95.1

Accuracy ¼ 83.5% (75.2%-89.9%); no information rate ¼ 46.8%; *P* (accuracy > no information rate) ¼ 2.3 × 10^{-15}.

joint registry (14.2%) met the inclusion criterion of preprocedural RHC. This is because contemporary guidelines from the European Society of Cardiology recommend restricting preprocedural RHC to situa- tions in which noninvasive evaluation of AS severity is inconclusive or discordant with clinical findings (11). Nonetheless, we focused on a cohort with pre- procedurally obtained RHC, as RHC represents the gold standard to assess the severity of hemodynamic impairment of the pulmonary circulation (12). To evaluate whether the study population is still repre- sentative, we compared classifications of PH with data from published research. Pooling 3 recently published datasets containing information about RHC from 2,336 patients with severe AS (2,13,14), we could observe similar proportions of patients with no and with isolated postcapillary PH, which together constitute the vast majority of cases (83.6% and 82.5% for the study population and for the reference from published research, respectively) (Supplemental Figure 2, Supplemental Table 3). As this study dem- onstrates, detailed assessment of hemodynamic sta- tus in pulmonary circulation and closely coupled right heart can refine diagnostic resolution and concomitant long-term risk stratification. At the same time, it could also prove beneficial for prognostic resolution prior to TAVR, if future ANNs can assign patients to predefined clusters solely according to echocardiographic parameters.

RELIABLE IMPUTATION OF MISSING VALUES USING A RANDOM FOREST ALGORITHM OPENS AN AVENUE FOR PROSPECTIVE STUDIES. Because calculating a distance matrix for subsequent cluster analysis re- quires complete observations, we initially applied a random forest algorithm to impute missing values.

Imputation by means of a random forest algorithm has been proved efficient in complex epidemiologic datasets (15), and the applied algorithm has been validated for a proportion of missing values of up to 30% (16). The imputation was based on the assump- tion that values are missing at random and can therefore be replaced by predictions based on the observed portion of variables (17). Observed and imputed values for the most frequently missing var- iable used for clustering (ie, right ventricular mean pressure) were compared, and no significant differ- ences were detected. The possibility of reliably imputing missing values could be a helpful approach, if software programs in future clinical practice should aim to prospectively assign patients to clusters that are characterized by the aforementioned 12 variables.

MACHINE LEARNING CAN ASSIST IN CLINICAL DECISION MAKING BY IDENTIFYING DIAMETRICALLY OPPOSED ENDS OF A CONTINUOUS SCALE OF DISEASE PROGRESSION WITH A DRAMATIC DIFFERENCE IN MORTALITY. After the onset of symptoms due to se- vere AS, mortality increases to more than 50% within 2 years, unless valve replacement is performed promptly (18). Clusters 1 and 3 represent diametri- cally opposed ends of a continuous scale of disease progression, with patients in cluster 3 being alarm- ingly dyspneic because of reduced cardiac output, massive PH, and severely impaired right heart func- tion. Interestingly, no differences in age could be found between clusters 1 and 3 (79.6 ! 6.0 years vs 79.8 ! 8.7 years). It therefore remains elusive whether patients in cluster 3 had more rapid pro- gression of disease, until they finally presented with New York Heart Association functional class > III or whether AS had evolved earlier but remained undi- agnosed and therefore untreated for a longer time. An increased HR of 2.6 (95% CI: 1.2-6.2) for death within 2 years after TAVR as observed in cluster 3 highlights the importance of early diagnosis and accurate timing of intervention.

RIGHT VENTRICULAR AND ATRIAL DILATATION WITH CONCOMITANT SEVERE TRICUSPID REGURGITATION DISTINGUISHES CLUSTERS 2 AND 4 WITH ISOLATED POSTCAPILLARY PH. With regard to the extent of PH subsequent to AS, clusters 2 and 4 both appeared as intermediate stages of disease progression at first sight but differed in mortality. Mortality in cluster 2 was similar to that in cluster 1, while mortality in cluster 4 was significantly increased and resided in a range similar to that in cluster 3, with the most extensive characteristics of cardiopulmonary failure. Possibly, dilatation of the right ventricle and atrium with subsequent high frequency of severe tricuspid

CENTRAL ILLUSTRATION Machine Learning Technology Aids in Capturing the Complexity of Clinical Presentations in Patients With Severe Aortic Stenosis
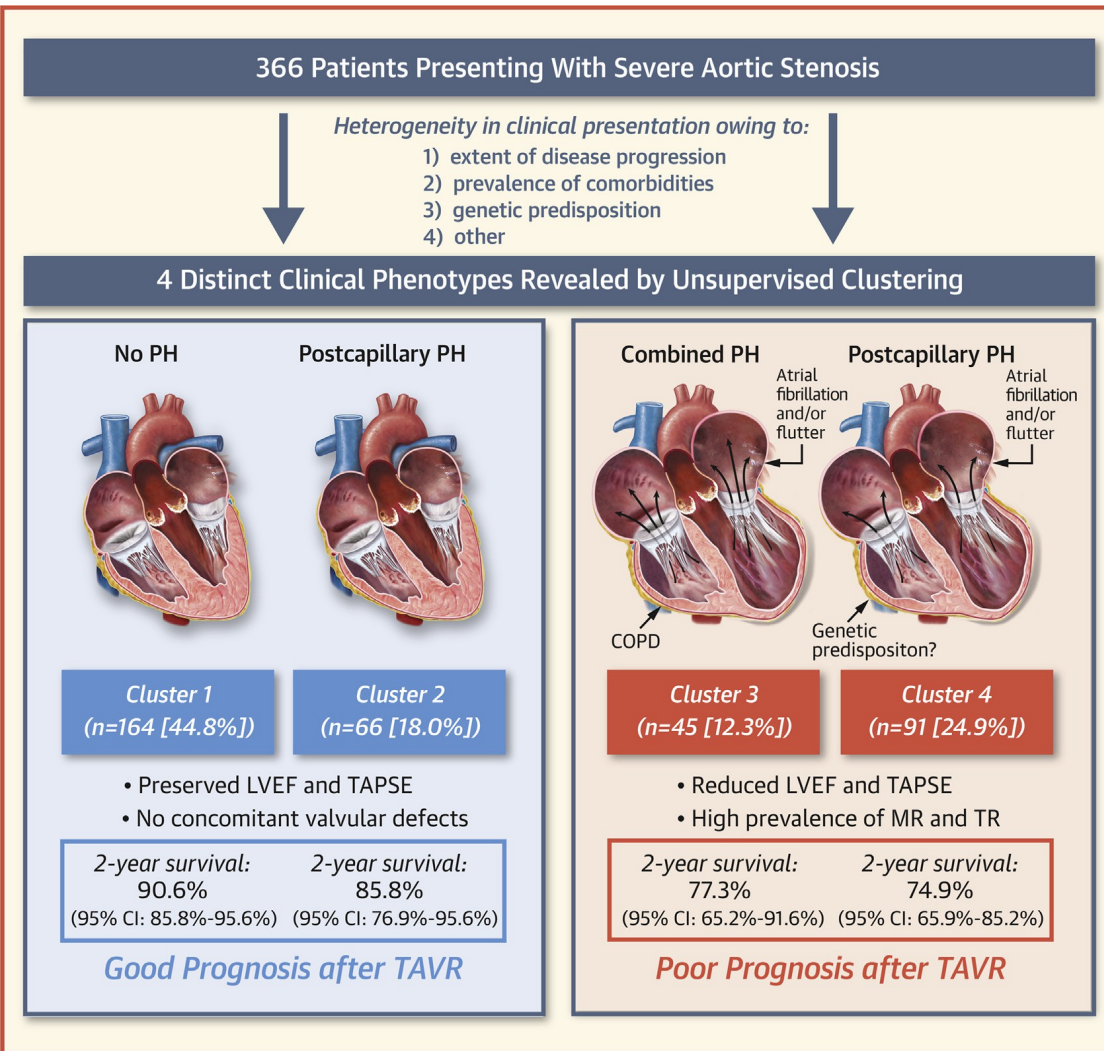
Lachmann, M. et al. J Am Coll Cardiol Intv. 2021;14(19):2127–2140.

Unsupervised clustering of patients with severe aortic stenosis (n ¼ 366) according to echocardiographic and hemodynamic characteristics reveals 4 distinct clinical phenotypes, which subsequently differ in survival after transcatheter aortic valve replacement. Both distinct extents of disease progression due to severe aortic stenosis and the aggravating impact of comorbidities such as atrial fibrillation and chronic obstructive pulmonary disease account for differences among clusters. Importantly, structural changes of the heart, possibly due to genetic predisposition, constitute an equally sensitive indicator of poor prognosis as compared with high-grade pulmonary hypertension. COPD ¼ chronic obstructive pulmonary disease; LVEF ¼ left ventricular ejection fraction; MR ¼ mitral regurgitation; PH ¼ pulmonary hypertension; TAPSE ¼ tricuspid annular plane systolic excursion; TAVR ¼ transcatheter aortic valve replacement; TR ¼ tricuspid regurgitation.

regurgitation in patients in cluster 4 ultimately resulted in death from right heart failure. We conclude that PH without right ventricular dilatation, as observed in cluster 2, does not impair outcomes at 2-year follow-up after TAVR, whereas cluster 4 with moderate and cluster 3 with severe PH form a group with distressing 2-year survival rates (Central

Illustration). Moreover, the remarkably similar sur- vival function among patients in clusters 3 and 4 accompanied by relevant differences in morphologic and hemodynamic right-sided parameters calls for future clinical investigation, where pre- and post- procedural management may have significant impact on segregating mid-term mortality among these

clusters. Our currently proposed machine learning approach may ultimately help identify these patients prior to TAVR. Functional tricuspid regurgitation ac- counts for approximately 80% of cases, and in pa- tients with long-standing AS, it reflects the consequence of advanced hemodynamic burden sec- ondary to left heart dysfunction, PH, and right ven- tricular and atrial dilatation (5). Because of its functional nature, it was classically believed that correction of the underlying AS would also improve tricuspid regurgitation. However, PH as a linking hub between left and right heart disease can persist in a substantial number of cases after TAVR, and persis- tence of PH translates into poor prognosis (19). Future echocardiographic follow-up studies after TAVR will shed light on the questionable reversibility of PH and right heart dysfunction in accordance with clus- ter assignment.

PATIENT SUSCEPTIBILITY AND GENETIC PREDISPOSITION TO VENTRICULAR FAILURE IN RESPONSE TO AS-INDUCED PRESSURE OVERLOAD REQUIRE FURTHER INVESTIGATION.

Interestingly, patients in cluster 4 showed dilata- tion of left ventricular end-diastolic diameter and right midventricular diameter, and concomitant with the structural changes, LVEF and TAPSE were impaired in comparison with patients in reference cluster 1. As the prevalence of coronary artery disease was equal in all clusters, future studies

should investigate individual genetic predisposition to ventricular failure in response to AS-induced

pressure overload. Notably, Généreux et al (4), who categorized patients with AS according to the extent of cardiac damage by applying conventional divisive, top- down clustering, also observed that extravalvular damage to the heart and pulmonary circulation does not necessarily occur in a sequential order of left heart failure, PH, and right heart dysfunction but may vary depending on pa- tient susceptibility and genetic predisposition. Apart from the possibly contributing influence of genetic predisposition, a self-amplifying loop of left and right atrial enlargement, atrial fibrillation and/or flutter, and advanced atrial remodeling fol- lowed by mitral and tricuspid regurgitation might have been triggered in patients in clusters 3 and 4 (Central Illustration).

STUDY LIMITATIONS.

As an elderly patient popula- tion with potentially multiple comorbidities approaching the end of life was analyzed, all-cause mortality within 2 years after TAVR was considered a clinically meaningful primary outcome measure, without further investigating the cause of death

(notably, increased HR for mortality in clusters 3 and 4 also remained significant if no censoring of survival data was applied) (Supplemental Figure 3). Patients unavoidably could have died from addi- tional cardiovascular diseases, such as coronary ar- tery disease or arrhythmic disorders including concomitant risk for stroke and bleeding from anti- coagulation therapy. Moreover, elderly patients and patients with more advanced cardiac diseases (as it is the case for patients in clusters 3 and 4 with EuroSCOREs of 29.1% ! 20.9% and 22.6% ! 14.6%,

respectively) often have reduced physiological reserve and are hence more vulnerable to any new insult, such as systemic infections. A retrospective, observational cohort study analyzing mortality in 1,197 patients with severe AS undergoing TAVR demonstrated that infection (14.0%) and malignancy (11.0%) were the leading causes of noncardiac death

(20). Moreover, Minamino-Muta et al (20) described age, male sex, low body mass index, diabetes mel- litus, anemia, and dialysis as factors to be inde- pendently associated with noncardiac death in patients with severe AS. Cox regression analysis performed in the present study confirmed well- established predictors for mortality such as the EuroSCORE, estimated glomerular filtration rate, chronic obstructive pulmonary disease, right atrial area, mPAP, and PVR.

Univariate analysis further suggested an associa-

tion between AVG_{mean} and mortality after TAVR. It is

important to note that traditional regression models assume a parametric linear relationship between predictor variable and outcome, which might hold true for the relationship between deteriorating car- diac output or renal function and mortality. However, the AVG initially increases with progressive narrow- ing of the aortic valve area but later decreases as the left ventricle decompensates ("low-flow, low- gradient AS"), as observed in high-risk clusters 3 and 4 (21). The benefit of the present unsupervised agglomerative clustering approach in combination with an ANN for prospective cluster assignment obviously lies in its ability to allow variables to act in concert without a priori assumptions, thereby deci- phering intricate interdependencies. As this is a bicentric, post hoc analysis with the majority of pa- tients excluded because of missing RHC, prospective and external validation on complete cohorts is required to confirm our model of 4 distinct clinical phenotypes among patients with severe AS and to demonstrate that the ANN can predict cluster assignment and concomitant mortality also in different settings.

CONCLUSIONS

Machine learning aids in capturing the complexity of clinical presentations observed in patients with severe AS by clustering basically 4 distinct clinical pheno- types, which differ in mortality depending on disease severity. Unsupervised agglomerative clustering de- serves particular attention, as the extent of cardiac and pulmonary circulatory impairments does not necessarily follow a sequential order from AS-induced left heart dysfunction, through PH, to ultimately right heart failure but is also influenced by comorbidities and ageing in general. As predefined cluster assign- ments could additionally be recapitulated by an ANN, this study advocates for the use of machine learning technology for individual cluster assignment and hence refined risk stratification prior to TAVR for pa- tients with severe AS in future clinical practice.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

ADDRESS FOR CORRESPONDENCE: Dr Christian Kupatt, First Department of Medicine, Klinikum rechts der Isar, Technical University of Munich, Ismaninger Straße 22, 81675 Munich, Germany. E-mail: christian.kupatt@tum.de.

PERSPECTIVES

WHAT IS KNOWN? AS can trigger a deleterious cascade of impairments, including left heart dysfunction, PH, and eventually right heart failure. Clinical presentations therefore appear het- erogeneous, depending on disease progression and possibly aggravated by comorbidities such as coronary artery disease, atrial fibrillation, and concomitant valvular defects.

WHAT IS NEW? Machine learning (ie, unsupervised agglom- erative, hierarchical clustering in conjunction with an ANN) aids in reducing the complexity of clinical presentations observed in patients with severe AS by distinguishing basically 4 distinct phenotypes, which reflect different extents of disease severity and hence differ in mortality. By distinguishing phenotypes without the constraint of any a priori assumption (ie, ignoring a hypothesized sequential order of isolated, AS-induced impair- ments of the heart and pulmonary circulation), this study sheds light from a new perspective on structural and hemodynamic impairments in patients with severe AS and finally demonstrates that structural alterations in left and right heart morphology constitute an equally sensitive indicator of poor prognosis compared with high-grade PH.

WHAT IS NEXT? Clusterwise comparison of future echocar- diographic follow-up data could reveal novel therapeutic targets beyond correction of severe AS. Addressing the irreversibility of PH and the persistence of severe tricuspid regurgitation should obtain paramount priority.

REFERENCES

1. Osnabrugge RLJ, Mylotte D, Head SJ, et al. Aortic stenosis in the elderly. *J Am Coll Cardiol*. 2013;62(11):1002–1012.

2. Weber L, Rickli H, Haager PK, et al. Haemody- namic mechanisms and long-term prognostic impact of pulmonary hypertension in patients with severe aortic stenosis undergoing valve replace- ment: impact of PH in severe aortic stenosis. *Eur J Heart Fail*. 2019;21(2):172–181.

3. Cavalcante JL, Simon MA, Chan SY. Compre- hensive right-sided assessment for transcatheter aortic valve replacement risk stratification: time for a change. *J Am Soc Echocardiogr*. 2017;30(1):47–51.

4. Généreux P, Pibarot P, Redfors B, et al. Staging classification of aortic stenosis based on the extent of cardiac damage. *Eur Heart J*. 2017;38(45):3351–3358.

5. Nishimura RA, Otto CM, Bonow RO, et al. 2014 AHA/ACC guideline for the management of pa- tients with valvular heart disease: executive sum- mary: a report of the American College of Cardiology/American Heart Association Task Force

on Practice Guidelines. *Circulation*. 2014;129(23): 2440–2492.

6. Du W, Elemento O. Cancer systems biology: embracing complexity to develop better anti- cancer therapeutic strategies. *Oncogene*. 2015;34(25):3215–3225.

7. Lancaster MC, Salem Omar AM, Narula S, Kulkarni H, Narula J, Sengupta PP. Phenotypic clustering of left ventricular diastolic function parameters. *J Am Coll Cardiol Img*. 2019;12(7): 1149–1161.

8. Tokodi M, Shrestha S, Bianco C, et al. Inter- patient similarities in cardiac function. *J Am Coll Cardiol Img*. 2020;13(5):1119–1132.

9. Kwak S, Lee Y, Ko T, et al. Unsupervised cluster analysis of patients with aortic stenosis reveals distinct population with different phenotypes and outcomes. *Circ Cardiovasc Imaging*. 2020;13(5): e009707.

10. Casaclang-Verzosa G, Shrestha S, Khalil MJ, et al. Network tomography for understanding

phenotypic presentations in aortic stenosis. *J Am Coll Cardiol Img*. 2019;12(2):236–248.

11. Baumgartner H, Falk V, Bax JJ, et al. 2017 ESC/EACTS guidelines for the management of valvular heart disease. *Eur Heart J*. 2017;38(36): 2739–2791.

12. Galiè N, Humbert M, Vachiery J-L, et al. 2015 ESC/ERS guidelines for the diagnosis and treat- ment of pulmonary hypertension: the Joint Task Force for the Diagnosis and Treatment of Pulmo- nary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS) endorsed by: Association for Euro- pean Paediatric and Congenital Cardiology (AEPC), International Society for Heart and Lung Trans- plantation (ISHLT). *Eur Heart J*. 2016;37(1):67–119.

13. O'Sullivan CJ, Wenaweser P, Ceylan O, et al. Effect of pulmonary hypertension hemodynamic presentation on clinical outcomes in patients with severe symptomatic aortic valve stenosis under- going transcatheter aortic valve implantation: in- sights from the new proposed pulmonary

hypertension classification. *Circ Cardiovasc Interv*. 2015;8(7):e002358.

14. Schewel J, Schmidt T, Kuck K-H, Frerker C, Schewel D. Impact of pulmonary hypertension hemodynamic status on long-term outcome after transcatheter aortic valve replacement. *J Am Coll Cardiol Intv*. 2019;12(21):2155–2168.

15. Shah AD, Bartlett JW, Carpenter J, Nicholas O, Hemingway H. Comparison of random forest and parametric imputation models for imputing missing data using MICE: a CALIBER study. *Am J Epidemiol*. 2014;179(6):764–774.

16. Stekhoven DJ, Buhlmann P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics*. 2012;28(1):112–118.

17. Rubin DB. Inference and missing data. *Biometrika*. 1976;63(3):581–592.

18. Makkar RR, Fontana GP, Jilaihawi H, et al. Transcatheter aortic-valve replacement for inoperable severe aortic stenosis. *N Engl J Med*. 2012;366(18):1696–1704.

19. Masri A, Abdelkarim I, Sharbaugh MS, et al. Outcomes of persistent pulmonary hypertension following transcatheter aortic valve replacement. *Heart*. 2018;104(10):821–827.

20. Minamino-Muta E, Kato T, Morimoto T, et al. Causes of death in patients with severe aortic stenosis: an observational study. *Sci Rep*. 2017;7(1):14723.

21. Awtry E, Davidoff R. Low-flow/low-gradient aortic stenosis. *Circulation*. 2011;124(23):e739–e741.

APPENDIX For supplemental figures and tables, please see the online version of this paper.