Title

Heart failure-type symptom scores in chronic kidney disease: The importance of body mass index

Journal

International Journal of Obesity, 46(10)

ISSN

0307-0565

Authors

Walther, Carl P
Benoit, Julia S
Gregg, L Parker
et al.

Publication Date

2022-10-01

DOI

10.1038/s41366-022-01208-x

Peer reviewed

HHS Public Access

Author manuscript

Int J Obes (Lond). Author manuscript; available in PMC 2023 April 01.

Published in final edited form as:

Int J Obes (Lond). 2022 October ; 46(10): 1910–1917. doi:10.1038/s41366-022-01208-x.

Heart failure-type symptom scores in chronic kidney disease: the importance of body mass index

Carl P. Walther, MD, MS[1], **Julia S. Benoit, PhD**[2], **L. Parker Gregg, MD, MS**[1,3], **Nisha Bansal, MD, MAS**[4], **Vijay Nambi, MD, PhD**[5,6], **Harold I. Feldman, MD, MSCE**[7], **Michael G. Shlipak, MD, MPH**[8,9], **Sankar D. Navaneethan, MD, MS, MPH**[1,3,10],

CRIC Study Investigators

[1]Selzman Institute for Kidney Health, Section of Nephrology, Baylor College of Medicine, Houston, TX

[2]Texas Institute for Measurement, Evaluation, and Statistics, University of Houston, Houston, TX

[3]Section of Nephrology, Michael E. DeBakey Veterans Affairs Medical Center, Houston, TX

[4]Kidney Research Institute and Division of Nephrology, University of Washington, Seattle, WA

[5]Section of Cardiology, Michael E. DeBakey Veterans Affairs Medical Center, Houston, TX

[6]Section of Cardiovascular Research, Baylor College of Medicine, Houston, TX

[7]Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, PA

[8]Kidney Health Research Collaborative, Department of Medicine, University of California San Francisco, San Francisco, CA

[9]General Internal Medicine Division, Medical Service, San Francisco Veterans Affairs Health Care System, San Francisco, CA

[10]Institute of Clinical and Translational Research, Baylor College of Medicine, Houston, TX

Abstract

Corresponding author: Carl P. Walther, MD, MS, One Baylor Plaza, Houston, TX 77030, 713-798-8350, carl.walther@bcm.edu.

Authors' Contributions

C.P.W. and S.D.N. contributed to the research idea, study design, and data acquisition. C.P.W., J.S.B., L.P.G., N.B., V.N., H.I.F., M.G.S., and S.D.N. contributed to the data analysis/interpretation. C.P.W. performed the statistical analysis. Each author contributed important intellectual content during manuscript drafting or revision and accepts accountability for the overall work.

Disclosures

Dr. Walther reports consulting fees from GlaxoSmithKline. Dr. Navaneethan reports personal fees from Bayer, personal fees from Boehringer-Ingelheim, personal fees from REATA, personal fees from Tricida, and grants from Keryx, outside the submitted work. Dr Nambi has a provisional patent along with Roche and Baylor College of Medicine for use of biomarkers in prediction of heart failure risk and was the site PI for studies sponsored by Merck and Amgen. Dr. Shlipak has reported consulting fees from Cricket Health and Intercept Pharmaceuticals.

Relationship with Industry

No relationships with industry exist related to the submitted work.

Disclaimer

The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or Veterans Administration.

CRIC Study Investigators

Lawrence J. Appel, MD, MPH, Jing Chen, MD, MMSc, MSc, Alan S. Go, MD, James P. Lash, MD, Robert G. Nelson, MD, PhD, MS, Mahboob Rahman, MD, Panduranga S. Rao, MD, Vallabh O Shah, PhD, MS, Raymond R. Townsend, MD, Mark L. Unruh, MD, MS

Objectives—This analysis sought to determine factors (including adiposity-related factors) most associated with HF-type symptoms (fatigue, shortness of breath, and edema) in adults with chronic kidney disease (CKD).

Background—Symptom burden impairs quality of life in CKD, especially symptoms that overlap with HF. These symptoms are common regardless of clinical HF diagnosis, and may be affected by subtle cardiac dysfunction, kidney dysfunction, and other factors. We used machine learning to investigate cross-sectional relationships of clinical variables with symptom scores in a CKD cohort.

Methods—Participants in the Chronic Renal Insufficiency Cohort (CRIC) with a baseline modified Kansas City Cardiomyopathy Questionnaire (KCCQ) score were included, regardless of prior HF diagnosis. The primary outcome was Overall Summary Score as a continuous measure. Predictors were 99 clinical variables representing demographic, cardiac, kidney and other health dimensions. A correlation filter was applied. Random forest regression models were fitted. Variable importance scores and adjusted predicted outcomes are presented.

Results—The cohort included 3,426 individuals, 10.3% with prior HF diagnosis. BMI was the most important factor, with BMI 24.3 kg/m^2 associated with the least symptoms. Symptoms worsened with higher or lower BMIs, with a potentially clinically relevant 5 point score decline at 35.7 kg/m^2 and a 1-point decline at the threshold for low BMI, 18.5 kg/m^2. The most important cardiac and kidney factors were heart rate and eGFR, the 4th and 5th most important variables, respectively. Results were similar for secondary analyses.

Conclusions—In a CKD cohort, BMI was the most important feature for explaining HF-type symptoms regardless of clinical HF diagnosis, identifying an important focus for symptom directed investigations.

Keywords

CKD; body mass index; heart failure; symptoms; machine learning

Introduction

Symptom burden impairs quality of life for persons with chronic kidney disease (CKD), even in the early stages, and identification and amelioration of symptoms is a crucial goal.(1) However, the large burden of comorbidities and complications in CKD make investigating relationships of risk factors with symptoms in individuals challenging, complicating management and clinical investigation.

Fatigue, shortness of breath, and edema—referred to herein as HF-type symptoms—are common in CKD populations, and can be caused by kidney, cardiovascular, and other processes. Fatigue is the most common symptom in CKD, affecting most individuals.(1, 2) Shortness of breath has been estimated to affect 28% of persons with CKD.(1, 3) Peripheral edema is a bothersome symptom common to more advanced or proteinuric CKD and HF. The complex comorbidities often present in persons with CKD leads to uncertainty about the causes of these HF-type symptoms.(4) Improved understanding of these relationships may

be an important step towards developing effective clinical interventions to reduce symptom burden.

Especially relevant to HF-type symptoms, cardiovascular comorbidities and overt and subtle cardiac abnormalities are prevalent even in early stage CKD, and structural abnormalities progress in tandem with CKD.(5, 6) Cardiac hypertrophy, which develops early and underlies much of the diastolic dysfunction observed in CKD, can manifest with HF symptoms.(7, 8)

To assess HF-type symptoms and their impact on functioning and quality of life —paraphrased as HF-type symptoms/health status—the Kansas City Cardiomyopathy Questionnaire (KCCQ) has been developed and validated for use in HF. It has subsequently been approved by the US Food and Drug Administration (FDA) as a Clinical Outcome Assessment for clinical trials in HF.(9) In CKD populations without diagnosed HF, scores from a modified KCCQ instrument in the Chronic Renal Insufficiency Cohort (CRIC) were associated with higher risk of incident HF hospitalization, independent of kidney function and other HF risk factors.(10, 11) A previous cross-sectional analysis of CRIC data used multivariable logistic regression with stepwise selection to investigate how a set of clinical variables related to a dichotomous outcome of significant HF-type symptoms. BMI (along with demographics, smoking status, asthma, coronary artery disease, diabetes, and peripheral arterial disease) was independently associated with risk of low KCCQ score, with a BMI >35 kg/m^2 associated with approximately 4-fold higher odds of having significant symptoms compared to a BMI <27 kg/m^2.(10)

The complexity of the CKD disease process means that numerous pathways have the potential to influence HF-type symptoms, most prominently, cardiovascular health, kidney health, and anemia. To learn more about how diverse health processes relate to HF-type symptoms in CKD, simultaneous analysis of comprehensive health metrics in a highly measured cohort may enable novel insight. To perform these analyses, where parameterizations of relationships are unknown, interactions are unknown and likely complex, and multiple related variables may provide complementary information yet be collinear, certain machine learning algorithms have demonstrated success.(12) Machine learning methods are available that can implicitly handle uncertain complex relationships and enable investigation in a data-driven, interpretable way (including ranking of variables in terms of importance), and these techniques have revealed novel insights in other complex areas in medicine.(13, 14)

To study the relationships of kidney, cardiac, and other clinical parameters with subjective HF-type symptom burden in CKD (a patient-centered metric) in a data-driven way without imposing assumptions about presence or lack of interactions, specific relationships, and parameterizations, we conducted a cross sectional analysis in a CKD cohort using the machine learning technique random forests.

Materials and Methods

Participants

We included persons enrolled in the CRIC study during Phase 1 & 2: adults 21–74 years old with CKD by eGFR thresholds (eGFR 20–70 ml/min/1.73m^2 for ages 21–44 years, 20–60 ml/min/1.73m^2 for 45–64 years, and 20–50 ml/min/1.73m^2 for 65–74 years) at seven clinical sites, 2003–2007, who attended the Year 1 visit (n=3520).(15, 16) Exclusion criteria were missing analysis baseline (CRIC Year 1) Overall Summary Score (n=30) and kidney failure with replacement therapy at baseline (n=64), resulting in 3,426 persons. Participants with and without prior clinical HF diagnosis were included. Sensitivity analyses were performed separately in those with and without prior HF diagnosis. Baylor College of Medicine's Institutional Review Board determined the analyses of de-identified data to not be human subjects research.

Outcomes

The primary outcome was the cross-sectional KCCQ Overall Summary Score as a continuous variable. A 5 point change is considered significant, and scores <75 are considered to reflect clinically important symptoms.(9, 11, 17, 18) The KCCQ is a 23-item self-administered questionnaire that quantifies physical limitations, symptom stability, self-efficacy, and social limitations related to dyspnea, fatigue, and edema.(19) The KCCQ was designed and validated to assess health status in persons with HF, and the FDA has determined that components of the KCCQ are appropriate clinical outcome assessments in HF.(20, 21) For application to persons with CKD regardless of HF diagnosis, the KCCQ administered in CRIC had minor modifications to remove references to HF diagnosis; scoring was not changed.(10) Secondary outcomes were the other KCCQ summary score (Clinical Summary) and 6 domain scores (Symptoms, Physical Function, Quality of Life, Social Limitation, Symptom Stability, and Self-Efficacy), treated as continuous variables. A sensitivity analysis was performed with dichotomized Overall Summary Score as the outcome (≥75 vs. <75), as this threshold has been used to distinguish clinically important symptoms.(9)

Predictors

Predictors were clinical and demographic variables related to kidney, cardiovascular, and general health status and functioning, in addition to social factors, that were available at the analysis baseline and that were missing for <20% of participants. This resulted in 99 variables categorized into demographics, social factors, comorbidities, medications, vital signs, laboratory measures, echocardiographic measures, and ECG measures (Table 1). To allow meaningful assessment of variable importance scores by reducing variable dilution, predictors were winnowed using an unsupervised correlation filter.(22) Sensitivity analyses were performed using different predictor sets: all predictors, a narrow subset of predictors (10) selected based on clinical reasoning of relevance to HF-type symptoms, and these 10 selected predictors plus N-terminal pro-brain natriuretic peptide (NT-proBNP) level, which was only available from CRIC Year 0 (one year prior to the analysis baseline). Finally, a post-hoc analysis was run which replaced BMI with waist circumference among the main set of predictor variables.

Statistical Methods

Analysis steps were 1) predictor selection, 2) missing value imputation, 3) random forest modeling, and 4) generation of interpretable summaries showing how clinical variables relate to symptom scores, including model performance metrics, variable importance scores, and partial dependence plots (Supplemental Figure 1). Partial dependence plots show the average predictions from the model when the variable of interest alone is varied for all participants (the marginal effect of changing the variable of interest).

Random forests is a non-parametric technique widely used in machine learning for its versatility and robustness, and can be used for regression or classification (Supplemental Figure 2).(23, 24) Although most widely known for prediction, random forests are widely used for descriptive learning in studying complex variable relationships in biomedical and other fields, as we use it.(12, 13, 25, 26) Random forests use sets of decision trees generated by recursive sampling of bootstrapped samples of data, with the ultimate goal of minimizing variance through averaging many noisy but unbiased models.(23, 27) Individual trees partition data into smaller groups that get progressively more homogeneous with respect to the outcome variable.(28) Individual trees, however, tend to overfit the data, so the ensemble method of random forests was developed. The random forest algorithm generates multiple bootstrapped samples of the original data and trains one decision tree on each sample. To train a decision tree, the algorithm selects at each split the best predictor from a random subset of the original predictors. The size of the random subsets we used was the total number of predictors divided by 3 for the regression models, and the square root of the number of predictors for the classification model. We checked for improvement with altering the number of predictors, but found none. Splitting is continued for each tree until a stopping criterion is met. This process is repeated the number of times set by the investigator, and then the final model is constructed by averaging the results from these trees. For this analysis, we chose to use 1000 trees for each model.(28) We then assessed qualitatively the overall relationship between the model and the primary and secondary outcomes using the percentage of variance explained and root mean squared error for regression, and the error rate for classification, which were calculated using out-of-bag predictions. Out-of-bag refers to the predictions being calculated using the observations randomly excluded from each tree’s training data, thus serving as internal validation data.

We used variable importance scores to assess which variables were most strongly related to KCCQ scores. The variable importance scores measure total decrease in node impurity from splitting on a variable, measured by residual sum of squares (regression) or decrease in the Gini index (classification).(29) As variable importance scores can be diluted by highly correlated variables, for the primary analysis we selected a subset of variables using a correlation filter. This filter was an unsupervised algorithm to select a group of variables with no pairwise correlations >0.5 (Spearman method), a threshold set by the investigators. We also ran sensitivity analyses using the different sets of predictors discussed above. We reran the primary analysis separately in those with and without clinically diagnosed HF.

We assessed relationships between clinical variables and symptom scores qualitatively using two and three dimensional partial dependence plots, which allow visualization of the complex, non-parametric modeled relationships.

Missing predictor values were imputed separately for each analysis using random forest imputation, which uses proximity matrices derived using random forests (with 300 trees and 5 iterations) to impute values.(29) Analyses were performed using R Version 4.0.2 (www.r-project.org) and the packages randomForest, caret, and recipes.

Results

In the 3,426 included participants, median (IQR) Overall Summary Score was 90 (IQR: 70–98). Clinical HF diagnosis was present in 354 (10.3%). Predictor values, stratified by Overall Summary Score ≥75 (low symptoms, n = 2,446 [71.4%]) versus <75 (significant symptoms, n = 980 [28.6%]), are summarized in Supplemental Table S 1. Standardized differences in predictor values (by Overall Summary Score <75 vs. ≥75) are shown in Figure 1. For the primary analysis, the random forest regression model explained 27.1% of the variance in Overall Summary Score (root mean square error 18.1) (Supplemental Table S 2). The model performance for other outcomes and sensitivity analyses were largely similar, except for the Self-Efficacy and Symptom Stability scores (0.1% and 0% variance explained), and the model for Overall Summary Score limited to persons with diagnosed HF (10.0% variance explained; RMSE 23.4).

BMI had the highest variable importance score in the primary analysis, and in all but one sensitivity analysis, and a marked difference in variable importance score was observed between BMI and the other variables (Figure 2, Supplemental Figures 3–6*).* The one sensitivity analysis where BMI did not have the highest variable importance was the model that included all predictors, including several highly correlated with BMI (weight, waist circumference, and body surface area), which diluted the variable importance score of BMI. In this model, income category had the highest importance, and BMI had the second highest. (Supplemental Figure 7).

Predicted Overall Summary Score was highest at BMI = 24.3 kg/m^2, and monotonically declined for higher BMIs, reaching a 5 point (that is, clinically important) decline at BMI = 35.7 kg/m^2 (Figure 3). A lesser decline in Overall Summary Score occurred with BMI <24.3 kg/m^2, with a 1.0-point decline at the low threshold for healthy BMI, 18.5 kg/m^2, and maximal decline of 1.9 points at the lowest BMI (15.1 kg/m^2). Hemoglobin level had the second highest variable importance score in the primary analysis (Figure 4). The highest score was with hemoglobin of 14.5 g/dl, followed by substantial decrease with lower hemoglobin levels, reaching a 5 point decline at 8.5 g/dl. With higher hemoglobin, a maximal decline of 2.0 points was observed with maximal hemoglobin (19.5 g/dl). There was no visual evidence of interaction between BMI and hemoglobin level (Supplemental Figure 8). Higher education level was associated with higher predicted Overall Summary Score, with highest attainment (college degree) having an adjusted predicted score 4 points higher than lowest attainment (no high school diploma).

Individual characteristics from each predictor category were represented among the most influential variables. Of kidney health/function metrics, eGFR demonstrated the highest variable importance (5th most important variable). Of cardiovascular metrics, heart rate was the most important (4th most important overall). The social variable educational category

was the 3rd most important variable in the primary analysis. In the analysis limited to persons with diagnosed HF, heart rate was the 2nd most important variable, following BMI, and the drop off in importance score after BMI was less marked than in other analyses (Supplemental Figure 5).

In the *post hoc* analysis which replaced BMI with waist circumference among the main predictor variable set, waist circumference had by far the highest variable importance score. The increase in node purity with for the 2nd most important variable, hemoglobin, was 42% that of waist circumference, and the shape of the relationship with Overall Summary Score was similar to that observed with BMI (Supplemental Figure 9).

Discussion

We have found that, among a large collection of clinical variables, BMI had by far the strongest relationship with a score targeting HF-type symptoms and related health status in a cohort of people with CKD with and without diagnosed HF. The modest performance of the models, however, suggests that other characteristics beyond what were evaluated in these analyses—for example, depression and depressive symptoms—may be important in explaining the observed variation in HF-type symptoms/health status.(30, 31)

It is notable that BMI, rather than cardiovascular or kidney metrics or hemoglobin level, had the strongest relationship with a HF-type symptom/health status score focused on the burden of fatigue, dyspnea, and edema. This cohort did have substantial kidney disease, clinical and subclinical cardiovascular abnormalities, diabetes, and anemia, and numerous metrics for these were included in the analyses, so insufficient variance in or capture of these factors do not seem likely explanations. What is striking is not the importance BMI, which has been previously reported, but that this factor could be the most important predictor.(10) Nevertheless, the study results do not imply causality.

First, we note that, as with almost all biometrics, extreme BMIs at either end were associated with worse outcome. This was most pronounced on BMI's long right tail, although the decline is notable with more modest BMI increases: at BMI of 35 kg/m^2, the 71.7th percentile in this cohort, the expected Overall Summary Score had dropped by nearly 5 points from the maximum, a potentially clinically relevant difference. Certainly the cross-sectional nature of our analysis does not help to point any causal arrows, but in the context of other evidence it seems possible that adiposity itself may be driving the symptoms. BMI may also be influenced by extracellular fluid volume, particularly relevant to HF symptoms, but it seems likely that the major share of BMI variance in this community dwelling cohort was due to adiposity, supported by the unchanged findings in the primary model when BMI was replaced by waist circumference. BMI remained the most important factor in subgroups with and without previously diagnosed HF.

In general populations, obesity is associated with fatigue, shortness of breath, and reduced quality of life. Fatigue is very common in obese US community dwelling individuals,(32) and a cohort study found that this relationship remained even after adjustment for depression and cytokine levels.(33) Obesity may contribute to fatigue through sleepiness driven by

sleep-related breathing disorders,(34) through depression,(35) and through neuromuscular pathways.(36) Obesity is strongly associated with shortness of breath as well, with multiple possible mechanisms, including alterations in respiratory mechanics and diaphragm structure and possibly airway obstruction.(37, 38) Obesity can also predispose to edema, through venous insufficiency, lymphedema, and other mechanisms.(39, 40) Mechanisms operating in the reverse direction—with fatigue, shortness of breath, and edema causing increased adiposity and obesity—may play an important role in our findings as well, and additional longitudinal investigations may help address this question. Finally, while the KCCQ was designed to identify social limitations and quality of life due specifically to HF-type symptoms, obesity itself may lead to social limitations and impaired quality of life, perhaps influencing KCCQ scores independently of the HF-type symptoms.(41)

Our findings expand on prior findings from CRIC showing that BMI, among other factors, was independently associated with HF-type symptoms/health status.(10) This prior analysis used a logistic regression model with the outcome of Overall Summary Score dichotomized at <75 versus 75 and above, and with forward and backward stepwise selection and found that of 11 variables, 8 were independently associated with the risk, including BMI. In this analysis, BMI was used as a categorical variable, with a baseline category of <27 kg/m^2, and other categories of 27 to <30, 30 to <35, and ≥35 kg/m^2. In this analysis, adjusted odds of having an Overall Summary Score <75 were increased by 1.52 for BMI 27 to <30, 1.78 for 30 to <35, and 3.99 for >35 kg/m^2. In other CKD cohorts obesity has been shown to be associated with more general reduced health-related quality of life metrics.(42, 43)

Further studies are needed to determine how longitudinal weight changes relate to HF-type symptoms/health status in CKD, and to determine the degree to which weights and weight changes in CKD cohorts reflect adiposity versus extracellular fluid volume and other factors. This BMI finding has potential importance for investigation as a target for symptom management in CKD given that several effective interventions are available for weight management, some of which have already demonstrated improvements in kidney outcomes.

After BMI, hemoglobin concentration was the most important variable in the primary model, with an optimal hemoglobin concentration (14.5 g/dl) substantially above the anemia cutoff. Therapeutic increase of hemoglobin levels into the normal range in CKD has previously been shown to improve subjective health status but not measures of cardiovascular health such as left ventricular mass, and thus the observed association may provide evidence that the KCCQ Overall Summary Score is more reflective of general subjective health status rather than cardiac health status in this cohort.(44) Education level was next in the primary model, likely reflecting the important relationship of sociodemographic factors and health status. Heart rate, influenced by relevant processes such as neurohormonal activation and cardiac function, and eGFR, a primary measure of kidney health/function, round out the 5 most important variables.

Limitations

Ours was a cross sectional study. Outcomes were self-reported and subject to underlying biases based on participants' other lifestyle factors that may not be captured; however, the subjective nature of the outcome may increase its relevance as a patient-centered metric.

The outcome is subject to recall bias. NT-proBNP levels were not available in the cross-sectional data (thus values from 1-year prior were used in sensitivity analyses). The models performance was modest, and much of the variation in KCCQ scores remained unexplained. The KCCQ has not been validated for use in non-HF populations. The CRIC study enrolled persons specifically with CKD, raising issues of potentially biased associations, including through collider bias.(45)

Conclusion

Adiposity-related measures, rather than kidney health/function or cardiovascular health/function measurements, showed the strongest relationship with subjective HF-type symptoms and related health status in a CKD cohort. Future investigations assessing whether adiposity is a potential modifiable causal factor for HF-type symptoms in CKD may be warranted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Dr. Navaneethan is supported by research funding from the Department of Veterans Affairs Health Services Research & Development (1I01HX002917–01A1) and a grant from the National Institutes of Health (NIDDK-R01DK101500). Dr. Gregg is supported in part by the Center for Innovations in Quality, Effectiveness and Safety (CIN 13–413), Michael E. DeBakey VA Medical Center, Houston, TX

Funding

Funding for this project was obtained through the CRIC Study Opportunity Pool Program. Funding for the CRIC Study was obtained under a cooperative agreement from National Institute of Diabetes and Digestive and Kidney Diseases (U01DK060990, U01DK060984, U01DK061022, U01DK061021, U01DK061028, U01DK060980, U01DK060963, U01DK060902 and U24DK060990). In addition, this work was supported in part by: the Perelman School of Medicine at the University of Pennsylvania Clinical and Translational Science Award NIH/NCATS UL1TR000003, Johns Hopkins University UL1 TR-000424, University of Maryland GCRC M01 RR-16500, Clinical and Translational Science Collaborative of Cleveland, UL1TR000439 from the National Center for Advancing Translational Sciences (NCATS) component of the National Institutes of Health and NIH roadmap for Medical Research, Michigan Institute for Clinical and Health Research (MICHR) UL1TR000433, University of Illinois at Chicago CTSA UL1RR029879, Tulane COBRE for Clinical and Translational Research in Cardiometabolic Diseases P20 GM109036, Kaiser Permanente NIH/NCRR UCSF-CTSI UL1 RR-024131, Department of Internal Medicine, University of New Mexico School of Medicine Albuquerque, NM R01DK119199.

Data availability

The data used for this study are available from the National Institute of Diabetes and Digestive and Kidney Diseases Central Repository: https://repository.niddk.nih.gov/studies/cric/.

Abbreviations

HF	heart failure
CKD	chronic kidney disease
BMI	body mass index

References

1. Brown SA, Tyrer FC, Clarke AL, Lloyd-Davies LH, Stein AG, Tarrant C, et al. Symptom burden in patients with chronic kidney disease not requiring renal replacement therapy. Clinical Kidney Journal. 2017;10(6):788–96. [PubMed: 29225808]
2. Abdel-Kader K, Unruh ML, Weisbord SD. Symptom burden, depression, and quality of life in chronic and end-stage kidney disease. Clinical journal of the American Society of Nephrology : CJASN. 2009;4(6):1057–64. [PubMed: 19423570]
3. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. J Ren Care. 2013;39(3):140–50. [PubMed: 23826803]
4. House AA, Wanner C, Sarnak MJ, Piña IL, McIntyre CW, Komenda P, et al. Heart failure in chronic kidney disease: conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference. Kidney International. 2019;95(6):1304–17. [PubMed: 31053387]
5. Park M, Hsu CY, Li Y, Mishra RK, Keane M, Rosas SE, et al. Associations between kidney function and subclinical cardiac abnormalities in CKD. J Am Soc Nephrol. 2012;23(10):1725–34. [PubMed: 22935481]
6. Bansal N, Roy J, Chen H-Y, Deo R, Dobre M, Fischer MJ, et al. Evolution of Echocardiographic Measures of Cardiac Disease From CKD to ESRD and Risk of All-Cause Mortality: Findings From the CRIC Study. American Journal of Kidney Diseases. 2018;72(3):390–9. [PubMed: 29784617]
7. Di Lullo L, Gorini A, Russo D, Santoboni A, Ronco C. Left ventricular hypertrophy in chronic kidney disease patients: from pathophysiology to treatment. Cardiorenal medicine. 2015;5(4):254–66. [PubMed: 26648942]
8. Zambroski CH, Moser DK, Bhat G, Ziegler C. Impact of symptom prevalence and symptom burden on quality of life in patients with heart failure. European Journal of Cardiovascular Nursing. 2005;4(3):198–206. [PubMed: 15916924]
9. Spertus JA, Jones PG, Sandhu AT, Arnold SV. Interpreting the Kansas City Cardiomyopathy Questionnaire in Clinical Trials and Clinical Care: JACC State-of-the-Art Review. J Am Coll Cardiol. 2020;76(20):2379–90. [PubMed: 33183512]
10. Shlipak MG, Lash JP, Yang W, Teal V, Keane M, Cappola T, et al. Symptoms characteristic of heart failure among CKD patients without diagnosed heart failure. Journal of cardiac failure. 2011;17(1):17–23. [PubMed: 21187260]
11. Mishra RK, Yang W, Roy J, Anderson AH, Bansal N, Chen J, et al. Kansas City Cardiomyopathy Questionnaire Score Is Associated With Incident Heart Failure Hospitalization in Patients With Chronic Kidney Disease Without Previously Diagnosed Heart Failure: Chronic Renal Insufficiency Cohort Study. Circ Heart Fail. 2015;8(4):702–8. [PubMed: 25985796]
12. Munger E, Choi H, Dey AK, Elnabawi YA, Groenendyk JW, Rodante J, et al. Application of machine learning to determine top predictors of noncalcified coronary burden in psoriasis: An observational cohort study. J Am Acad Dermatol. 2020;83(6):1647–53. [PubMed: 31678339]
13. Munger E, Hickey JW, Dey AK, Jafri MS, Kinser JM, Mehta NN. Application of machine learning in understanding atherosclerosis: Emerging insights. APL Bioeng. 2021;5(1):011505. [PubMed: 33644628]
14. Chen X, Wang CC, Yin J, You ZH. Novel Human miRNA-Disease Association Inference Based on Random Forest. (2162–2531 (Print)).
15. Lash JP, Go AS, Appel LJ, He J, Ojo A, Rahman M, et al. Chronic Renal Insufficiency Cohort (CRIC) Study: Baseline Characteristics and Associations with Kidney Function. Clinical Journal of the American Society of Nephrology. 2009;4(8):1302. [PubMed: 19541818]
16. Feldman HI, Appel LJ, Chertow GM, Cifelli D, Cizman B, Daugirdas J, et al. The Chronic Renal Insufficiency Cohort (CRIC) Study: Design and Methods. J Am Soc Nephrol. 2003;14(7 Suppl 2):S148–53. [PubMed: 12819321]
17. Heidenreich PA, Spertus JA, Jones PG, Weintraub WS, Rumsfeld JS, Rathore SS, et al. Health status identifies heart failure outpatients at risk for hospitalization or death. Journal of the American College of Cardiology. 2006;47(4):752–6. [PubMed: 16487840]

18. Joseph SM, Novak E, Arnold SV, Jones PG, Khattak H, Platts AE, et al. Comparable performance of the Kansas City Cardiomyopathy Questionnaire in patients with heart failure with preserved and reduced ejection fraction. Circulation: Heart Failure. 2013;6(6):1139–46. [PubMed: 24130003]
19. Spertus JA, Jones PG. Development and validation of a short version of the Kansas City Cardiomyopathy Questionnaire. Circulation: Cardiovascular Quality and Outcomes. 2015;8(5):469–76. [PubMed: 26307129]
20. Green CP, Porter CB, Bresnahan DR, Spertus JA. Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure. J Am Coll Cardiol. 2000;35(5):1245–55. [PubMed: 10758967]
21. U.S. Food and Drug Administration. Qualified Clinical Outcome Assessments (COA) Silver Spring, Maryland: Department of Health and Human Services; 2020 [updated 4/14/2020. Available from: https://www.fda.gov/drugs/clinical-outcome-assessment-coa-qualification-program/qualified-clinical-outcome-assessments-coa.
22. Kuhn M, Johnson K. Feature engineering and selection: A practical approach for predictive models: CRC Press; 2019.
23. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. 2nd Edition, Corrected 12th Printing ed. New York, NY: Springer; 2017.
24. Walsh CG, Ribeiro JD, Franklin JC. Predicting risk of suicide attempts over time through machine learning. Clinical Psychological Science. 2017;5(3):457–69.
25. Ambale-Venkatesh B, Yang X, Wu CO, Liu K, Hundley WG, McClelland R, et al. Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis. Circ Res. 2017;121(9):1092–101. [PubMed: 28794054]
26. Huynh-Thu VA, Saeys Y, Wehenkel L, Geurts P. Statistical interpretation of machine learning-based feature importance scores for biomarker discovery. Bioinformatics. 2012;28(13):1766–74. [PubMed: 22539669]
27. Biau G, Scornet E. A random forest guided tour. TEST. 2016;25(2):197–227.
28. Kuhn M, Johnson K. Applied Predictive Modeling. Corrected 5th Printing ed. New York, NY: Springer; 2016.
29. Liaw A, Wiener M. Classification and regression by randomForest. R news. 2002;2(3):18–22.
30. Gregg LP, Bossola M, Ostrosky-Frid M, Hedayati SS. Fatigue in CKD: Epidemiology, Pathophysiology, and Treatment. Clin J Am Soc Nephrol. 2021.
31. Rumsfeld JS, Havranek E, Masoudi FA, Peterson ED, Jones P, Tooley JF, et al. Depressive symptoms are the strongest predictors of short-term declines in health status in patients with heart failure. J Am Coll Cardiol. 2003;42(10):1811–7. [PubMed: 14642693]
32. Resnick HE, Carter EA, Aloia M, Phillips B. Cross-sectional relationship of reported fatigue to obesity, diet, and physical activity: results from the third national health and nutrition examination survey. Journal of Clinical Sleep Medicine. 2006;2(02):163–9. [PubMed: 17557490]
33. Lim W, Hong S, Nelesen R, Dimsdale JE. The association of obesity, cytokine levels, and depressive symptoms with diverse measures of fatigue in healthy subjects. Archives of internal medicine. 2005;165(8):910–5. [PubMed: 15851643]
34. Beccuti G, Pannain S. Sleep and obesity. Current opinion in clinical nutrition and metabolic care. 2011;14(4):402. [PubMed: 21659802]
35. Preiss K, Brennan L, Clarke D. A systematic review of variables associated with the relationship between obesity and depression. Obesity Reviews. 2013;14(11):906–18. [PubMed: 23809142]
36. Mehta R Impacts of obesity and stress on neuromuscular fatigue development and associated heart rate variability. International Journal of Obesity. 2015;39(2):208–13. [PubMed: 25042859]
37. Buras ED, Converso-Baran K, Davis CS, Akama T, Hikage F, Michele DE, et al. Fibro-adipogenic remodeling of the diaphragm in obesity-associated respiratory dysfunction. Diabetes. 2019;68(1):45–56. [PubMed: 30361289]
38. Brock JM, Billeter A, Müller-Stich BP, Herth F. Obesity and the Lung: What We Know Today. Respiration. 2020;99(10):856–66. [PubMed: 33242862]
39. Davies HO, Popplewell M, Singhal R, Smith N, Bradbury AW. Obesity and lower limb venous disease–The epidemic of phlebesity. Phlebology. 2017;32(4):227–33. [PubMed: 27178403]

40. Savetsky IL, Torrisi JS, Cuzzone DA, Ghanta S, Albano NJ, Gardenier JC, et al. Obesity increases inflammation and impairs lymphatic function in a mouse model of lymphedema. American Journal of Physiology-Heart and Circulatory Physiology. 2014;307(2):H165–H72. [PubMed: 24858842]

41. Kolotkin RL, Andersen JR. A systematic review of reviews: exploring the relationship between obesity, weight loss and health-related quality of life. Clinical obesity. 2017;7(5):273–89. [PubMed: 28695722]

42. Schweitzer ML, Stengel B, Legrand K, Briancon S, Jacquelinet C, Combe C, et al. Obesity phenotype and patient-reported outcomes in moderate and severe chronic kidney disease: a cross-sectional study from the CKD-REIN cohort study. Qual Life Res. 2019;28(7):1873–83. [PubMed: 30659448]

43. Perlman RL, Finkelstein FO, Liu L, Roys E, Kiser M, Eisele G, et al. Quality of life in chronic kidney disease (CKD): a cross-sectional analysis in the Renal Research Institute-CKD study. Am J Kidney Dis. 2005;45(4):658–66. [PubMed: 15806468]

44. Drüeke TB, Locatelli F, Clyne N, Eckardt K-U, Macdougall IC, Tsakiris D, et al. Normalization of hemoglobin level in patients with chronic kidney disease and anemia. New England Journal of Medicine. 2006;355(20):2071–84. [PubMed: 17108342]

45. Munafo MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol. 2018;47(1):226–35. [PubMed: 29040562]

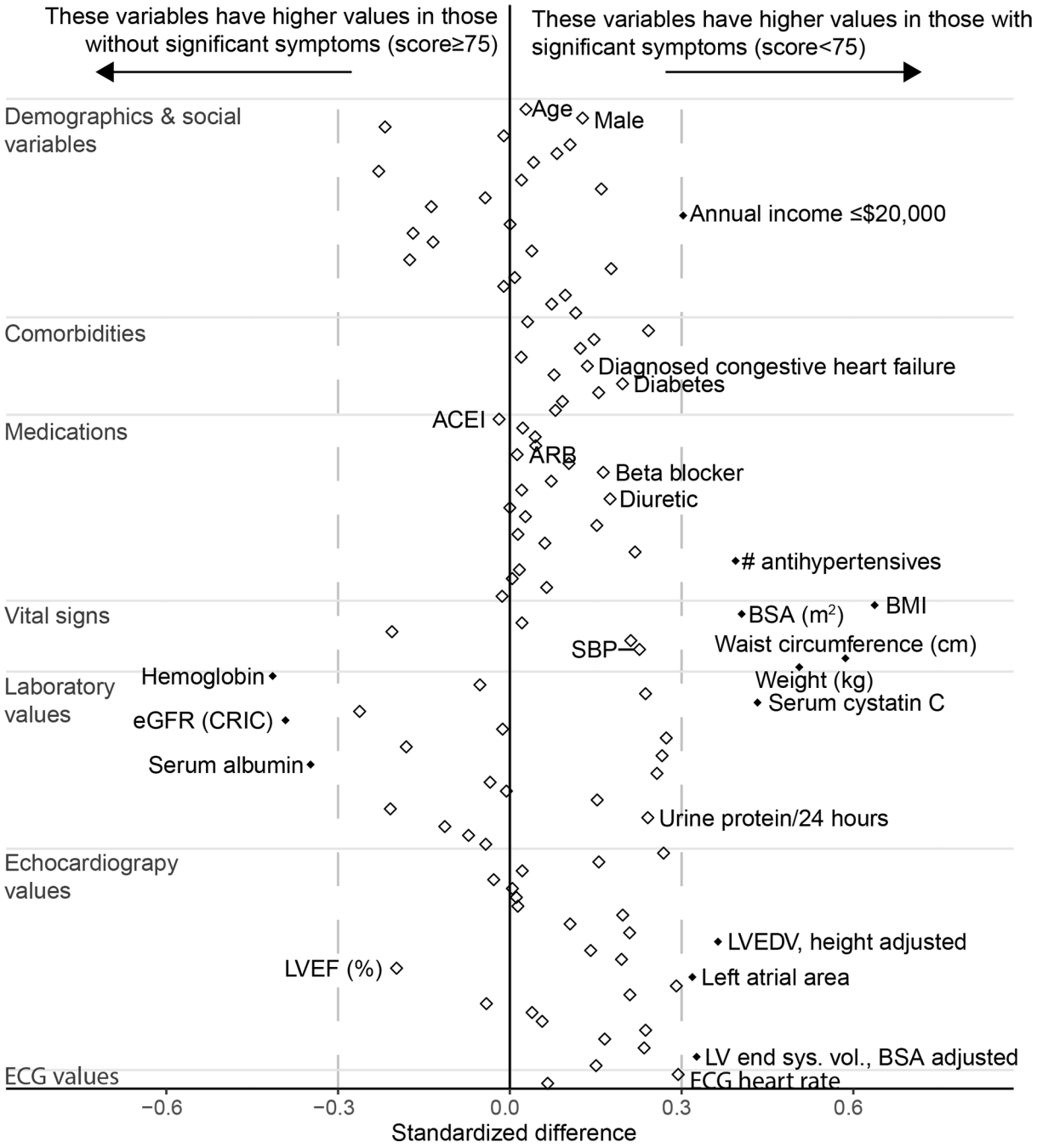

Figure 1. Standardized differences of predictor variables: KCCQ Overall Summary Score <75 (clinically important symptoms) vs. ≥75 (absence of clinically important symptoms)

Notes: Because of the large number of variables, values are labeled for variables with absolute value of standardized difference ≥0.3 and for selected variables of clinical interest. Standardized mean differences were used for continuous variables, and raw difference in proportion was used binary variables. Categorical variables were divided into binary subcategories to calculate standardized differences.

Abbreviations: KCCQ, Kansas City Cardiomyopathy Questionnaire; ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin II receptor blocker; BMI, body mass index;

BSA, body surface area; SBP, systolic blood pressure; eGFR, estimated glomerular filtration rate; CRIC, Chronic Renal Insufficiency Cohort; LVEDV, left ventricular end diastolic volume; LVEF, left ventricular ejection fraction.

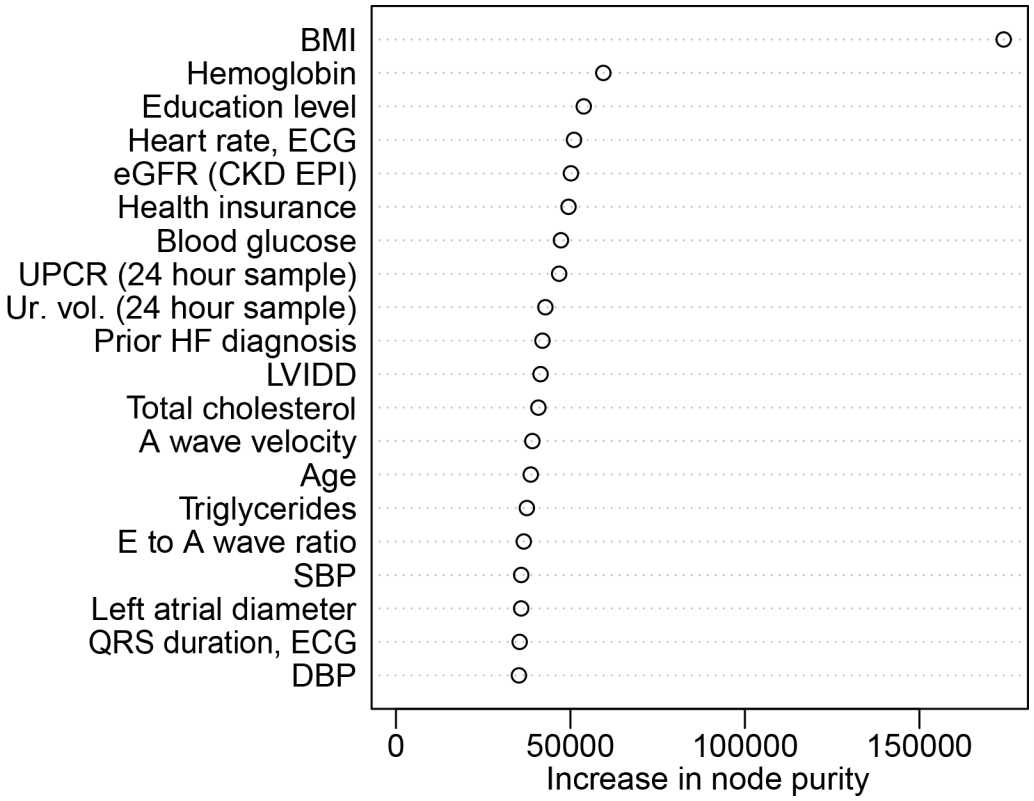

Figure 2. Variable importance scores for predicting KCCQ Overall Summary Score, for the 20 most important variables

Notes: Overall Summary Score reflects total symptoms (fatigue, shortness of breath, peripheral edema), and their effects on physical function, social limitation, and quality of life. Increase in node purity is the total decrease in the residual sum of squares from splitting on a variable, and can be interpreted in terms of relative values among the variables.

Abbreviations: BMI, body mass index; ECG, electrocardiogram; eGFR, estimated glomerular filtration rate; CKD EPI, Chronic Kidney Disease Epidemiology Collaboration; UPCR, urine protein-to-creatinine ratio; Ur. vol., urine volume; LVIDD, left ventricular internal diameter-end diastole; SBP, systolic blood pressure; DBP, diastolic blood pressure.

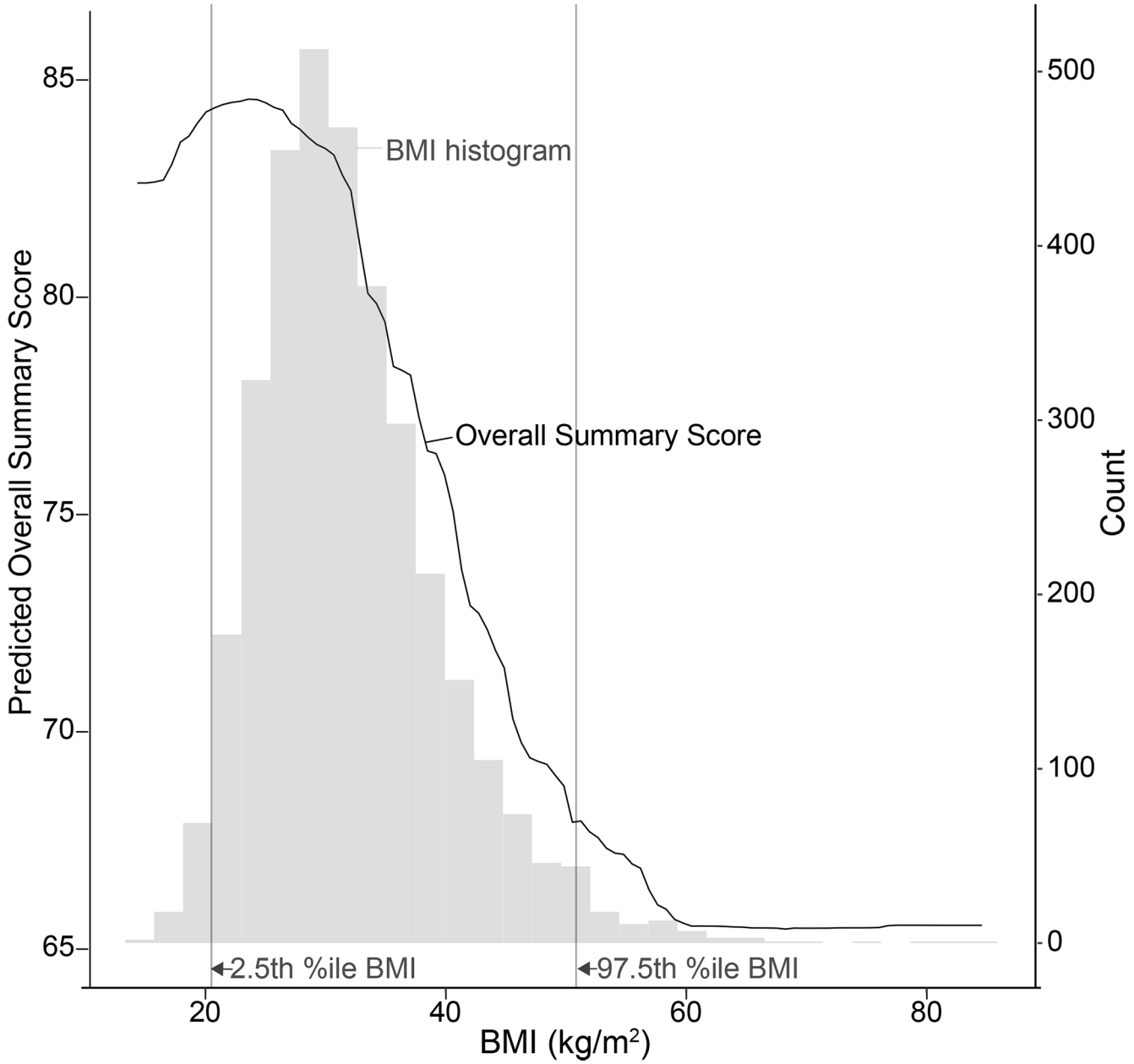

Figure 3. Adjusted* KCCQ Overall Summary Scores over BMI from the primary model (line), and distribution of BMI in the study cohort (bars)

*Partial dependence plots demonstrating the predicted score adjusted for the 63 other variables in the primary model (bolded variables in Table 1). Partial dependence plots show the average predictions from the model when the variable of interest alone is varied for all participants.

Note: The region between the two vertical gray lines highlights the middle 95% of the BMI distribution. The BMI distribution is provided to demonstrate participant BMIs available for model fitting, and does not represent the participants used for predicted scores (all participants were included to create predicted scores at all BMIs).

Abbreviations: KCCQ, Kansas City Cardiomyopathy Questionnaire; BMI, body mass index

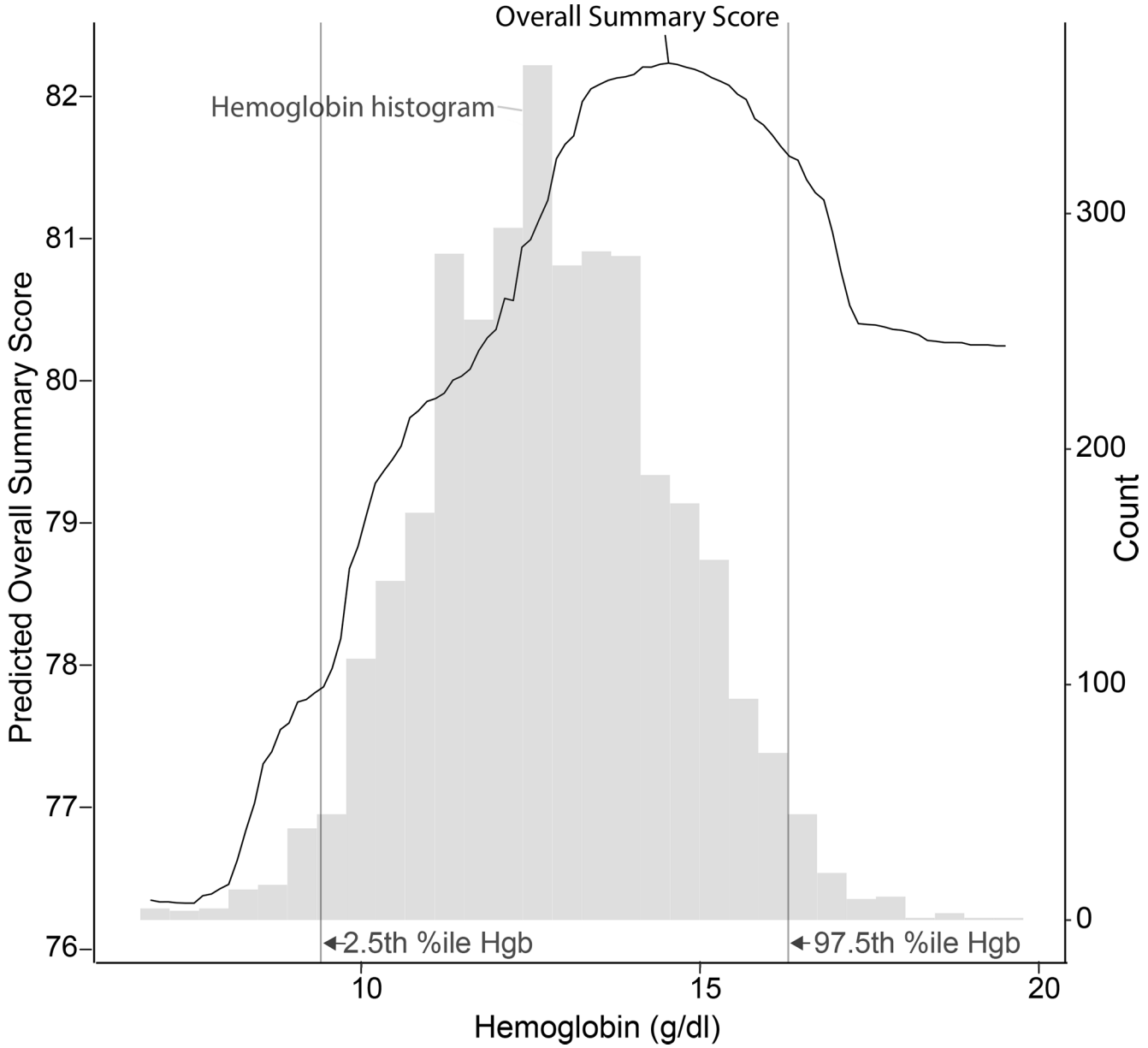

Figure 4. Adjusted* KCCQ Overall Summary Scores over hemoglobin from the primary model (line), and distribution of hemoglobin in the study cohort (bars)

*Partial dependence plots demonstrating the predicted score adjusted for the 63 other variables in the primary model (bolded variables in Table 1). Partial dependence plots show the average predictions from the model when the variable of interest alone is varied for all participants.

Note: The region between the two vertical gray lines highlights the middle 95% of the hemoglobin distribution. The hemoglobin distribution is provided to demonstrate participant hemoglobins available for model fitting, and does not represent the participants used for predicted scores (all participants were included to create predicted scores at all hemoglobin levels).

Abbreviations: KCCQ, Kansas City Cardiomyopathy Questionnaire; Hgb, hemoglobin

Table 1.

Predictors of interest for investigation of the relationship to KCCQ HF-type symptom score (cross sectional)

Category	Variable
Demographics (n = 3)	**Age**, Sex, **Race/ethnicity** (Non-Hispanic white, Non-Hispanic Black, Hispanic, Other)
Social Factors (n = 8)	**Alcohol use**, **Recreational drug use**, **Education level** (less than high school graduation, High school graduate, Some college, College graduate), **Health Insurance** (None, Medicaid, Other, Unknown), Income per year ($20,000 or less, $20,001–$50,000, $50,001–$100,000, >$100,000), **Smoked >100 tobacco cigarettes**, **Current tobacco smoking**, **prior nephrologist visit**
Comorbidities (n = 12)	Any cardiovascular disease, **Diagnosed heart failure**, **Coronary heart disease**, **Peripheral arterial disease**, **Stroke**, **Atrial fibrillation**, **Amputation**, **Arthritis**, **Asthma**, **Cancer besides skin cancer in prior 5 years**, **COPD**, Diabetes
Medications (n = 21)	**ACEI**, Aldosterone antagonist, **Alpha 2 agonist**, **Alpha 1 blocker**, **ARB**, **Glycemic therapy**, **Beta blocker**, **Calcium channel blocker**, **Digoxin**, **Diuretic**, **DPP4 inhibitor**, **ESA**, **Insulin**, **Potassium sparing diuretic**, Lipid therapy, Loop diuretic, Number of antihypertensive medications, **NSAID**, **Non-statin lipid therapy**, **Statin**, **Thiazide diuretic**
Clinical measures (n = 8)	**SBP (mmHg)**, **DBP (mmHg)**, Pulse (per minute), **BMI (kg/m2)**, BSA (m2), Height (cm), Waist circumference (cm), Weight (kg)
Laboratory Measures (n = 20)	Serum creatinine, Serum cystatin C, **eGFR (CKD EPI, ml/min/1.73m2)**, eGFR (CRIC, ml/min/1.73m2), **PCR (g/g, 24 hour urine sample)**, BUN (mg/dl), Urine creatinine excretion (g, 24 hour sample), Urine protein (g, 24 hour sample), Urine creatinine concentration (mg/dl, 24 hour sample), **Urine volume (ml, 24 hour sample)**, Urine sodium (mmol/L, 24 hour sample), **Hemoglobin (g/dl)**, **CO2 (mEq/L)**, **Fasting for labs**, **Blood glucose (mg/dl)**, **HDL-C (mg/dl)**, **Serum albumin (g/dl)**, **Serum sodium (mmol/L)**, **Total cholesterol (mg/dl)**, **Triglycerides (mg/dl)**
Echocardiography (n = 25)	A wave velocity (cm/s), **Aortic valve peak velocity (cm/s)**, **Transmitral E wave deceleration time (msec)**, **Degree of aortic valve thickening (scale from 0–2)**, **Degree of aortic regurgitation (scale from 0–2)**, **Degree of mitral annulus calcification (scale from 0–2)**, **Degree of mitral regurgitation (scale from 0–3)**, LVEDD, 4 chamber view (cm2), LVEDL, 4 chamber view (cm), LVEDV, 4 chamber view (ml), LVEDV, height adjusted (ml/m2.7), E wave velocity (cm/s), LVEF (%), Left atrial area (cm2), **Left atrial diameter (cm)**, **LVIDD (cm)**, **E to A wave ratio**, MR, 4 chamber area (cm2), **Pericardial effusion**, (scale from 0–2), LV end systolic area (cm2), LV end systolic length (cm), LV end systolic volume (ml), LV end sys. vol., adj (ml/m2.7), Diastolic sphericity index, **Systolic sphericity index**
ECG (n = 2)	**QRS duration (ms)**, **Heart rate (per minute)**

Notes: **Bolded variables** are those selected using the correlation filter and included in the primary analysis. All variables were included in sensitivity analyses, and the 10 variables shaded in gray were used in separate sensitivity analyses. These 10 variables were selected based on physiologic considerations and based or high standardized differences with KCCQ Overall Summary Score dichotomized at 75. NT-proBNP, available only from 1 year prior to the analysis baseline, was included in a sensitivity analysis as well. Variables were chosen based on plausible relationships to kidney, cardiovascular, and general health status and functioning.

Abbreviations: COPD, chronic obstructive pulmonary disease; ACEI, angiotensin-converting enzyme inhibitor ; ARB, angiotensin II receptor blocker; DPP4, dipeptidyl peptidase 4; ESA, erythropoiesis-stimulating agent; NSAID, non-steroidal anti-inflammatory drug; SBP, systolic blood pressure; DBP, diastolic blood pressure; BMI, body mass index; eGFR, estimated glomerular filtration rate; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration; PCR, protein-to-creatinine ratio; LVEDD, left ventricular end diastolic diameter; LVEDL, left ventricular end diastolic length; LVEDV, left ventricular end diastolic volume; LVEF, left ventricular ejection fraction; LVIDD, left ventricular internal diameter-diastole.