

Machine learning-based mortality prediction of patients undergoing cardiac resynchronization therapy: the SEMMELWEIS-CRT score

Márton Tokodi ^{1†}, Walter Richard Schwertner^{1†}, Attila Kovács ¹, Zoltán Tóser ², Levente Staub², András Sárkány ², Bálint Károly Lakatos ¹, Anett Behon ¹, András Mihály Boros¹, Péter Perge¹, Valentina Kutýifa ^{1,3}, Gábor Széplaki¹, László Gellér¹, Béla Merkely ^{1*‡}, and Annamária Kosztin ^{1‡}

¹Heart and Vascular Center, Semmelweis University, 68 Városmajor St., Budapest 1122, Hungary; ²Argus Cognitive, Inc., 16 Cavendish Ct., Lebanon, NH 03766, USA; and ³Heart Research Follow-up Program, Cardiology Division, University of Rochester Medical Center, 265 Crittenden Blvd., Box 653, Rochester, NY 14642, USA

Received 7 August 2019; revised 24 October 2019; editorial decision 2 December 2019; accepted 3 December 2019; online publish-ahead-of-print 10 January 2020

Aims

Our aim was to develop a machine learning (ML)-based risk stratification system to predict 1-, 2-, 3-, 4-, and 5-year all-cause mortality from pre-implant parameters of patients undergoing cardiac resynchronization therapy (CRT).

Methods and results

Multiple ML models were trained on a retrospective database of 1510 patients undergoing CRT implantation to predict 1- to 5-year all-cause mortality. Thirty-three pre-implant clinical features were selected to train the models. The best performing model [SEMMELWEIS-CRT score (personalized assessment of estimated risk of mortality with machine learning in patients undergoing CRT implantation)], along with pre-existing scores (Seattle Heart Failure Model, VALID-CRT, EAARN, SCREEN, and CRT-score), was tested on an independent cohort of 158 patients. There were 805 (53%) deaths in the training cohort and 80 (51%) deaths in the test cohort during the 5-year follow-up period. Among the trained classifiers, random forest demonstrated the best performance. For the prediction of 1-, 2-, 3-, 4-, and 5-year mortality, the areas under the receiver operating characteristic curves of the SEMMELWEIS-CRT score were 0.768 (95% CI: 0.674–0.861; $P < 0.001$), 0.793 (95% CI: 0.718–0.867; $P < 0.001$), 0.785 (95% CI: 0.711–0.859; $P < 0.001$), 0.776 (95% CI: 0.703–0.849; $P < 0.001$), and 0.803 (95% CI: 0.733–0.872; $P < 0.001$), respectively. The discriminative ability of our model was superior to other evaluated scores.

Conclusion

The SEMMELWEIS-CRT score (available at semmelweisrtscore.com) exhibited good discriminative capabilities for the prediction of all-cause death in CRT patients and outperformed the already existing risk scores. By capturing the non-linear association of predictors, the utilization of ML approaches may facilitate optimal candidate selection and prognostication of patients undergoing CRT implantation.

Keywords

Heart failure • Cardiac resynchronization therapy • Risk stratification • Mortality prediction • Machine learning • Precision medicine

* Corresponding author. Tel: +361-458-68-10, Fax: +361-458-68-17, Email: merkely.study@gmail.com

† These authors are joint first authors.

‡ These authors are joint last authors.

© The Author(s) 2020. Published by Oxford University Press on behalf of the European Society of Cardiology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Introduction

Cardiac resynchronization therapy (CRT) is a key component in the management of symptomatic heart failure with reduced ejection fraction and wide QRS complex.¹ Based on the report of the European Heart Rhythm Association, over 90 CRT implantations per million population are performed annually in the ESC countries.² Although CRT improves mortality, functional capacity, clinical symptoms, and quality of life in a certain patient subpopulation, not everyone benefits equally and mortality rates still remain high among these patients.^{3–7}

The recognition of this variability in outcomes has prompted efforts in the risk stratification of CRT patients based on pre-implant assessments. However, the currently available risk scores have several shortcomings (e.g. lack of generalizability and impact analyses, omitting routinely assessed, powerful predictors) which hamper their utilization in the everyday clinical practice.⁸ Therefore, more precise and personalized methods are required. The recent improvements in computation power and software technologies have led to the flourishing of machine learning (ML), a field of artificial intelligence (AI), which seems to be a promising tool to meet this compelling demand.⁹

Machine learning refers to a collection of techniques that gives AI the ability to learn complex rules and to identify patterns from multidimensional datasets, without being explicitly programmed or applying any a priori assumptions. It has been effectively utilized in many areas of cardiology such as precision phenotyping, diagnostics, and prognostication including the prediction of hospital readmissions and mortality.^{10–12} Although, heart failure patients undergoing CRT implantation represent another important target population for mortality prediction, only few studies have applied ML to tackle this issue.^{13–15}

Accordingly, our aim was to design and evaluate a ML-based risk stratification system to predict 1-, 2-, 3-, 4-, and 5-year mortality from pre-implant parameters of patients undergoing CRT implantation. We hypothesized that ML can capture high-dimensional, non-linear relationships among clinical features and a risk stratification system can be developed that predicts mortality for individual patients more accurately than the currently available risk scores.

Methods

Study population and protocol

We identified 2282 patients who underwent successful CRT implantation at the Heart and Vascular Center of Semmelweis University (Budapest, Hungary) between September 2000 and December 2017. For each of these patients, pre-implant clinical characteristics such as demographics, medical history, physical status and vitals, currently applied medical therapy, electrocardiogram, echocardiographic, and laboratory parameters were extracted retrospectively from electronic medical records and entered to our structured database.

An additional prospective database of patients undergoing CRT implantation between January 2009 and December 2011 was also utilized. Patients included in both the retrospective and the prospective databases were removed from the retrospective database. In this way, the two cohorts were completely independent and they could be used as training and test cohorts for ML algorithms.

The study protocol (Supplementary material online, Figure S1) complies with the Declaration of Helsinki and it was approved by the Regional

and Institutional Committee of Science and Research Ethics (approval No. 161/2019).

Study outcomes

Follow-up data [status (dead or alive), date of death] was obtained for all patients from the National Health Insurance Database of Hungary. Patients with shorter than 5-year follow-up duration (614 patients in the retrospective and 0 patient in the prospective database) were excluded from all analyses. The primary endpoint of our study was all-cause mortality.

Feature selection and data pre-processing

Our structured database initially comprised over 100 easily obtainable clinical variables (so-called features). Firstly, features included in both the retrospective and the prospective databases were identified ($n=49$). Then, features missing for >40% of cases ($n=16$) were excluded. The final set of input features included 33 pre-implant clinical variables (Supplementary material online, Table S1).

Missing values were imputed using the mean imputation method which replaces the missing values of a certain variable with the mean of the available cases. As the range of different features varied widely and some of the utilized algorithms required the data to be normalized, Z-score normalization was performed after imputation.

Model development

We used the follow-up data to generate six classes of possible outcomes: death during the 1st (class 1), the 2nd (class 2), the 3rd (class 3), the 4th (class 4), the 5th year after CRT implantation (class 5), and no death during the first 5 years following the implantation (class 6). The task of ML algorithms was to predict the probability distribution (i.e. class membership probabilities) of each patient over these classes based on the pre-implant clinical features.

Model development included trials of several ML classifiers such as logistic regression, ridge regression, support vector machines, k -nearest neighbours classifier, gradient boosting classifier, random forest, conditional inference random forest, and multi-layer perceptron. Models were trained with stratified 10-fold cross-validation on the training cohort and a grid search approach was used to tune the hyper-parameters of each ML algorithm (Supplementary material online, Table S2).

The outputs of each model were series of six values representing the previously defined class membership probabilities (Figure 1A). The sum of these probabilities is equal to one in each patient. To create binary classifiers, we calculated cumulative class membership probabilities by summing these values until the given year of follow-up (Figure 1B). The computed cumulative probabilities were then calibrated using Platt's scaling and the survival curve could be plotted for each patient (Figure 1C). The calibration of the model was evaluated using Brier score which is defined as the mean squared difference between the observed and the predicted outcome. Expected survival was also calculated from the annual calibrated cumulative probabilities (Figure 1D).

To quantify the model's discriminative capabilities in each year, receiver operating characteristic (ROC) curve analysis was performed and area under the curve (AUC) was calculated. The mean AUC of 1-, 2-, 3-, 4-, and 5-year calibrated cumulative probabilities was calculated and it served as the major metric to assess a model's performance.

Model testing

The model with the highest mean AUC was selected for further evaluation and it is referenced as the SEMMELWEIS-CRT (personalized assessment of estimated risk of mortality with machine learning in patients undergoing CRT implantation) score throughout the entire

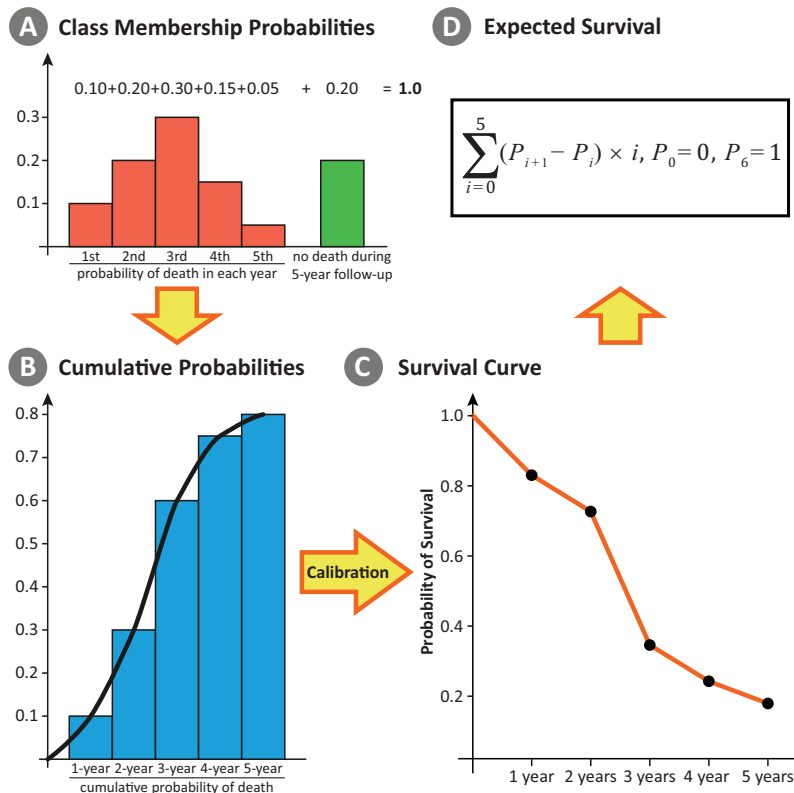


Figure 1 Computation of survival probabilities: illustrating the methodology through an example case. (A) The outputs of each model were series of class membership probabilities. (B) Cumulative probabilities were calculated by summing these values until the given year of follow-up. (C) To calibrate the cumulative probabilities, Platt's scaling was performed. Using these calibrated cumulative probabilities, the survival curve could be plotted for each patient. (D) Then, the expected survival time of each patient was estimated from the annual survival probabilities. P_i , the calibrated cumulative probability of all-cause mortality at year i .

manuscript. To determine whether the model remains accurate when new data are fed into it, we tested it on the patients of the test cohort.

For each patient in the test cohort, we also computed pre-existing risk scores (Seattle Heart Failure Model, VALID-CRT, EAARN, ScREEN, and CRT-score).^{16–20} Their prediction capabilities were quantified annually with AUCs and they were compared with SEMMELWEIS-CRT score using the DeLong test.

Feature importances

To determine the major predictors of all-cause mortality in our patient population, permutation feature importances were computed from the final model. Permutation feature importance measures the importance of an input feature by calculating the increase in the model's prediction error after permuting its values. A feature is considered important if shuffling its values decreases the model's discriminative capability as the model relies heavily on that feature for the prediction. A feature is unimportant if shuffling its values leaves the AUC unchanged as in this case the model ignores the feature for the prediction.

Results

Baseline clinical characteristics

The final training cohort included 1510 patients [66 ± 10 years, 1141 (76%) males] who underwent CRT implantation. A total of 158 CRT

patients [67 ± 10 years, 127 (80%) males] were prospectively enrolled and entered to the test database. During the 5-year follow-up period, 805 (53%) patients died in the training cohort and there were 80 (51%) deaths in the test cohort. [Supplementary material online, Table S3](#) shows the baseline characteristics of both cohorts and the comparisons between patients who were dead and alive at 5-year follow-up.

Prediction of all-cause mortality

Among the evaluated ML classifiers, random forest (i.e. SEMMELWEIS-CRT score) yielded the highest AUCs for the prediction of all-cause mortality at 1-, 2-, 3-, 4-, and 5-year follow-up in the test cohort ([Table 1](#) and [Figure 2](#)). Calibration improved the Brier scores of the final model ([Supplementary material online, Table S5](#)).

Comparison of the SEMMELWEIS-CRT score with the pre-existing scores

When compared with the pre-existing risk scores, the SEMMELWEIS-CRT score demonstrated significantly better response prediction and greater discrimination of mortality ([Table 1](#)). The CRT-score exhibited the best performance among the pre-existing risk scores; however, our random forest-based classifier was still superior to it for the prediction of 5-year outcome. Regarding

Table 1 Area under the receiver operating characteristic curve of the different scores

	1 year	2 years	3 years	4 years	5 years	Mean
SEMMELEIS-CRT	0.768 (0.674–0.861)	0.793 (0.718–0.867)	0.785 (0.711–0.859)	0.776 (0.703–0.849)	0.803 (0.733–0.872)	0.785
SHFM ²⁰	0.537 (0.426–0.647)*	0.543 (0.445–0.642)*	0.539 (0.447–0.632)*	0.544 (0.453–0.635)*	0.544 (0.454–0.634)*	0.541
EAARN ¹⁸	0.602 (0.505–0.699)*	0.627 (0.539–0.714)*	0.653 (0.570–0.736)*	0.649 (0.566–0.731)*	0.643 (0.560–0.726)*	0.635
VALID-CRT ¹⁶	0.529 (0.416–0.643)*	0.618 (0.523–0.713)*	0.638 (0.552–0.725)*	0.637 (0.550–0.724)*	0.650 (0.564–0.737)*	0.614
CRT-score ¹⁷	0.722 (0.637–0.806)	0.743 (0.667–0.818)	0.732 (0.657–0.807)	0.720 (0.644–0.795)	0.693 (0.615–0.771)*	0.722
ScREEN ¹⁹	0.595 (0.516–0.673)*	0.555 (0.477–0.633)*	0.536 (0.460–0.612)*	0.525 (0.449–0.601)*	0.549 (0.474–0.624)*	0.552

*P < 0.05 vs. SEMMELEIS-CRT, DeLong test. Cell contents are areas under the receiver operating characteristic curves with 95% confidence intervals. SHFM, Seattle Heart Failure Model.

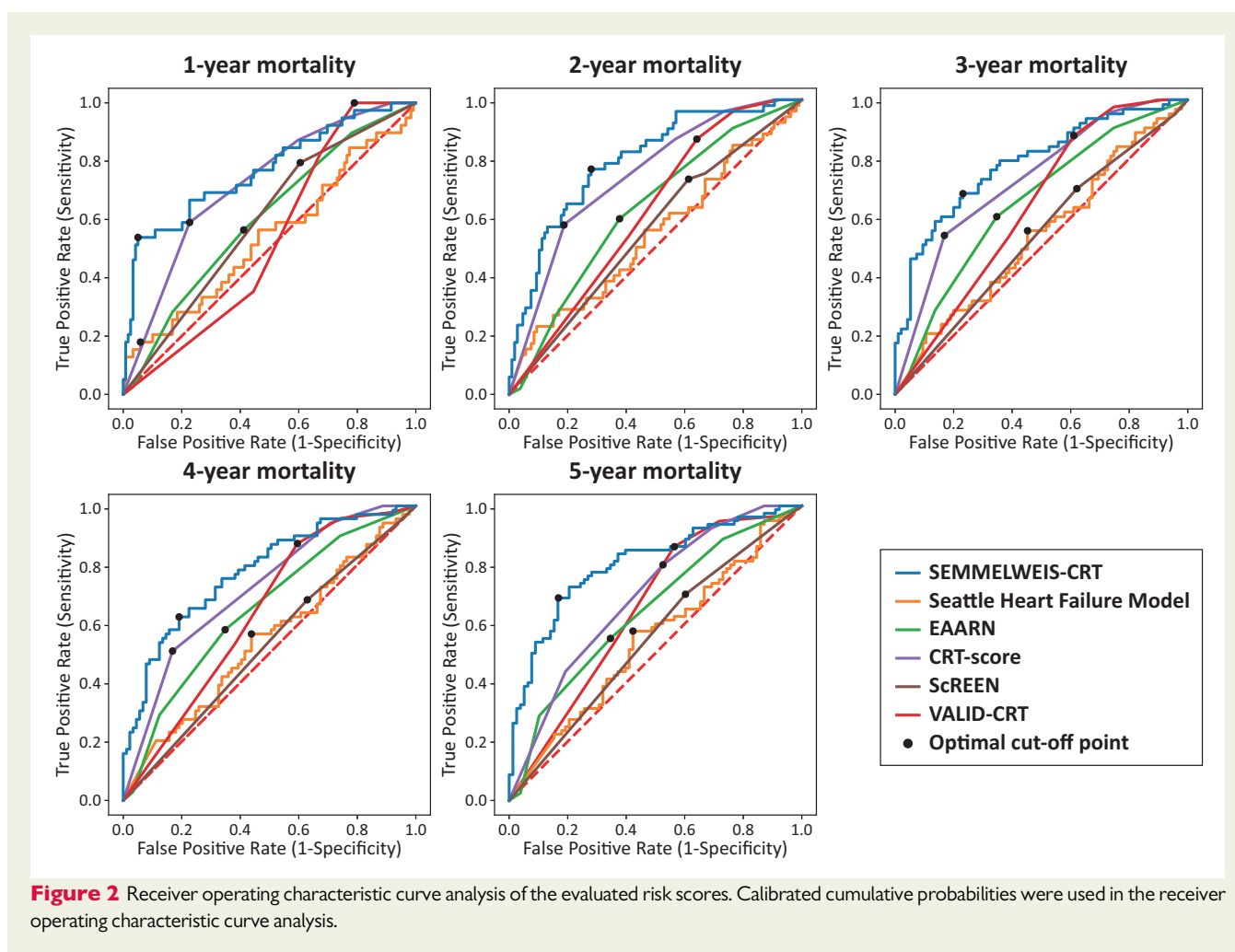


Figure 2 Receiver operating characteristic curve analysis of the evaluated risk scores. Calibrated cumulative probabilities were used in the receiver operating characteristic curve analysis.

the rest of the risk scores, the SEMMELWEIS-CRT score significantly outperformed them at all of the investigated time points.

Most important predictors of mortality as assessed with machine learning

Leading predictors of all-cause mortality are presented on *Figure 3* and the full list of feature importances is provided as [Supplementary material online, Table S6](#). Older age, higher serum levels of creatinine, lower values of left ventricular ejection fraction, serum sodium, haemoglobin concentration, and glomerular filtration rate were associated with higher predicted probability of all-cause mortality (*Figure 4*). However, as random forest captures complex high-level interactions among a multitude of variables, it is challenging to determine the effect of a single feature on the predicted probability of mortality and these individual relationships should be interpreted with caution.

Machine learning-based risk stratification

Based on the predicted probability of death, patients were split into four quartiles at each year of follow-up. As depicted by Kaplan–Meier curves, there was significant difference in the distribution of events across the quartiles at all years and a graded increase in event rates could be observed while moving from the 2nd quartile through the 4th quartile (*Figure 5*). At 1-year follow-up, being categorized to the 4th quartile was associated with a more than 7-fold increased risk of death compared with those in the 1st quartile (*Table 2*). At 2-, 3-, 4-, and 5-year follow-up, patients in 3rd and 4th quartiles exhibited a significantly increased risk of mortality compared with those in the 1st quartile (*Table 2*). The expected survival of patients was monotonously decreasing from the 1st through the 4th quartile in each year ([Supplementary material online, Table S7](#)).

Discussion

In the present study, we developed and tested a ML-based risk stratification tool to predict all-cause mortality of CRT patients during a 5-year follow-up period (*Take home figure*). Among the evaluated ML classifiers, random forest demonstrated the best performance; therefore, this algorithm was used to create the SEMMELWEIS-CRT score. With an average AUC over 0.700, the SEMMELWEIS-CRT score significantly outperformed the other currently available risk scores. We also developed an online calculator (available at semmelweiscrtscore.com) to enable a convenient, interactive, and personalized calculation of predicted mortality in patients undergoing CRT implantation.

Importance of risk assessment in patients undergoing cardiac resynchronization therapy implantation

Cardiac resynchronization therapy induces reverse left ventricular remodelling and improves outcomes in a certain subgroup of heart failure patients.^{2,21} Despite these well-known beneficial effects, individual outcomes vary substantially. In the past years, several studies have investigated predictors that contribute to this variation and numerous prognostic models have been developed by combining multiple risk factors.^{16–19} However, these currently available risk scores

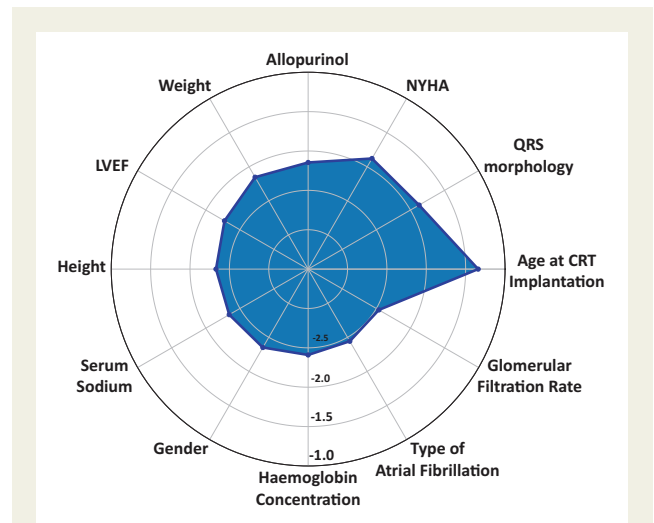


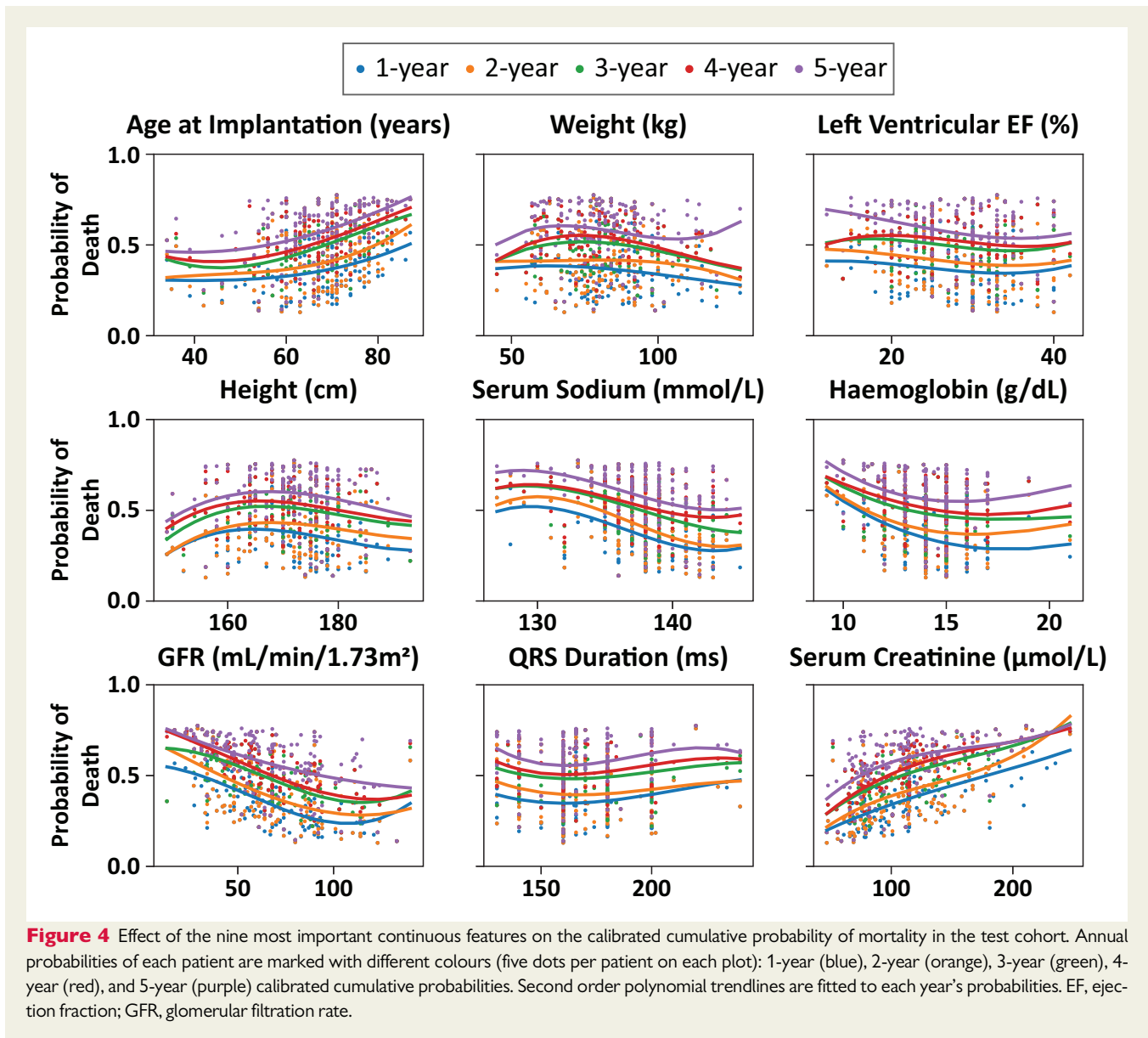
Figure 3 The 12 most important predictors of all-cause mortality as assessed by the SEMMELWEIS-CRT score. The importance of each feature was quantified by calculating the decrease in the model's performance (area under the receiver operating characteristic curve) after permuting its values (permutation feature importances method). The higher its value, the more important the feature is. As the values of feature importances were spread over a wide range (more orders of magnitude), base-10 logarithmic transformation was performed to facilitate plotting. CRT, cardiac resynchronization therapy; LVEF, left ventricular ejection fraction; NYHA, New York Heart Failure Association functional class.

have shortcomings and physicians are still reluctant to use them in daily clinical practice.⁸

The major limitation is the insufficient reliability and ineffectiveness for risk assessment at the individual patient level as outcome estimates have been extrapolated from large clinical trials. Although, these scores offer general guidance and they are effective at predicting outcomes at the population level, there remains a significant gap in the capability to predict outcomes for an individual patient.²³ On the other hand, individual prognostication remains essential to develop appropriate personalized treatment plans and to make critical medical decisions based on life expectancy. These facts emphasize the need for more precise assessment through capturing the complex underlying interactions of predictors. With the SEMMELWEIS-CRT score, we intended to develop a more personalized approach for the risk assessment of patients undergoing CRT implantation.

Risk stratification with machine learning

Simultaneously interpreting the myriad risk predictors in an individual patient is challenging for clinicians. As a vast number of clinical variables associated with mortality needs to be considered, the complexity of assessment increases, making it more difficult for clinicians to draw an overall conclusion regarding risk in an individual patient. Moreover, the potential influence of complex and hidden interactions between several weaker predictors is often overlooked. In this study, we demonstrated that ML is capable to overcome these challenges by leveraging complex higher-level interactions among a multitude of



clinical features. Accordingly, our model exhibited improved discrimination and predictive range with respect to all-cause mortality compared with the pre-existing risk scores. Moreover, the SEMMELWEIS-CRT score was capable of identifying patients with robustly increased risk of all-cause mortality (4th quartile) during the entire follow-up period.

With the increasing availability of enormous electronic datasets, ML algorithms have emerged as highly effective methods for medical prediction problems, with the potential to augment risk stratification.⁹ By making no a priori assumptions about causative factors, ML enables an agnostic exploration of all available data for non-linear patterns that may predict a particular individual's risk, i.e. personalized risk stratification.

Our evaluation of ML algorithms was rigorous, including trials of numerous different classifiers within a wide hyper-parameter space. Among the evaluated algorithms, the best performing model was the

random forest classifier which is consistent with previous studies using ML to predict clinical endpoints.^{13,24–26}

There are various risk models available for the risk assessment of patients from the entire heart failure spectrum.^{20,27} However, in our analysis, we focused exclusively on CRT recipients and we generated models that recognize patterns in the clinical characteristics of this specific subset of heart failure patients. Moreover, many of the pre-existing scores provide risk estimates for only a distinct time interval. In contrast, our goal was to build a model that could assess the risk of mortality annually from 1 to 5 years. Recently, Kalscheur et al.¹³ have developed a ML-based risk assessment tool and their model exhibited comparable discriminative capabilities to ours. However, their model was limited to predict 1-year outcomes, while the SEMMELWEIS-CRT score offers prediction of mortality risk at 1-, 2-, 3-, 4-, and 5-year follow-up.

Ideally, ML models, such as the one developed in the present study, will be integrated into electronic medical record systems and they

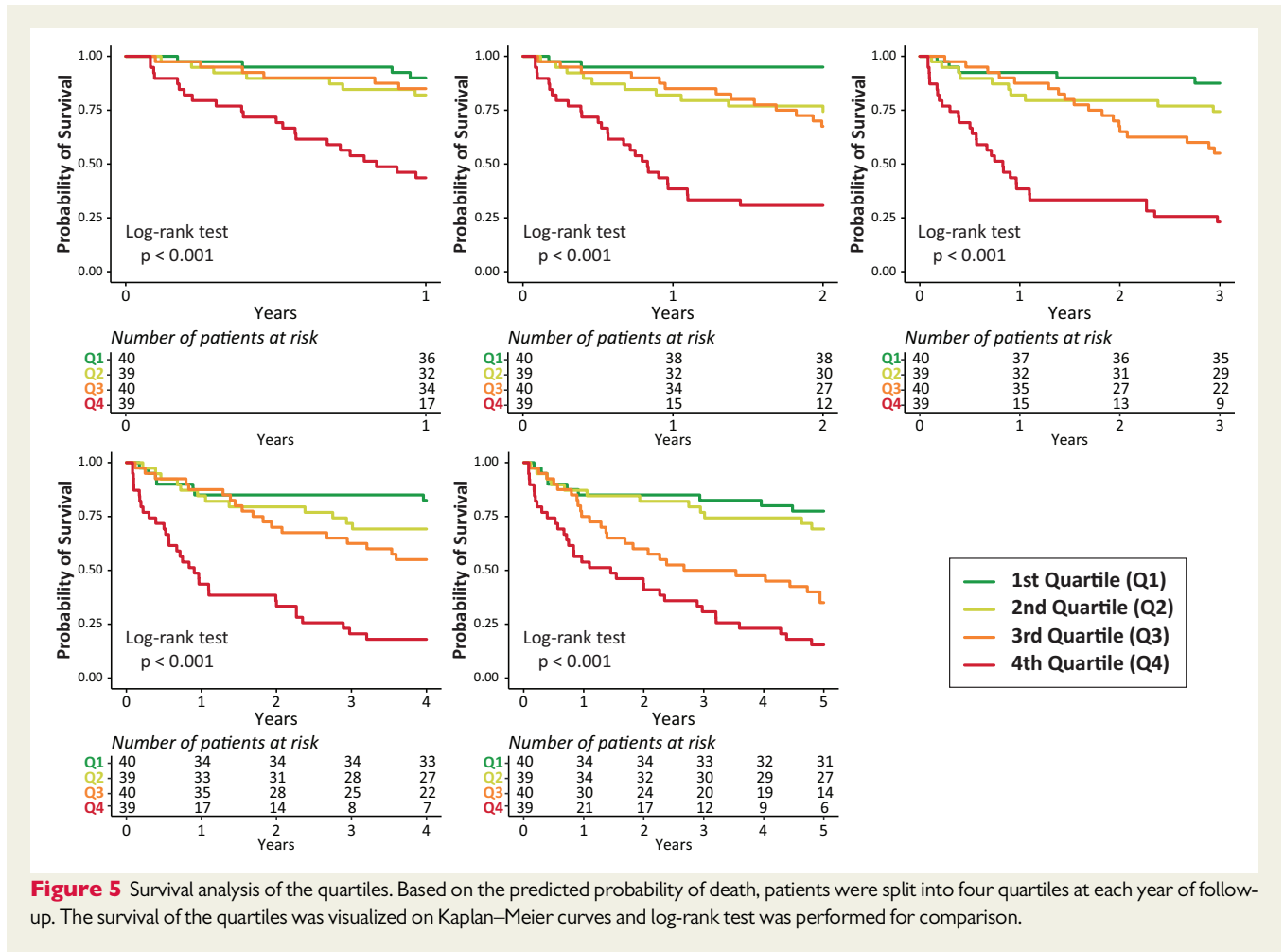


Table 2 Hazard ratios of all-cause mortality in different quartiles

	1 year	2 years	3 years	4 years	5 years
2nd vs. 1st quartile	1.89 (0.55–6.45) <i>P</i> = 0.301	5.55 (1.22–25.35) <i>P</i> = 0.010	2.18 (0.44–3.36) <i>P</i> = 0.142	1.81 (0.717–4.60) <i>P</i> = 0.203	1.40 (0.59–3.33) <i>P</i> = 0.439
3rd vs. 1st quartile	1.56 (0.44–5.52) <i>P</i> = 0.487	7.30 (1.65–32.37) <i>P</i> < 0.001	4.18 (1.55–11.28) <i>P</i> = 0.002	2.88 (1.20–6.90) <i>P</i> = 0.012	3.75 (1.75–8.04) <i>P</i> < 0.001
4th vs. 1st quartile	7.92 (2.72–23.07) <i>P</i> < 0.001	21.55 (5.10–91.06) <i>P</i> < 0.001	10.59 (4.07–27.56) <i>P</i> < 0.001	8.16 (3.56–18.72) <i>P</i> < 0.001	6.71 (3.17–14.21) <i>P</i> < 0.001

Patients were split (repeatedly) into four quartiles based on the predicted probably of death in each year. As the quartiles in each year might contain different set of patients, row-wise evaluation of hazard ratios should be avoided. Cell contents are hazard ratios (95% confidence interval) with *P*-values calculated using Cox proportional-hazards models.

will operate in the background providing real-time, personalized risk assessment based on the electronically available clinical features. Consequently, clinicians do not have to calculate a patient’s risk manually that may enhance the model’s feasibility in clinical practice. Another potential benefit of ML algorithms is the capability to

assimilate new data in real-time to continuously improve its own predictive accuracy. The SEMMELWEIS-CRT score uses 33 clinical variables. Majority of them are routinely assessed during the management of heart failure; therefore, they are readily available from electronic medical

SEMMELWEIS-CRT score

personalized assessment of estimated risk of mortality with machine learning in patients undergoing CRT implantation

WebCalculator is available at semmelweiscrtscore.com

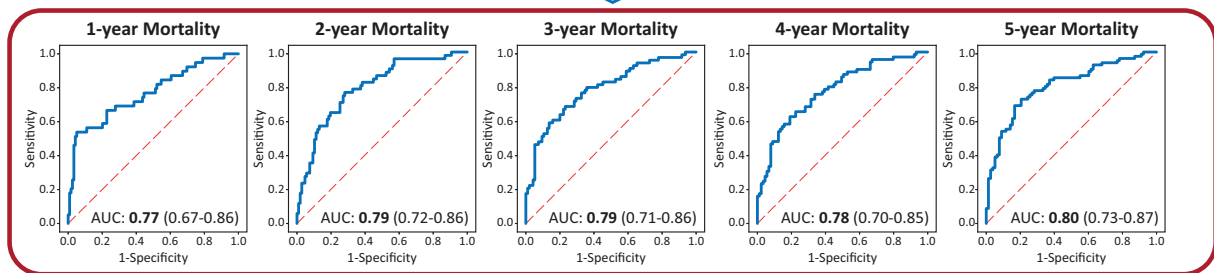


Heart Failure Patients Undergoing CRT Implantation
Training Cohort: 1510 patients; Validation Cohort: 158 patients

Pre-Implant Assessment



Machine Learning (Random Forest) - Based Mortality Prediction



Take home figure Using commonly available pre-implant clinical variables, the machine learning-based SEMMELWEIS-CRT score (available at semmelweiscrtscore.com) can effectively predict all-cause mortality of patients undergoing cardiac resynchronization therapy. AUC, area under the receiver operating characteristic curve; CRT, cardiac resynchronization therapy; ECG, electrocardiogram.

records. Moreover, our model was designed in a way to tolerate moderate number of missing parameters, however, with special regards to the most important features, high percentage of missing values may reduce the reliability of the prediction.

We also identified the most important predictors of all-cause mortality in this patient cohort. Many of these features have been described previously as influencing CRT outcomes, such as advanced age, male gender, non-left bundle branch block QRS morphology, history of or present atrial fibrillation at implantation, impaired renal function, and increased comorbidity burden.²⁸⁻³⁰ However, it is challenging to assess the independent impact of each variable on the predicted risk of mortality as ML models capture higher dimensional, non-linear interactions among features.

Future perspectives

The observed high efficacy of our random forest model suggests that ML should be integrated into the individual risk assessment of patients undergoing CRT implantation. We foresee that the role of ML-based prognostic risk scores will become increasingly relevant in

the near future and structured, dense databases in combination with state-of-the-art analytic approaches will pave the way to precision cardiovascular medicine.

Limitations

This study has several strengths and limitations to be acknowledged. To ensure the generalizability of our model, we trained our models with 10-fold cross-validation on a large database and we performed additional testing of the final model on an independent cohort of patients. However, our study represents results from a single centre; therefore, the SEMMELWEIS-CRT score should be validated in external centres to confirm its generalizability. Our score requires a broad spectrum of input variables that might discourage clinicians from its utilization at first glance. Thus, we designed our score in a way to tolerate missing values, nevertheless, it might be less reliable with a large number of missing variables. In spite of including established predictors of mortality in the final model, some relevant input features were excluded during model development due to the proportion of missing values. Inclusion of the omitted parameters (eg,

other comorbidities) might further improve the predictive capabilities of our model. Besides missing values, the relatively long-time course of retrospective data collection bears inherent limitations also regarding the changes in the guideline directed medical therapy. Another major limitation of risk score models is the lack of impact analyses to determine how the utilization of the models improves patient care and outcomes. Accordingly, future investigations should target the identification of treatment plans that specifically fit different levels of risk assessed by the SEMMELWEIS-CRT score. As the application of ML depends on the robustness of the database, practical use of our model in patient care would require careful and structured collection of data. However, this issue will resolve soon as large and structured databases are becoming widely available. Moreover, our model could be linked with electronic medical record systems to automatically calculate risk score obviating the manual computation of patients' risk and potentially increasing the model's use in clinical practice.

Conclusions

Using commonly available clinical variables, we developed and tested a random forest-based risk stratification system, the SEMMELWEIS-CRT score to effectively predict all-cause mortality in patients undergoing CRT implantation. Our ML-based risk assessment tool outperformed the pre-existing conventional risk scores. By capturing the non-linear association of predictors, the SEMMELWEIS-CRT score effectively outlined patient subgroups at high risk for mid- and long-term mortality. Therefore, the integration of these approaches into daily clinical practice may facilitate optimal candidate selection and prognostication of patients undergoing CRT implantation.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Funding

This work was supported by the National Research, Development and Innovation Office of Hungary (NKFI; NVKP_16-1-2016-0017 National Heart Program) and the Higher Education Institutional Excellence Program of the Ministry for Innovation and Technology in Hungary, within the framework of the Therapeutic Development thematic program of the Semmelweis University.

Conflicts of interest: B.M. receives lecture fees from Biotronik, Medtronic and Abbott. Other authors declare no conflicts of interest regarding this manuscript.

References

- Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, Falk V, González-Juanatey JR, Harjola V-P, Jankowska EA, Jessup M, Linde C, Nihoyannopoulos P, Parissis JT, Pieske B, Riley JP, Rosano GMC, Ruitlope LM, Ruschitzka F, Rutten FH, van der Meer P; ESC Scientific Document Group. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J* 2016;**37**:2129–2200.
- Raatikainen MJP, Arnar DO, Merkely B, Nielsen JC, Hindricks G, Heidebuchel H, Camm J. A decade of information on the use of cardiac implantable electronic devices and interventional electrophysiological procedures in the European Society of Cardiology Countries: 2017 report from the European Heart Rhythm Association. *Europace* 2017;**19**:ii1–ii90.
- Goldenberg I, Kutiyafa V, Klein HU, Cannom DS, Brown MW, Dan A, Daubert JP, Estes NAM, Foster E, Greenberg H, Kautzner J, Klempfner R, Kuniss M, Merkely B, Pfeffer MA, Quesada A, Viskin S, McNitt S, Polonsky B, Ghanem A, Solomon SD, Wilber D, Zareba W, Moss AJ. Survival with cardiac-resynchronization therapy in mild heart failure. *N Engl J Med* 2014;**370**:1694–1701.
- Normand C, Kaye DM, Povsic TJ, Dickstein K. Beyond pharmacological treatment: an insight into therapies that target specific aspects of heart failure pathophysiology. *Lancet* 2019;**393**:1045–1055.
- Steffel J, Ruschitzka F. Superresponse to cardiac resynchronization therapy. *Circulation* 2014;**130**:87–90.
- Kutiyafa V, Geller L, Bogyi P, Zima E, Aktas MK, Ozcan EE, Becker D, Nagy VK, Kosztin A, Szilagy S, Merkely B. Effect of cardiac resynchronization therapy with implantable cardioverter defibrillator versus cardiac resynchronization therapy with pacemaker on mortality in heart failure patients: results of a high-volume, single-centre experience. *Eur J Heart Fail* 2014;**16**:1323–1330.
- Nagy KV, Szeplaki G, Perge P, Boros AM, Kosztin A, Apor A, Molnar L, Szilagy S, Tahin T, Zima E, Kutiyafa V, Geller L, Merkely B. Quality of life measured with EuroQol-five dimensions questionnaire predicts long-term mortality, response, and reverse remodelling in cardiac resynchronization therapy patients. *Europace* 2018;**20**:1506–1512.
- Canepa M, Fonseca C, Chioncel O, Laroche C, Crespo-Leiro MG, Coats AJS, Mebazaa A, Piepoli MF, Tavazzi L, Maggioni AP. Performance of prognostic risk scores in chronic heart failure patients enrolled in the European Society of Cardiology Heart Failure Long-Term Registry. *JACC Heart Fail* 2018;**6**:452–462.
- Al'Aref SJ, Anchouche K, Singh G, Slomka PJ, Kolli KK, Kumar A, Pandey M, Maliakal G, van Rosendaal AR, Beecy AN, Berman DS, Leipsic J, Nieman K, Andreini D, Pontone G, Schoepf UJ, Shaw LJ, Chang HJ, Narula J, Bax JJ, Guan Y, Min JK. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *Eur Heart J* 2019;**40**:1975–1986.
- Ahmad T, Lund Lars H, Rao P, Ghosh R, Warier P, Vaccaro B, Dahlström U, O'Connor Christopher M, Felker GM, Desai Nihar R. Machine learning methods improve prognostication, identify clinically distinct phenotypes, and detect heterogeneity in response to therapy in a large cohort of heart failure patients. *J Am Heart Assoc* 7:e008081.
- Motwani M, Dey D, Berman DS, Germano G, Achenbach S, Al-Mallah MH, Andreini D, Budoff MJ, Cademartiri F, Callister TQ, Chang HJ, Chinnaiyan K, Chow BJ, Cury RC, Delago A, Gomez M, Gransar H, Hadamitzky M, Hausleiter J, Hindoyan N, Feuchtner G, Kaufmann PA, Kim YJ, Leipsic J, Lin FY, Maffei E, Marques H, Pontone G, Raff G, Rubinshtein R, Shaw LJ, Stehli J, Villines TC, Dunning A, Min JK, Slomka PJ. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *Eur Heart J* 2017;**38**:500–507.
- Motwani BJ, Downing NS, Bucholz EM, Dharmarajan K, Manhapra A, Li SX, Negahban SN, Krumholz HM. Analysis of machine learning techniques for heart failure readmissions. *Circ Cardiovasc Qual Outcomes* 2016;**9**:629–640.
- Kalscheur MM, Kipp RT, Tattersall MC, Mei C, Buhr KA, DeMets DL, Field ME, Eckhardt LL, Page CD. Machine learning algorithm predicts cardiac resynchronization therapy outcomes: lessons from the COMPANION trial. *Circ Arrhythm Electrophysiol* 2018;**11**:e005499.
- Feeny AK, Rickard J, Patel D, Toro S, Trulock KM, Park CJ, LaBarbera MA, Varma N, Niebauer MJ, Sinha S, Gorodeski EZ, Grimm RA, Ji X, Barnard J, Madabhushi A, Spragg DD, Chung MK. Learning prediction of response to cardiac resynchronization therapy. *Circ Arrhythm Electrophysiol* 2019;**12**:e007316.
- Cikes M, Sanchez-Martinez S, Claggett B, Duchateau N, Piella G, Butakoff C, Poulaur AC, Knappe D, Biering-Sorensen T, Kutiyafa V, Moss A, Stein K, Solomon SD, Bijns B. Machine learning-based phenogrouping in heart failure to identify responders to cardiac resynchronization therapy. *Eur J Heart Fail* 2019;**21**:74–85.
- Gasparini M, Klersy C, Leclercq C, Lunati M, Landolina M, Auricchio A, Santini M, Boriani G, Proclemer A, Leyva F. Validation of a simple risk stratification tool for patients implanted with cardiac resynchronization therapy: the VALID-CRT risk score. *Eur J Heart Fail* 2015;**17**:717–724.
- Hoke U, Mertens B, Khidir MJH, Schallig MJ, Bax JJ, Delgado V, Ajmone Marsan N. Usefulness of the CRT-SCORE for shared decision making in cardiac resynchronization therapy in patients with a left ventricular ejection fraction of ≤ 35 . *Am J Cardiol* 2017;**120**:2008–2016.
- Khatib M, Tolosana JM, Trucco E, Borràs R, Castel A, Berrueto A, Doltra A, Sitges M, Arbelo E, Matas M, Brugada J, Mont L. EAARN score, a predictive score for mortality in patients receiving cardiac resynchronization therapy based on pre-implantation risk factors. *Eur J Heart Fail* 2014;**16**:802–809.
- Providencia R, Marijon E, Barra S, Reitan C, Breitenstein A, Defaye P, Papageorgiou N, Duehmke R, Winnik S, Ang R, Klug D, Gras D, Oezkartal T, Segal OR, Deharo JC, Leclercq C, Lambiase PD, Fauchier L, Bordachar P, Steffel J, Sadoul N, Piot O, Borgquist R, Agarwal S, Chow A, Boveda S. Usefulness of a

- clinical risk score to predict the response to cardiac resynchronization therapy. *Int J Cardiol* 2018;**260**:82–87.
20. Levy WC, Mozaffarian D, Linker DT, Sutradhar SC, Anker SD, Cropp AB, Anand I, Maggioni A, Burton P, Sullivan MD, Pitt B, Poole-Wilson PA, Mann DL, Packer M. The Seattle Heart Failure Model: prediction of survival in heart failure. *Circulation* 2006;**113**:1424–1433.
 21. Bristow MR, Saxon LA, Boehmer J, Krueger S, Kass DA, De Marco T, Carson P, DiCarlo L, DeMets D, White BG, DeVries DW, Feldman AM. Cardiac-resynchronization therapy with or without an implantable defibrillator in advanced chronic heart failure. *N Engl J Med* 2004;**350**:2140–2150.
 22. Moss AJ, Hall WJ, Cannom DS, Klein H, Brown MW, Daubert JP, Estes NAM, Foster E, Greenberg H, Higgins SL, Pfeffer MA, Solomon SD, Wilber D, Zareba W. Cardiac-resynchronization therapy for the prevention of heart-failure events. *N Engl J Med* 2009;**361**:1329–1338.
 23. Allen LA, Matlock DD, Shetterly SM, Xu S, Levy WC, Portalupi LB, Mcllvannan CK, Gurwitz JH, Johnson ES, Smith DH, Magid DJ. Use of risk models to predict death in the next year among individual ambulatory patients with heart failure. *JAMA Cardiol* 2017;**2**:435–441.
 24. Sengupta PP, Kulkarni H, Narula J. Prediction of abnormal myocardial relaxation from signal processed surface ECG. *J Am Coll Cardiol* 2018;**71**:1650–1660.
 25. Churpek MM, Yuen TC, Winslow C, Meltzer DO, Kattan MW, Edelson DP. Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. *Crit Care Med* 2016;**44**:368–374.
 26. Weiss JC, Natarajan S, Peissig PL, McCarty CA, Page D. Machine learning for personalized medicine: predicting primary myocardial infarction from electronic health records. *AI Mag* 2012;**33**:33–45.
 27. Pocock SJ, Ariti CA, McMurray JJV, Maggioni A, Køber L, Squire IB, Swedberg K, Dobson J, Poppe KK, Whalley GA, Doughty RN; on behalf of the Meta-Analysis Global Group in Chronic Heart Failure. Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies. *Eur Heart J* 2013;**34**:1404–1413.
 28. Daimee UA, Moss AJ, Biton Y, Solomon SD, Klein HU, McNitt S, Polonsky B, Zareba W, Goldenberg I, Kutiyifa V. Long-term outcomes with cardiac resynchronization therapy in patients with mild heart failure with moderate renal dysfunction. *Circ Heart Fail* 2015;**8**:725–732.
 29. Cleland JG, Abraham WT, Linde C, Gold MR, Young JB, Claude Daubert J, Sherfese L, Wells GA, Tang AS. An individual patient meta-analysis of five randomized trials assessing the effects of cardiac resynchronization therapy on morbidity and mortality in patients with symptomatic heart failure. *Eur Heart J* 2013;**34**:3547–3556.
 30. Zeitler EP, Friedman DJ, Daubert JP, Al-Khatib SM, Solomon SD, Biton Y, McNitt S, Zareba W, Moss AJ, Kutiyifa V. Multiple comorbidities and response to cardiac resynchronization therapy: MADIT-CRT long-term follow-up. *J Am Coll Cardiol* 2017;**69**:2369–2379.