



Published in final edited form as:

JACC Clin Electrophysiol. 2021 December ; 7(12): 1505–1515. doi:10.1016/j.jacep.2021.06.009.

Using Machine-Learning for Prediction of the Response to Cardiac Resynchronization Therapy: the SMART-AV Study

Stacey J. Howell, MD¹, Tim Stivland, BSME, MBA², Kenneth Stein, MD, FHRS², Kenneth A. Ellenbogen, MD, FHRS³, Larisa G. Tereshchenko, MD, PhD, FHRS, CCDS^{1,4}

¹Knight Cardiovascular Institute, Oregon Health & Science University, Portland, OR;

²Boston Scientific; Department of Medicine, Johns Hopkins School of Medicine, Baltimore, MD.

³Medical College of Virginia/VCU School of Medicine; Department of Medicine, Johns Hopkins School of Medicine, Baltimore, MD.

⁴Cardiovascular Division, Department of Medicine, Johns Hopkins School of Medicine, Baltimore, MD.

Abstract

Objectives—We aimed to apply machine learning (ML) to develop a prediction model for short-term CRT response to identifying CRT candidates for early multidisciplinary CRT-heart failure (HF) care.

Background—Multidisciplinary optimization of cardiac resynchronization therapy (CRT) delivery can improve long-term CRT outcomes but requires substantial staff resources.

Methods—Participants from the SmartDelay Determined AV Optimization trial (n=741; age, 66±11 yrs; 33% female; 100% New York Heart Association HF class III-IV; 100% ejection fraction 35%) were randomly split into training & testing (80%; n=593), and validation (20%; n=148) samples. Baseline clinical, ECG, echocardiographic, biomarker characteristics, and left ventricular (LV) lead position (43 variables) were included in 8 ML models (random forests, convolutional neural network, lasso, adaptive lasso, plugin lasso, elastic net, ridge, and logistic regression). A composite of freedom from death and HF hospitalization and a >15% reduction in LV end-systolic volume index at 6-months post-CRT was the endpoint.

Results—The primary endpoint was met by 337 patients (45.5%). The adaptive lasso model was the most more accurate (AUC 0.759; 95% CI 0.678–0.840), well-calibrated, and parsimonious (19 predictors; nearly half are potentially modifiable). Participants in the 5th quintile as compared to those in the 1st quintile of the prediction model had 14-fold higher odds of composite CRT

Correspondence: Larisa Tereshchenko, 3181 SW Sam Jackson Park Rd; UHN62; Portland, OR, 97239. tereshch@ohsu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclosures: SMART AV trial was sponsored by Boston Scientific.

Clinical trial registration—[ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT00677014) Identifier: NCT00677014

Twitter: #MachineLearning predicts short-term (6-mo) CRT response and thus may help with CRT procedure and early post-CRT care planning. (add Figure 1 and hashtags)

response (OR 14.0; 95%CI 8.0–14.4). The model predicted CRT response with 70% accuracy, 70% sensitivity, and 70% specificity, and should be further validated in prospective studies.

Conclusions—ML predicts short-term CRT response and thus may help with CRT procedure and early post-CRT care planning.

Condensed abstract

We analyzed the large sample (n=741) of cardiac resynchronization therapy (CRT) recipients, participants of the SMART-AV randomized controlled trial. Using machine learning, we developed and validated a parsimonious model that is comprised of routinely available baseline clinical, ECG, and echocardiographic characteristics (19 predictor variables). Participants in the 5th quintile compared to those in the 1st quintile of the prediction model had 14-fold higher odds of composite CRT response. The model outperformed the current guidelines and predicted CRT response with 70% accuracy, 70% sensitivity, and 70% specificity and should be further validated in prospective studies.

Keywords

cardiac resynchronization therapy; machine learning

Introduction

Cardiac resynchronization therapy (CRT) is an established treatment for patients with systolic heart failure (HF) and ventricular dyssynchrony.(1) However, despite proven benefits, nearly a third of CRT recipients are considered to be “non-responders.”(2) CRT-eligible HF patients face high mortality: 50% of them die within 5 years. There has been no improvement in HF prognosis over several decades.(3)

Guided left ventricular (LV) lead placement considering the timing of LV activation and electrical delay(4), together with dynamic atrioventricular (AV) optimization(5), can potentially reduce the CRT non-response rate. Previous analysis of the SMART-AV (SmartDelay Determined AV Optimization: A Comparison to Other AV Delay Methods Used in Cardiac Resynchronization Therapy) study suggested a strategy for using measures of LV electrical delay at implantation to guide LV lead placement.(6) However, a complex interaction between cardiac veins anatomy and cardiomyopathy substrate can make guided LV lead placement procedure technically difficult. Furthermore, observational studies suggested that integrated multidisciplinary care delivered within the first 6 months post-CRT might improve long-term clinical outcomes.(7–9) However, such a multidisciplinary approach requires substantial resources, and its cost-effectiveness has not been evaluated. Prediction of the probability of a short-term CRT response might help with resource allocation and CRT procedure/post-procedure CRT optimization planning. Figure 1 outlines possible pathways for patients with different probabilities of short-term CRT response.

Machine learning (ML) has taken hold in a number of fields to improve risk prediction as compared to traditional methods.(10,11) Several studies have applied ML to address the clinical challenge of CRT patient selection and showed that ML algorithms perform better than guidelines-recommended QRS duration and bundle branch block (BBB) morphology.

(12–15) However, all previous ML-prediction models targeted the long-term (1 year) CRT outcomes, focusing on selecting the “most appropriate” CRT candidate. At present, there is no short-term (6-month) CRT response prediction tool that can be used to plan CRT delivery optimization and early post-CRT care.

We conducted the current study with the goal to use ML to predict short-term (6-month) response to CRT.

Methods

The authors used the deidentified SMART-AV study dataset provided by the executive study committee. The Oregon Health & Science University Institutional Review Board determined the deidentified nature of the dataset. Open-source code for statistical data analysis is provided at <https://github.com/Tereshchenkolab/statistics>. The CRT response prediction calculator is provided at <http://www.ecgpredictscd.org/crt>, and as a supplement.

Study population

The SMART-AV was a randomized, multicenter, single-blinded clinical trial(16,17) that sought to determine whether AV delay optimization would improve CRT response six months post-implant. The trial enrolled New York Heart Association (NYHA) class III-IV HF patients with left ventricular ejection fraction (LVEF) \geq 35% despite optimal medical therapy, and QRS duration \geq 120 ms, in sinus rhythm. HF patients who were in complete heart block, could not tolerate right ventricular (RV) pacing at VVI-40 for up to two weeks, or previously received CRT were excluded. Enrollment was completed from May 2008 through December 2009. In the current study, we excluded participants with missing candidate predictor variables and lost follow-up. Of the 980 randomized SMART-AV participants, 741 CRT recipients were included in this study.

Candidate predictor variables

At the enrollment visit, baseline clinical characteristics data were collected, which included medical history, current cardiovascular evaluation (NYHA class) and medications list, the 6-minute walk test, quality of life (Minnesota Living with Heart Failure Questionnaire), and blood draw for biomarkers.(16,17) We calculated estimated glomerular filtration rate (eGFR) using the chronic kidney disease Epidemiology Collaboration equation (CKD-EPI). (18) LV lead location was selected at the discretion of the implanting physician. Baseline ECG and echocardiogram were recorded post-implant (no biventricular pacing).(16,17) We normalized LV volumes and dimensions by body surface area. Measurement of biomarkers in SMART-AV study has been previously described.(19)

The study endpoint

We defined the primary endpoint as a composite of freedom from death and HF hospitalization and a $>$ 15% reduction(5,6,20,21) in LV end-systolic volume index (LVESVI) at six months of follow-up. LVESVI was the primary endpoint in the SMART-AV trial. (16,17) A single core laboratory performed all echocardiographic measurements in a blinded fashion.

Statistical machine learning analysis

We randomly split the study population into two non-overlapping samples: training&testing (80%; n=593), and validation (20%; n=148). Considering future clinical implementation, we included routinely available predictor variables that describe baseline clinical, ECG, echocardiographic and biomarker characteristics, and LV lead position (Input #1; 43 variables, Table 1).

We fitted eight different models (random forests(22), convolutional neural network (CNN) (23), lasso, adaptive lasso, plugin lasso, elastic net, ridge, and logistic regression).

To train the random forests algorithm, we arranged the data in a randomly sorted order and tuned the number of subtrees and variables to investigate at each split randomly. We calculated both out-of-bag training error (tested against training data subsets that are not included in subtree construction) and a testing error (tested against the testing data) to find the model with the highest testing accuracy. In tuning the random forests algorithm, we observed that both out-of-bag training error and testing error stabilized after 300 iterations at 30–35% (Supplemental Figure 1), and we conservatively chose 500 subtrees. The minimum testing error was observed for 7 variables, and we chose 7 variables to investigate at each split randomly.

We trained the CNN with 20 hidden layers, using 500 iterations with a training factor 2 and 4 normalization parameters. The network was comprised of 3 layers, 64 neurons per layer, and 901 synapse weights.

The lasso family (least absolute shrinkage and selection operator) models employed ten-fold cross-validation in the training&testing sample. In the lasso model, cross-validation selected the tuning parameter λ that minimized the out-of-sample deviance in training&testing sample. The adaptive lasso performed multistep cross-validation, performing the second cross-validation step among the covariates selected in the first cross-validation step. The plugin lasso used partialing-out estimators to determine which covariates belong in the model, achieving an optimal bound on the number of covariates it included.(24,25) The elastic net permitted retention of correlated covariates. In the ridge model, the penalty parameter used squared terms and kept all predictors in the model.

We validated the predictive accuracy of the models by comparing the area under the receiver operator curve (ROC AUC) in the validation sample. To assess calibration, we compared the observed and predicted proportions within the groups formed by the Hosmer-Lemeshow test(26) and used the calibration belt(27) to examine the relationship between out-of-sample estimated probabilities and observed CRT response rates. For the lasso family of models, we also calculated the deviance and deviance ratio (goodness-of-fit).

The selection of “the best” ML model was guided by discrimination (ROC AUC) and calibration of the model in validation sample, and then the most parsimonious model was selected.

We used quintiles of the endpoint’s predicted probability to divide the population into five equally numerous subsets. To illustrate the final selected model’s discrimination capacity,

unadjusted logistic regression (model 1) compared the odds of the endpoint in participants in the 2nd, 3rd, 4th, and 5th quintiles compared to those in the 1st quintiles. To assess how the SMART-AV intervention affected the model's predictability and test for a possible interaction between the endpoint's predicted probability and the intention-to-treat (ITT) SMART-AV intervention, we adjusted the logistic regression for the SMART-AV treatment group and their interaction term.

To test the final prediction model's performance across different patient populations, we calculated ROC AUC in validation sample in men and women, white and non-white, with and without diabetes, age ≥ 65 y and < 65 y, and ITT subgroups.

The final equation that calculates the probability of a 6-month CRT response is provided in the Supplement. We selected the threshold of predictive function corresponding to 70% accuracy, 70% sensitivity, and 70% specificity.

We compared the performance of the selected model to the current 2013 American College of Cardiology Foundation/American Heart Association class I guideline criteria (QRS >150 ms and the presence of LBBB). (28)

In addition, as better CRT response was observed in women,(21) we constructed sex-specific ML models. To explore broader range of sex-specific predictors, we added 10 biomarkers (Input #2; 53 variables): extracellular matrix-metalloproteinases (MMP-2 and MMP-9), soluble interleukin-2 receptor (sIL-2r), glycoprotein 130 (sGP-130), soluble suppressor of tumorigenicity-2 (sST-2), interferon gamma (IFNG), soluble tumor necrosis factor receptor-II (sTNFr-II), tissue inhibitor of extracellular matrix-metalloproteinases (TIMP-1, TIMP-2, TIMP-4). Because these biomarkers are not available in everyday clinical practice, to preserve clinical utility and generalizability of the overall model, we considered two types of input separately.

Statistical analysis was performed using STATA MP 16.1 (StataCorp LP, College Station, TX). P -value < 0.05 was considered statistically significant.

Results

The SMART-AV study population characteristics are shown in Table 1 and have been previously reported elsewhere.(21) The primary endpoint was met by 337 patients (45.5%). Out of 404 participants who failed to respond, 13 died, 75 participants were hospitalized because of HF, and 316 participants failed to achieve a volumetric response. Out of 741 study participants, echocardiographic data was not available for 31 participants. Out of 31 participants with missing echocardiographic data, 13 died, and the other 18 were hospitalized because of HF exacerbation within a 6-month follow-up.

A comparison of the prediction models' performance is shown in Table 2. The CNN demonstrated the highest predictive accuracy in the training&testing sample, with a final error of only 6%. However, the CNN model's calibration was unsatisfactory (Hosmer-Lemeshow test $P<0.0001$; Supplemental Figure 2), and predictive accuracy in the validation sample did not differ from the lasso family of models.

Several models (lasso, adaptive lasso, elastic net, ridge, and logistic regression) demonstrated similar good fit and high predictive accuracy (Table 2), which was significantly higher than for random forests and plugin lasso models, as well as current class I clinical guidelines (AUC 0.639; 95% CI 0.554–0.722), $P<0.0001$. Supplemental Figure 3 shows the cross-validation function and selected λ for each model.

The random forests model reported substantial 26% error in validation sample; it correctly predicted CRT response in only 38 out of 65 individuals (sensitivity 58.5%), and predicted freedom from composite CRT response endpoint in 71 out of 83 participants (specificity 85.5%), having a positive predictive value of 76% and negative predictive value of 72.4%. The single most important predictor was diabetes (Supplemental Figure 4), which, together with demographic characteristics (age, sex, race) and other comorbidities (hypertension, smoking), comprised six of the most important predictors.

Only a few models (logistic regression, adaptive lasso, and plugin lasso) showed satisfactory out-of-sample calibration (Figure 2). However, the plugin lasso model had significantly lower ROC AUC than the adaptive lasso and logistic regression (Table 2). Ultimately, we selected the adaptive lasso model as the most accurate, well-calibrated, and parsimonious (19 predictors listed in Supplemental Table 1).

In the adaptive lasso model, the most important predictors (Central illustration) characterized dyssynchrony (ventricular conduction type, QRS duration), underlying disease substrate (cardiomyopathy type, primary prevention indication), and potentially modifiable characteristics (NT-proBNP, systolic blood pressure), including PR interval. Nonischemic cardiomyopathy, female sex, primary prevention indication, history of valvular heart disease and cancer, higher QRS duration, systolic blood pressure, LVEDVI, and 6-min walk distance, eGFR_{CKD-EPI}, and age were associated with CRT response. Non-LBBB, AV block I-II, and higher NT-proBNP, CRP, PR interval, LVEF, LVESDI, and weight were associated with non-response. Participants in the 5th quintile as compared to those in the 1st quintile had 14-fold higher odds of composite CRT response (Central Illustration). The online calculator is available at <http://www.ecgpredictscd.org/crt>. The final model performance was consistent across the subgroups (Figure 3).

Adjustment for the SMART-AV treatment group attenuated the association of ML-predicted probability of CRT response with the study endpoint (Figure 4A), suggesting that AV-optimization is one of the mechanisms responsible for the composite CRT response outcome. Furthermore, both SmartDelay algorithm-optimized and echo-optimized AV delay hinted a higher probability of the composite CRT response than the fixed AV delay. However, consistently with the reported SMART-AV results, the difference did not reach statistical significance (Figure 4B).

The performance of sex-specific models was poor (Supplemental Table 2). Out of all sex-specific ML models, only the adaptive lasso model had satisfactory calibration both in men and women, which allowed us to compare retained predictors (Figure 5). Male-specific predictors included quality of life, sleep apnea, and biomarkers (MMP-2, sPG-130, sTNFr-

II). Female-specific predictors included height, smoking, biomarker TIMP-2, and treatment (use of ACEI/ARBs, aldosterone antagonists, LV lead location).

Discussion

In this study, using the ML approach, we developed a parsimonious model for predicting short-term CRT response that comprises routinely available baseline clinical, ECG, and echocardiographic characteristics - measures of the disease substrate, dyssynchrony, and comorbidities. Several included predictors could be potentially modifiable. Developed in this study, the CRT response prediction model opens an avenue for a future randomized controlled trial, testing CRT delivery strategy, incorporating targeted lead placement, dynamic AV optimization programming,(5,6) and integrated multidisciplinary care.(7–9) Importantly, identified sex-specific predictors of CRT response provide insight into sex differences in underlying disease substrate.

Our results of sex-specific ML analysis were consistent with the previously developed Biomarker CRT score.(19) A Biomarker CRT Score of 4 included 3 biomarkers identified in this study as important predictors of CRT response in men (MMP-2, sTNFr-II, and CRP). Overall, retained biomarkers reflect kidney failure (MMP-2 and TIMP-2), inflammation (CRP, sTNFr-II), and HF (NT-proBNP). Identified in our study sex-specific predictors of CRT response indicate the importance of integrated multidisciplinary care for men and women in different ways. Treatment of sleep apnea is especially important for men, whereas the use of ACEI/ARB and aldosterone antagonists, weight management, and smoking quit is especially important for women.

It has been previously shown that increasing degrees of interventricular (rather than intraventricular) dyssynchrony is expected to result in improved rates of clinical CRT response.(29) Previous analysis of the SMART-AV study showed that optimally timed AV delay provides an incremental benefit to the substantial interventricular conduction delay(5,6), suggesting that both LV and RV lead placement should target maximizing RV-LV delay.

Pre-procedural planning of LV and RV lead placement maximizing RV-LV delay may involve expensive and time-consuming cardiac imaging. Our risk score can predict the probability of the short-term composite CRT response and, therefore, can help to preserve resources while improving clinical outcomes. Careful pre-procedural planning would be particularly critical for CRT candidates with a moderate or low probability of CRT response.

Notably, both the baseline PR interval and I-II AV block's presence were selected by the adaptive lasso model as essential predictors in the model, indicating the likely benefit of dynamic AV optimization. In this ML model, consistently with previous findings, a longer PR interval indicated a lesser probability of CRT response.(30) However, it is essential to remember that the ML model strives to achieve the best prediction of the outcome but does not answer whether each predictor variable reflects an independent mechanism of outcome. Our study provided additional evidence about the importance of adaptive, optimally timed AV delay for the composite CRT response. Both SmartDelay algorithm-optimized and

echo-optimized AV delay hinted at a higher probability of short-term CRT response than fixed AV delay.

Pre-procedure, our calculator tool could be potentially used for shared decision-making(31) and set-up of management goals (e.g., target weight, target systolic blood pressure, 6-min walk distance, biomarkers level). Such discussion with a CRT candidate might motivate compliance to diet, fluid restriction, and medication adherence.

Furthermore, our prediction model can be used to identify CRT recipients with a low probability of CRT response. They must be referred to a multidisciplinary CRT-HF clinic very early, immediately after CRT implantation.(7–9) Considering the grim prognosis of CRT-eligible HF patients, CRT delivery’s adequacy should be scrutinized, and modifiable predictors of CRT response should be targeted early. Notably, the range of optimization interventions goes beyond our predictors’ list and should be patient-specific.

Consistent with prior studies(6,12–15), we confirmed that ML model performs better than current guidelines. The strength of ML algorithms is the ability to capture complex interactions.(32) Several prior studies have used ML to predict CRT response. Kalscheur et al analyzed 595 COMPANION NYHA III/IV patients,(12) Cikes et al studied 1106 MADIT-CRT NYHA class II patients,(15) Feeny et al evaluated 470 NYHA I-IV patients from an observational cohort, and Hu et al. retrospectively analyzed 990 predominately NYHA II-III patients from a single-center cohort.(33) Of note, all previous studies considered long-term CRT benefits, answering a question of CRT candidate selection. In contrast, our prediction model focuses on a short-term CRT response and can help plan the CRT delivery and early, aggressive optimization strategy.

In this study, the absence of sustained ventricular tachyarrhythmia (primary prevention indication) was an important predictor of CRT response. This finding is consistent with previous studies that showed the antiarrhythmic effect of CRT and reversed electrical remodeling(34), which can be facilitated by the autonomic nervous system response.(35)

A comparison of ML models and selection of the “best” model also deserves discussion. We observed similar accuracy in all but one (plugin lasso) models, leaving seven models for consideration. However, only two of them (logistic regression and adaptive lasso) demonstrated satisfactory calibration. The parsimonious model (adaptive lasso) won because it is more simple (19 versus 43 predictors). The most important predictors in the adaptive lasso model provide a meaningful characterization of the disease substrate and its electrophysiology (a type of cardiomyopathy and conduction abnormality, QRS duration, history of sustained ventricular tachyarrhythmia or cardiac arrest, NT-proBNP and systolic blood pressure), which can guide CRT delivery. Importantly, the model performed equally well in clinically important subgroups.

Strengths and Limitations

SMART-AV is a large multicenter randomized control trial with careful phenotyping that included blinded analysis of echocardiograms and biomarkers in core laboratories, and appropriate follow-up, providing an opportunity to study composite CRT response. A

strength of the present study was the use of a composite endpoint of clinical outcomes (death, HF hospitalization) and volumetric remodeling. Inclusion of participants who died or was hospitalized and thus missed 6-month follow-up echocardiogram, strengthened the study and reduced attrition bias.

Another strength of the study is the definition of volumetric response. A decrease in LVESV better than an increase in LVEF reflects reverse remodeling. LVESV is influenced by fiber shortening and, to a lesser degree, by end-diastolic volume. LVEF is influenced to a greater extent by end-diastolic volume and heart rate and is, therefore, less suitable as a surrogate marker of long-term CRT response.(36) LVESV change is the strongest predictor of mortality among the three measures of LV remodeling (LVEF, LVEDV, LVESV) in the setting of either low LVEF or high LVEDV.(37)

However, the limitations of the study have to be taken into account. The study population was predominantly men, although this is characteristic and similar to other CRT trials. Another common CRT field limitation was selection bias. The study included only participants who had successfully implanted CRT device and excluded those without suitable cardiac veins, whose procedures may have been aborted due to difficult anatomy. We limited candidate predictor variables by currently widely available and did not include novel ECG measures of dyssynchrony that can further improve prediction.(20,38) Baseline predictors were measured only once. It is possible that repeated assessment can improve accuracy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding:

This work was supported in part by HL118277, Medical Research Foundation of Oregon and OHSU President Bridge funding (LGT).

Abbreviations

CRT	Cardiac resynchronization therapy
HF	Heart failure
LV	left ventricular
RV	right ventricular
AV	atrioventricular
SMART-AV	SmartDelay Determined AV Optimization: A Comparison to Other AV Delay Methods Used in Cardiac Resynchronization Therapy study
ML	machine learning

LBBB	left bundle branch block
NYHA	New York Heart Association
LVEF	left ventricular ejection fraction
CKD-EPI	chronic kidney disease Epidemiology Collaboration equation
LVESVI	left ventricular end-systolic volume index
LVEDVI	left ventricular end-diastolic volume index
ROC AUC	area under the receiver operator curve
LVEDSI	left ventricular end-systolic dimension index
LVEDDI	left ventricular end-diastolic dimension index
ITT	intention to treat
CNN	convolutional neural network
CI	confidence interval
SD	standard deviation
IQR	inter-quartile range

References

1. Tracy CM, Epstein AE, Darbar D et al. 2012 ACCF/AHA/HRS Focused Update of the 2008 Guidelines for Device-Based Therapy of Cardiac Rhythm Abnormalities: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Heart Rhythm* 2012;9:1737–53. [PubMed: 22975672]
2. Chatterjee NA, Singh JP. Cardiac resynchronization therapy: past, present, and future. *Heart Fail Clin* 2015;11:287–303. [PubMed: 25834976]
3. Conrad N, Judge A, Canoy D et al. Temporal Trends and Patterns in Mortality After Incident Heart Failure: A Longitudinal Analysis of 86 000 Individuals. *JAMA Cardiology* 2019;4:1102–1111. [PubMed: 31479100]
4. Singh JP, Fan D, Heist EK et al. Left ventricular lead electrical delay predicts response to cardiac resynchronization therapy. *Heart Rhythm* 2006;3:1285–1292. [PubMed: 17074633]
5. Gold MR, Yu Y, Singh JP et al. Effect of Interventricular Electrical Delay on Atrioventricular Optimization for Cardiac Resynchronization Therapy. *Circ Arrhythm Electrophysiol* 2018;11:e006055. [PubMed: 30354310]
6. Field ME, Yu N, Wold N, Gold MR. Comparison of measures of ventricular delay on cardiac resynchronization therapy response. *Heart Rhythm* 2020;17:615–620. [PubMed: 31765805]
7. Altman RK, Parks KA, Schlett CL et al. Multidisciplinary care of patients receiving cardiac resynchronization therapy is associated with improved clinical outcomes. *European Heart Journal* 2012;33:2181–2188. [PubMed: 22613342]
8. Gorodeski EZ, Magnelli-Reyes C, Moennich LA, Grimaldi A, Rickard J. Cardiac resynchronization therapy-heart failure (CRT-HF) clinic: A novel model of care. *PLoS One* 2019;14:e0222610. [PubMed: 31536565]
9. Mullens W, Grimm RA, Verga T et al. Insights from a cardiac resynchronization optimization clinic as part of a heart failure disease management program. *J Am Coll Cardiol* 2009;53:765–73. [PubMed: 19245967]

10. Deo RC. Machine Learning in Medicine. *Circulation* 2015;132:1920–30. [PubMed: 26572668]
11. Haq KT, Howell SJ, Tereshchenko LG. Applying Artificial Intelligence to ECG Analysis: Promise of a Better Future. *Circ Arrhythm Electrophysiol* 2020;13:e009111. [PubMed: 32809878]
12. Kalscheur MM, Kipp RT, Tattersall MC et al. Machine Learning Algorithm Predicts Cardiac Resynchronization Therapy Outcomes: Lessons From the COMPANION Trial. *Circ Arrhythm Electrophysiol* 2018;11:e005499. [PubMed: 29326129]
13. Tokodi M, Schwertner WR, Kovács A et al. Machine learning-based mortality prediction of patients undergoing cardiac resynchronization therapy: the SEMMELWEIS-CRT score. *European Heart Journal* 2020;41:1747–1756. [PubMed: 31923316]
14. Feeny AK, Rickard J, Patel D et al. Machine Learning Prediction of Response to Cardiac Resynchronization Therapy: Improvement Versus Current Guidelines. *Circ Arrhythm Electrophysiol* 2019;12:e007316. [PubMed: 31216884]
15. Cikes M, Sanchez-Martinez S, Claggett B et al. Machine learning-based phenotyping in heart failure to identify responders to cardiac resynchronization therapy. *European journal of heart failure* 2019;21:74–85. [PubMed: 30328654]
16. Stein KM, Ellenbogen KA, Gold MR et al. SmartDelay determined AV optimization: a comparison of AV delay methods used in cardiac resynchronization therapy (SMART-AV): rationale and design. *Pacing Clin Electrophysiol* 2010;33:54–63. [PubMed: 19821938]
17. Ellenbogen KA, Gold MR, Meyer TE et al. Primary results from the SmartDelay determined AV optimization: a comparison to other AV delay methods used in cardiac resynchronization therapy (SMART-AV) trial: a randomized trial comparing empirical, echocardiography-guided, and algorithmic atrioventricular delay programming in cardiac resynchronization therapy. *Circulation* 2010;122:2660–8. [PubMed: 21098426]
18. Matsushita K, Selvin E, Bash LD, Astor BC, Coresh J. Risk implications of the new CKD Epidemiology Collaboration (CKD-EPI) equation compared with the MDRD Study equation for estimated GFR: the Atherosclerosis Risk in Communities (ARIC) Study. *Am J Kidney Dis* 2010;55:648–659. [PubMed: 20189275]
19. Spinale FG, Meyer TE, Stolen CM et al. Development of a biomarker panel to predict cardiac resynchronization therapy response: Results from the SMART-AV trial. *Heart Rhythm* 2019;16:743–753. [PubMed: 30476543]
20. Tereshchenko LG, Cheng A, Park J et al. Novel measure of electrical dyssynchrony predicts response in cardiac resynchronization therapy: Results from the SMART-AV Trial. *Heart Rhythm* 2015;12:2402–10. [PubMed: 26272523]
21. Cheng A, Gold MR, Waggoner AD et al. Potential mechanisms underlying the effect of gender on response to cardiac resynchronization therapy: insights from the SMART-AV multicenter trial. *Heart Rhythm* 2012;9:736–41. [PubMed: 22182496]
22. Schonlau M, Zou RY. The random forest algorithm for statistical learning. *The Stata Journal* 2020;20:3–29.
23. Doherr T BRAIN: Stata module to provide neural network. 1 ed. Boston: Boston College Department of Economics, 2018:Boston College Department of Economics. 2018. <https://ideas.repec.org/c/boc/bocode/s458566.html>.
24. Belloni A, Chen D, Chernozhukov V, Hansen C. Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain. *Econometrica* 2012;80:2369–2429.
25. StataCorp. Stata 16 Base Reference Manual. College Station, TX: Stata Press. 2019.
26. Lemeshow S, Hosmer DW Jr, A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982;115:92–106. [PubMed: 7055134]
27. Nattino G, Lemeshow S, Phillips G, Finazzi S, Bertolini G. Assessing the calibration of dichotomous outcome models with the calibration belt. *Stata Journal* 2017;17:1003–1014.
28. Yancy CW, Jessup M, Bozkurt B et al. 2013 ACCF/AHA guideline for the management of heart failure: executive summary: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation* 2013;128:1810–52. [PubMed: 23741057]

29. Waks JW, Perez-Alday EA, Tereshchenko LG. Understanding Mechanisms of Cardiac Resynchronization Therapy Response to Improve Patient Selection and Outcomes. *Circ Arrhythm Electrophysiol* 2018;11:e006290. [PubMed: 29654133]
30. Januskiewicz Ł, Vegh E, Borgquist R et al. Prognostic implication of baseline PR interval in cardiac resynchronization therapy recipients. *Heart Rhythm* 2015;12:2256–62. [PubMed: 26066291]
31. Seaburg L, Hess EP, Coylewright M, Ting HH, McLeod CJ, Montori VM. Shared Decision Making in Atrial Fibrillation. *Circulation* 2014;129:704–710. [PubMed: 24515956]
32. Rahman QA, Tereshchenko LG, Kongkatong M, Abraham T, Abraham MR, Shatkay H. Utilizing ECG-Based Heartbeat Classification for Hypertrophic Cardiomyopathy Identification. *IEEE Trans Nanobioscience* 2015;14:505–12. [PubMed: 25915962]
33. Hu SY, Santus E, Forsyth AW et al. Can machine learning improve patient selection for cardiac resynchronization therapy? *PloS one* 2019;14:e0222397. [PubMed: 31581234]
34. Tereshchenko LG, Henrikson CA, Stempniewicz P, Han L, Berger RD. Antiarrhythmic effect of reverse electrical remodeling associated with cardiac resynchronization therapy. *Pacing Clin Electrophysiol* 2011;34:357–64. [PubMed: 21091740]
35. Tereshchenko LG, Henrikson CA, Berger RD. Strong coherence between heart rate variability and intracardiac repolarization lability during biventricular pacing is associated with reverse electrical remodeling of the native conduction and improved outcome. *J Electrocardiol* 2011;44:713–7. [PubMed: 21944164]
36. Cohn JN, Ferrari R, Sharpe N. Cardiac remodeling—concepts and clinical implications: a consensus paper from an international forum on cardiac remodeling. *Journal of the American College of Cardiology* 2000;35:569–582. [PubMed: 10716457]
37. White HD, Norris RM, Brown MA, Brandt PW, Whitlock RM, Wild CJ. Left ventricular end-systolic volume as the major determinant of survival after recovery from myocardial infarction. *Circulation* 1987;76:44–51. [PubMed: 3594774]
38. Jacobsson J, Borgquist R, Reitan C et al. Usefulness of the Sum Absolute QRST Integral to Predict Outcomes in Patients Receiving Cardiac Resynchronization Therapy. *Am J Cardiol* 2016;118:389–95. [PubMed: 27265674]

Perspectives

Competency in medical knowledge:

Machine learning could improve patient selection for CRT therapy beyond current guidelines. Parsimonious model for short-term (6-months) CRT response prediction (comprised of routinely available baseline clinical, ECG, and echocardiographic characteristics) predicts CRT response with 70% accuracy, 70% sensitivity, and 70% specificity. Patients in the 5th versus the 1st quintile of the prediction model have 14-fold higher odds of composite CRT response.

Translational outlook:

Future randomized controlled trials are needed to test the hypothesis that pre-procedure planning and aggressive early (within first 6 months) management of modifiable risk factors and CRT delivery optimization can improve outcomes in CRT recipients with predicted moderate or poor CRT response.

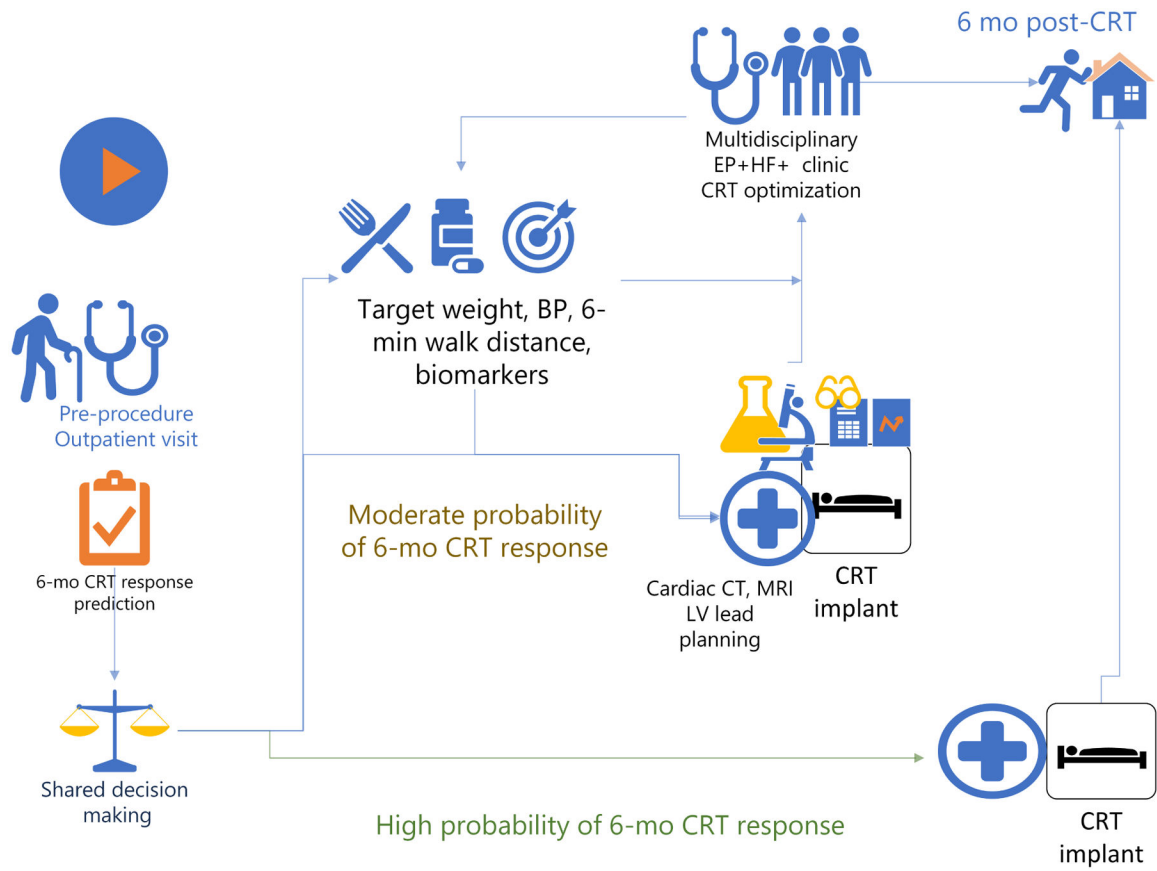


Figure 1. Patient management scenarios.

Possible patient management scenarios based on a pre-procedural prediction of short-term (6-months) CRT response.

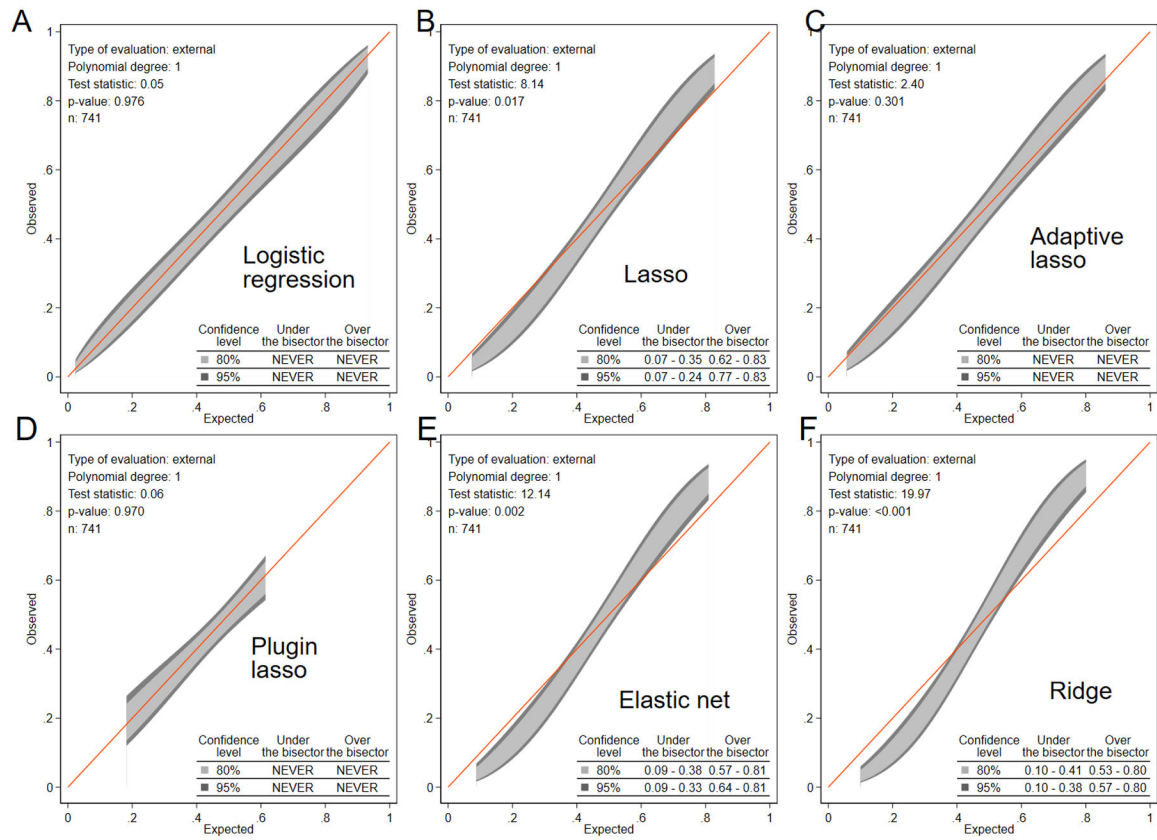


Figure 2. Calibration.

The calibration belt with 80% and 95% CI on the external sample shows the observed and predicted CRT response proportions in (A) logistic regression, (B) lasso, (C) adaptive lasso, (D) plugin lasso, (E) elastic net, and (F) ridge models for all participants. The appropriately calibrated model fits the red diagonal line, and neither 80% nor 95% CI is entirely under or above the red line (never crosses the red line).

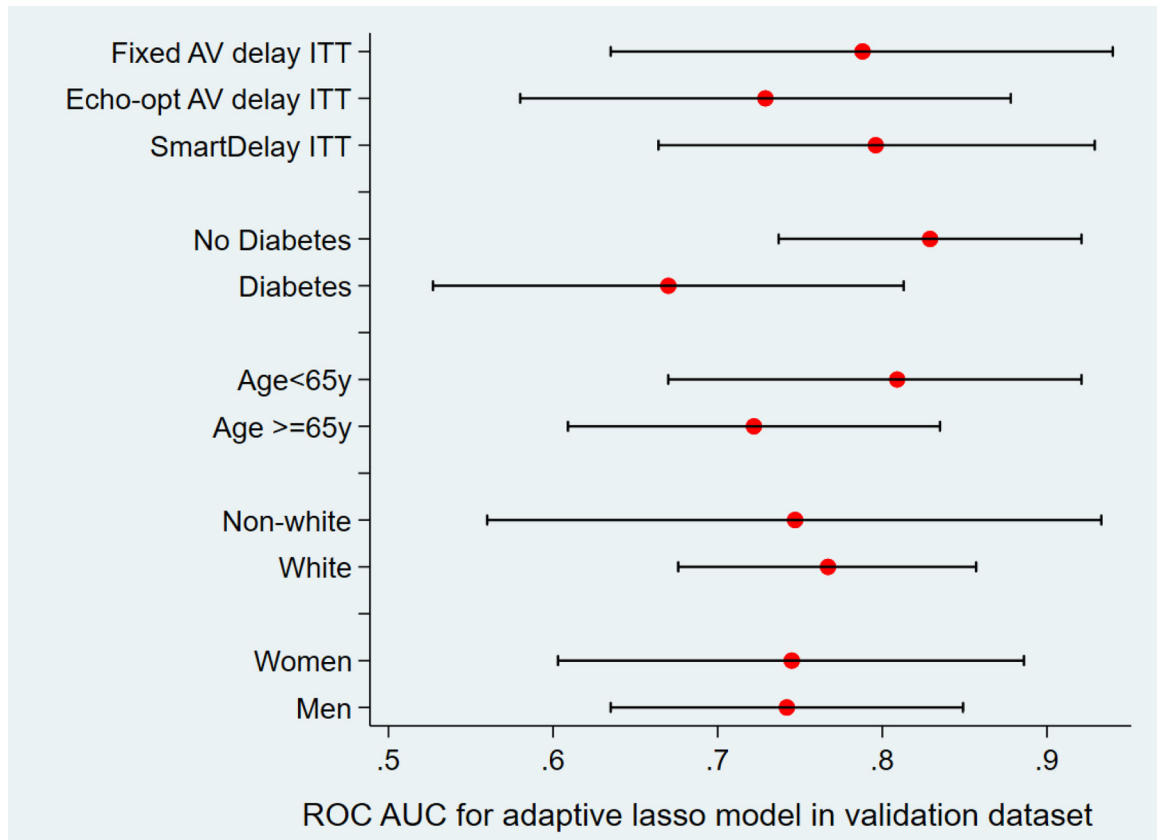


Figure 3. The final adaptive lasso model performance in subgroups.
 ROC AUC in the validation sample, in subgroups.

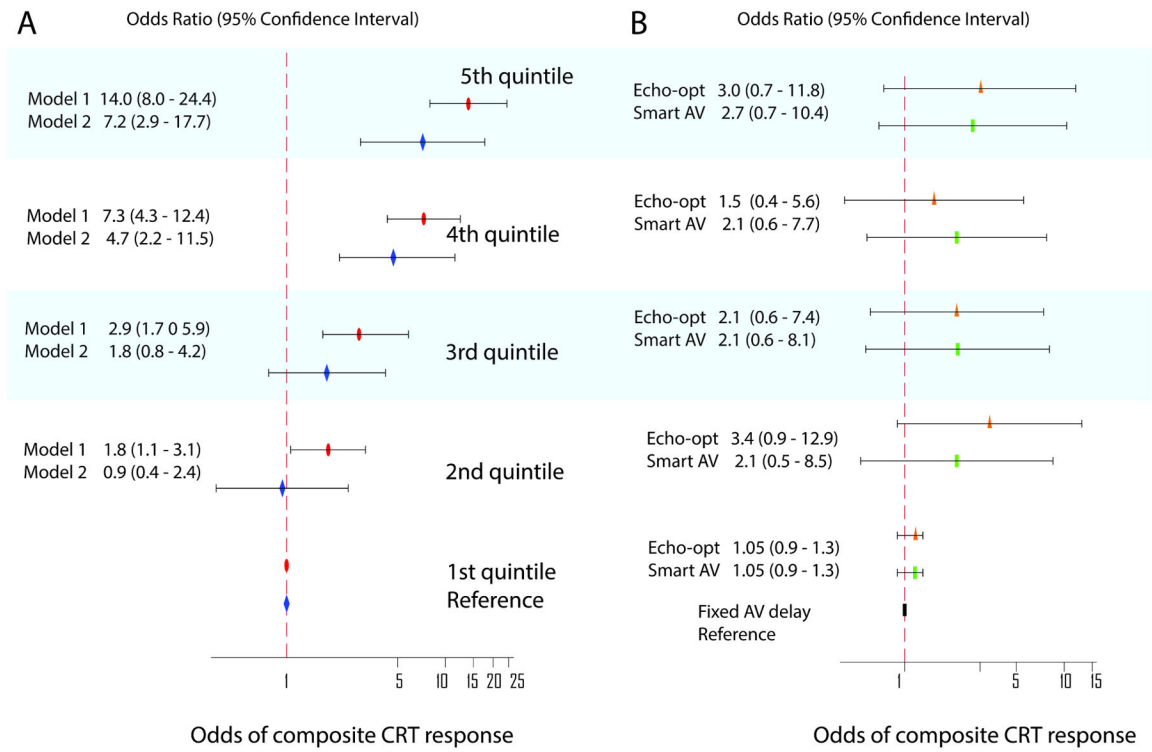


Figure 4. Odds of composite CRT response. Probabilities of composite CRT response by quintiles of the adaptive lasso model. The 1st quintile is a reference. **A.** Unadjusted (model 1, red ovals) and adjusted by the ITT AV-delay optimization assignment (model 2, blue diamonds). **B.** Relative odds of composite CRT response in SMART-AV participants randomized to echo-optimized AV delay (orange triangles) and Smart-AV-delay algorithm (green rectangles). Fixed AV delay group was a reference. Black lines correspond to 95% CI bounds.

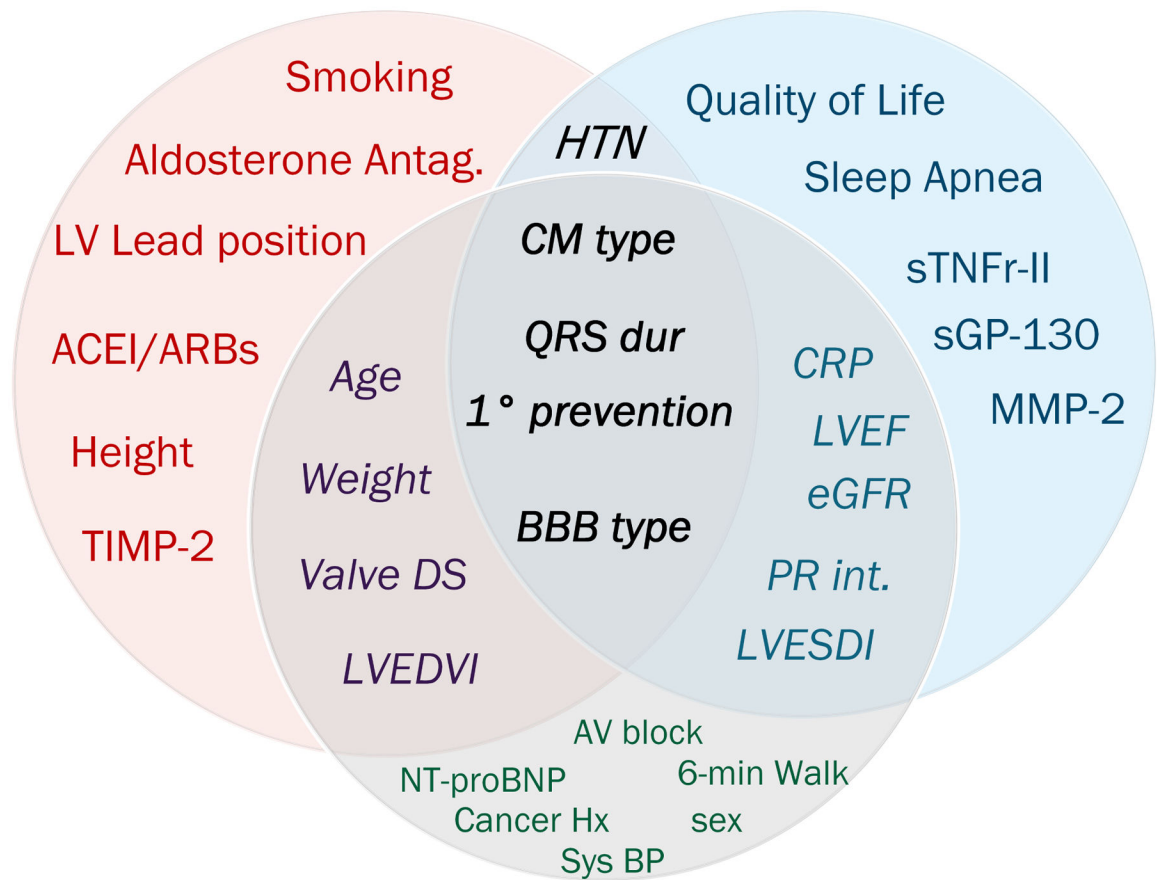
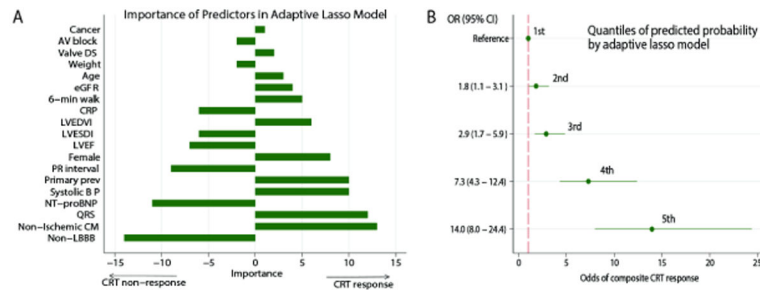


Figure 5. Sex-specific predictors of CRT response.
 Venn diagram shows sex-specific predictors of CRT response in women (red), men (blue), and all participants (gray), and those predictors that overlap.



Central illustration. Prediction of short-term CRT response.

A. Importance of the selected predictors in the adaptive lasso model. The most important predictors were added to the model early. **B.** Probabilities of composite CRT response by quintiles of the adaptive lasso model. The odds of the endpoint in participants in the 2nd, 3rd, 4th, and 5th quintiles of CRT-response-predicted probability are compared to those in the 1st quintile.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Baseline Clinical Characteristics in study participants

Characteristics	All (n=741)	Training (n=593)	Validation (n=148)
The main input #1 included 43 variables			
Age(SD), y	66.0(11.0)	66.0(11.0)	65.8(11.0)
Female, n(%)	241(32.5)	293(32.6)	48(32.4)
White, n(%)	575(77.6)	465(78.4)	110(74.3)
LVEF(SD), %	27.5(8.7)	27.6(8.6)	27.4(9.2)
Weight(SD), kg	87.4(20.8)	87.4(20.4)	87.2(22.3)
Height(SD), cm	171.6(10.3)	171.8(10.3)	170.9(10.3)
Body mass index (SD), kg/m ²	29.6(6.2)	29.5(6.2)	29.8(6.5)
BP systolic(SD), mmHg	124.5(20.9)	124.8(21.0)	123.5(20.8)
BP diastolic(SD), mmHg	71.4(12.7)	72.0(12.6)	68.6(12.7)
Ischemic cardiomyopathy Hx, n(%)	426(57.5)	343(57.8)	83(56.1)
Primary prevention, n(%)	589(79.5)	474(79.9)	115(77.0)
Smoking Hx(current or former), n(%)	461(62.2)	380(64.1)	81(54.7)
Hypertension Hx, n(%)	528(71.3)	434(73.2)	94(63.5)
Diabetes Hx, n(%)	289(39.0)	219(36.9)	70(47.3)
Revascularization Hx, n(%)	380(51.3)	304(51.3)	76(51.4)
Autoimmune disease Hx, n(%)	19.0(2.6)	15(2.5)	4(2.7)
Sleep apnea Hx, n(%)	89(12.0)	66(11.1)	23(15.5)
Cancer Hx, n(%)	67(9.0)	53(8.9)	14(9.5)
Renal disease Hx, n(%)	119(16.1)	90(15.2)	29(19.6)
COPD Hx, n(%)	109(14.7)	89(15.0)	20(13.5)
Valve disease Hx, n(%)	40(5.4)	31(5.2)	9(6.1)
Pacemaker implant Hx, n(%)	15(2.0)	12(2.0)	3(2.0)
AV block I-II, n(%)	138(18.6)	108(18.2)	30(20.3)
PR interval(SD), ms	198.2(50.4)	197.0(50.4)	203.1(50.9)
Heart rate(SD), bpm	71.3(12.5)	71.3(12.8)	71.1(11.3)
QRS duration(SD), ms	151.8(19.9)	151.3(19.3)	153.7(22.2)
Conduction disease:LBBB, n(%)	552(74.5)	443(74.7)	109(73.7)
RBBB	81(10.9)	62(10.5)	19(12.8)
IVCD	86(11.6)	70(11.8)	16(10.8)
RBBB+left hemiblock	22(3.0)	18(3.0)	4(2.7)
NYHA class II, n (%)	21(2.8)	18(3.0)	2(1.4)
III	698(94.2)	560(94.4)	138(93.2)
IV	22 (3.0)	15(2.5)	7(4.7)
6-minute walk(SD), m	268.2(124.7)	269.3(122.7)	263.8(132.8)
Quality of life(SD), points	47.2(25.0)	46.9(24.9)	48.1(25.6)
Potassium(SD), mmol/L	4.3(0.5)	4.3(0.5)	4.3(0.5)
Sodium(SD), mmol/L	138.7(3.1)	138(3.3)	138(2.8)
C-reactive protein(SD), ng/mL	6,438(4,425)	6,407(4,409)	6,559(4,500)

Characteristics	All (n=741)	Training (n=593)	Validation (n=148)
NT-proBNP median(IQR), pmol/L	1,691(863–3,952)	1,656(853–3,952)	1,895(889–3,948)
eGFR _{CKD-EPI} (SD), mL/min/1.73 m ²	63.6(22.8)	63.9(22.9)	62.5(22.4)
Use of ACEI/ARB, n (%)	485(65.5)	398(67.1)	87(58.8)
Use of beta blocker, n(%)	681(91.9)	556(93.8)	125(84.5)
Use of aldosterone antagonist, n(%)	262(35.4)	208(35.1)	54(36.5)
LV end systolic volume index (SD), mL/m ²	64.7(29.8)	64.4(29.6)	66.2(29.7)
LV end diastolic volume index (SD), mL/m ²	87.0(32.0)	86.6(32.1)	88.6(31.7)
LV end systolic dimension index (SD), cm/m ²	2.8(0.5)	2.8(0.5)	2.8(0.5)
LV end diastolic dimension index (SD), cm/m ²	3.2(0.5)	3.2(0.5)	3.2(0.5)
Lead location Apical n(%)	98(13.2)	82(13.8)	16(10.8)
Basal	47(6.3)	35(5.9)	12(8.1)
Mid	596(80.4)	476(80.3)	120(81.1)
Additional 10 biomarkers added in the input #2 (43+10=53 variables)			
MMP-2 median (IQR), ng/mL	733(526–1093)	725(526–1077)	817(525–1197)
MMP-9 median (IQR), ng/mL	107(68–172)	105(66–167)	115(74–185)
sGP-130 median (IQR), ng/mL	196(154–243)	195(154–243)	200(154–250)
sIL-2r median (IQR), ng/mL	1.0(0.7–1.4)	1.0(0.7–1.4)	1.0(0.7–1.5)
sTNFr-II median (IQR), ng/mL	7.6(5.3–10.8)	7.5(5.4–10.5)	8.0(5.1–11.2)
IFNG median (IQR), pg/mL	2.9(2.6–3.2)	2.9(2.6–3.2)	2.9(2.7–3.3)
sST-2 median (IQR), ng/mL	28.3(20.1–41.8)	28.0(20.0–40.8)	30.0(20.8–46.4)
TIMP-1 median (IQR), ng/mL	122(90–175)	123(91–173)	120(89–184)
TIMP-2 median (IQR), ng/mL	102(87–122)	101(87–121)	104(87–125)
TIMP-4 median (IQR), ng/mL	2.3(1.5–3.2)	2.3(1.6–3.1)	2.3(1.5–3.6)

Table 2.

Development and validation of composite CRT response prediction tool

Model	Training & testing sample (All N=593)					Validation sample (All N=148)				
	Deviance	Deviance ratio	Number of predictors	ROC AUC (95%CI)	P-value	Deviance	Deviance ratio	Number of predictors	ROC AUC (95%CI)	P-value versus #
Class I guidelines	n/a	n/a	2	0.644(0.602–0.685)	<0.0001	n/a	n/a	2	0.639(0.554–0.722)	0.0004
Ridge	1.201	0.129	43	0.753(0.714–0.792)	0.746	1.164	0.151	43	0.778(0.699–0.857)	0.161
Elastic net	1.196	0.133	30	0.751(0.711–0.790)	0.968	1.163	0.152	30	0.769(0.688–0.849)	0.258
Lasso	1.187	0.140	29	0.752(0.713–0.792)	0.673	1.155	0.158	29	0.770(0.690–0.850)	0.123
Adaptive lasso	1.184	0.142	19	0.751(0.712–0.790)	Reference	1.169	0.148	19	0.759(0.678–0.840)	Reference
Logistic regression	1.147	0.168	43	0.768(0.730–0.805)	0.040	1.135	0.172	43	0.774(0.697–0.851)	0.398
CNN	-	-	43	0.979(0.966–0.993)	<0.0001	-	-	43	0.759(0.682–0.837)	1.000
Random forests	-	-	43	0.642(0.600–0.683)	<0.0001	-	-	43	0.720(0.649–0.791)	0.228
Plugin lasso	1.354	0.019	2	0.655(0.613–0.696)	<0.0001	1.349	0.016	2	0.667(0.582–0.751)	0.003

All coefficients are penalized. The smaller deviance & the larger deviance ratio—the better model. P-value is for difference with adaptive lasso model.