

Development and evaluation of cardiovascular disease risk prediction models for patients with type 2 diabetes

Received: 6 November 2025

Accepted: 17 March 2026

Published online: 31 March 2026

Cite this article as: Yang Y., Liu T., Liao C. *et al.* Development and evaluation of cardiovascular disease risk prediction models for patients with type 2 diabetes. *Sci Rep* (2026). <https://doi.org/10.1038/s41598-026-45129-5>

Yang Yang, Tian Liu, Che-Yi Liao, Sun Ju Lee, Esmail Keyvanshokoh, Hui Shao, Mary Beth Weber, Francisco J. Pasquel & Gian-Gabriel P. Garcia

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

Title: Development and Evaluation of Cardiovascular Disease Risk Prediction Models for Patients with Type 2 Diabetes

Authors: Yang Yang*, BS¹; Tian Liu*, MS²; Che-Yi Liao, MS¹; Sun Ju Lee, PhD²; Esmaeil Keyvanshokoo, PhD³; Hui Shao, PhD⁴; Mary Beth Weber, PhD⁴; Francisco J. Pasquel, MD, MPH⁵; Gian-Gabriel P. Garcia, PhD^{6†} (garciagg@uw.edu)

¹H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA

²Freeman College of Management, Bucknell University, Lewisburg, PA, USA

³Department of Information and Operations Management, Mays Business School, Texas A&M University, College Station, TX, USA

⁴Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA

⁵Division of Endocrinology, Metabolism, and Lipids, Department of Medicine, Emory University School of Medicine, Atlanta, GA, USA

⁶Department of Industrial & Systems Engineering, University of Washington, Seattle, WA, USA

* denotes equal contribution.

†Corresponding Author: Gian-Gabriel P. Garcia, PhD, Department of Industrial & Systems Engineering, University of Washington, Seattle, WA (garciagg@uw.edu)

Abstract

To facilitate treatment decisions in people at risk of Cardiovascular Disease (CVD), several risk equations such as the Pooled Cohort Equations and Predicting Risk of Cardiovascular Disease Events (PREVENT) equations have been developed to estimate CVD risk for primary prevention patients. However, it is unclear whether these equations achieve high predictive accuracy and fairness in patients with type 2 diabetes (T2D), and whether a T2D-specific risk equation is needed. Accordingly, we developed a Weibull Accelerated Failure Time (AFT) survival model for predicting the 3-year CVD risk in 23,795 patients with T2D from the All of Us dataset, using sociodemographic information, physical measurements, medication, and CVD history. Among patients without CVD history, our Weibull AFT (vs. PREVENT) achieved a greater C-index (0.646 vs. 0.465), greater Concordance Fractions (0.610-0.674 vs 0.541-0.600), and comparable Concordance Imparity (0.006 vs. 0.002) across sex and race/ethnicity (0.065 vs. 0.058) subgroups. Our findings highlight the need for a T2D-specific CVD risk equation and demonstrate the value of diverse datasets for developing fair and accurate predictive models.

**Keywords: Risk Equation Development, Survival Modeling,
Cardiovascular Diseases, Type 2 diabetes, Fairness
Evaluation**

ARTICLE IN PRESS

Introduction

Cardiovascular disease (CVD) is a leading cause of death worldwide. In the United States, national CVD mortality rose during the COVID-19 pandemic and has remained high; more than 200,000 CVD deaths occurred from 2020 to 2022, representing a 9% increase over projections based on historical CVD deaths from 2010 to 2019.¹ Notably, adults with type 2 diabetes (T2D) experience CVD at elevated rates due to commonly co-occurring risk factors such as hypertension and dyslipidemia. Moreover, CVD is the leading cause of morbidity and mortality in this population.² To facilitate timely prevention and treatment among at-risk patients, accurate estimation of CVD risk is critical.

To facilitate treatment decisions for at-risk patients, significant efforts have been made to develop CVD risk prediction models. In 2013, the American Heart Association (AHA) and the American College of Cardiology (ACC) developed the pooled cohort equations (PCEs), which are sex and race-specific models to estimate the 10-year atherosclerotic CVD (ASCVD) event risk for African-American and White men and women aged 40 to 79 years without previous CVD history.^{3,4} The PCEs are based on age, total cholesterol, high-density lipoprotein cholesterol, systolic BP, diabetes mellitus, and current smoking status. Although the PCEs were endorsed by the 2019 AHA/ACC Primary Prevent Guidelines, experts raised many concerns. First, the PCEs did not account for heart failure, which is a rising prevalent CVD

subtype.^{5,6} In addition, the PCEs do not generalize to other race and ethnicity groups such as Asians and Hispanics, which were not included in its derivation.⁷ Moreover, the PCEs were developed over a decade ago and thus may not reflect recent population-level changes in risk factors and better preventive treatments.^{8,9}

With the goal of more accurate and equitable CVD risk assessment, in 2023, the AHA Cardiovascular-Kidney-Metabolic Scientific Advisory Group developed and externally validated the Predicting Risk of Cardiovascular Disease EVENTS (PREVENT) risk equations using 25 datasets that contain a total of 6 million participants collected between 1992 and 2017.^{10,11} PREVENT focused on predicting the risk of ASCVD, heart failure, and total CVD for primary prevention patients, providing sex-specific 10-year risk estimates for individuals 30-79 years of age and 30-year risk estimates for individuals 30-59 years of age. Enhanced models that consider Hemoglobin A1c, urine albumin-to-creatinine ratio, and social deprivation index were also developed.

While risk equations such as the PCEs and PREVENT have positively impacted clinical care for CVD, there remain several opportunities to further advance the utility of these tools. First, there is a dearth of CVD risk equations specifically for adults with T2D. Accordingly, there is limited evidence on whether existing risk equations such as PREVENT can adequately predict

risk of CVD among adults with T2D compared to specific risk equations developed for this population. Second, it is unclear whether widely used risk equations generate fair predictions across different demographic subgroups. For example, despite the extensive external validation performed on PREVENT, there exists limited literature which examines its predictive fairness across different racial subgroups.¹² Lastly, existing risk equations focused on primary prevention and thus may not generalize to patients with a history of CVD.

To bridge these gaps, we developed a Weibull Accelerated Failure Time (AFT) survival model using a diverse cohort of patients with T2D from the National Institutes of Health (NIH) All of Us dataset to predict 3-year risk of CVD among adults with T2D. We assess this model across several dimensions, including variable importance, model calibration, predictive accuracy, and fairness. Then, we benchmark against PREVENT across these metrics. Our results demonstrate that our Weibull AFT model outperforms PREVENT in calibration, predictive accuracy, and group fairness when validated against a diverse cohort of adult patients with T2D.

Methods

Data Source

We leveraged the All of Us dataset from the NIH,¹³ a longitudinal cohort study aiming to advance precision medicine. The All of Us dataset integrated diverse individual-level data from electronic health records (EHRs),

participant-derived information from surveys, physical measurements, biospecimens, and wearables. As of August 2024, the program has enrolled over 832,000 participants, covering a diverse population from historically underrepresented groups in biomedical research (>80%), e.g., racial and ethnic minorities (45%).¹⁴ Every participants in the All of Us program completes a survey at enrollment which gathers information about a participant's demographics, lifestyle, personal and family health history, healthcare access and utilization, and social determinants of health among other topics.

Data Preprocessing

Using the cloud-based Research Workbench, we followed three steps to preprocess the data. First, we selected patients with T2D and extracted demographic characteristics (age, sex, race, ethnicity), socio-economic factors (e.g., employment status, insurance coverage), clinical features (CVD/kidney disease history), medication history, and relevant biomarkers (see Table 1 for more information). These predictors were chosen for their clinical relevance and availability in primary care. See eAppendix A for additional details on data preprocessing. Second, we filtered out patients younger than 40 to align with AHA guidelines, which recommend risk estimation for adults aged 40 to 74, considering that CVD typically occurs in older populations.¹⁵ Third, we used Random Forest to impute missing continuous data (< 6%) and one-hot encoded the categorical variables. We

split the data into 80% training and 20% testing sets and normalized continuous variables before model fitting. Fourth, we added interaction terms to capture the non-additive effects between pre-survey medication history and biomarkers (see details in the supplementary risk calculation spreadsheet).

Next, we formatted our data so that it was amenable to survival analysis, i.e., time-to-event modeling. Since patients may develop competing diseases (e.g., kidney disease) before CVD, our study focuses on time-to-first-event (i.e., *duration*) modeling.¹⁶ That is, for patients who develop CVD, the duration is the interval between the survey and diagnosis dates. For right-censored patients,¹⁷ the duration is the time between the survey and last clinical visit unless they are censored due to a competing event. If a competing disease occurs before CVD, we use the interval between the survey date and the competing disease diagnosis date.

Survival Model Development

Similarly to PREVENT, our primary outcome was risk of CVD event, comprised of myocardial infarction (i.e., heart attack), stroke, and heart failure. Since 96% of durations in our data were 3 years or fewer, we focused on 3-year CVD risk. Moreover, we opted to use a Weibull Accelerated Failure Time (Weibull AFT) model for CVD risk prediction because of its intuitive interpretation in survival analysis, as demonstrated in previous works.^{18,19} In

our analysis, we fit a Weibull AFT model to the training data with demographic characteristics, socio-economic factors, clinical features, medication history, and relevant biomarkers as predictors. See eAppendix B for additional details on comparisons between our modeling variables and those in PREVENT. Also, see eAppendix C for an analysis on the impact of race and ethnicity as predictors and eAppendix D for an analysis of our model trained only on the subset of patients with history of CVD. Finally, since PREVENT was developed for patients without CVD history, we performed all comparisons between PREVENT and the Weibull AFT model on a subset of patients without CVD history in the testing set as well as the entire testing set. Comparisons were made on calibration, predictive accuracy, and fairness as are described in further detail below. 10-year risks from PREVENT were approximately converted to 3-year risks (see eAppendix E).

Variable Importance

We evaluated the impact of predictors on the performance of our Weibull AFT model using permutation importance.²⁰ Permutation importance measures each predictor's impact on predictive accuracy by perturbing its values while keeping others constant and calculating the resulting change in performance as measured by area under the receiver operating characteristic curve (AUC). To ensure reliable estimates, we bootstrapped 100 iterations, sampling the training data, training a model, calculating baseline AUC, and determining importance scores for each iteration. We report the average permutation

importance scores, where a higher importance indicates stronger model dependence on that variable to make accurate predictions.

Model Calibration

We evaluated the calibration of our Weibull AFT model and PREVENT by comparing predicted 3-year CVD risks with observed event rates across sex subgroups and race and ethnicity subgroups in the test set using risk calibration plots.²¹ For each subgroup, we calculated both the Weibull AFT model's predicted 3-year CVD risk and the scaled risk from PREVENT equations. After grouping patients into deciles based on predicted risk scores, we computed the mean predicted risk and observed event rate for each subgroup. Plotting observed rates against mean predicted risks revealed any discrepancies in risk estimation across predicted risk levels.

Accuracy Evaluation

We evaluated the predictive accuracy of the Weibull AFT model and PREVENT using the concordance index (C-index) in the testing set. The C-index describes the proportion of time that, when two patients are randomly selected, the patient with a shorter time-to-CVD event is given a higher risk score than the patient with the longer time-to-CVD event. Greater C-index values are indicative of better predictive accuracy.

Fairness Evaluation

We compared the fairness of the Weibull AFT model with PREVENT across different sex subgroups and race and ethnicity subgroups using the Concordance Fraction (CF) and Concordance Imparity (CI).²² Compared to other existing fairness metrics that are commonly used in binary classification, CF and CI account for censorship in the data to ensure rigorous fairness evaluation in survival analysis.²³ Concretely, CF measures individual-level prediction accuracy within each subgroup by comparing model predictions with true outcomes. It excludes incomparable pairs where the shorter duration is censored or both durations are censored with identical survival times. CF for a subgroup is computed as the proportion of all correctly predicted pairs to all comparable pairs. A higher CF for a subgroup indicates better prediction accuracy for that group. We remark that CF and C-index have similar interpretations. However, a subgroup-specific C-index would only include comparisons within that subgroup and would fail to include comparisons to individuals outside of that subgroup.

Our second metric, CI, measures the disparity in prediction performance across subgroups. It is calculated as the absolute difference between CFs of the two subgroups that result in the largest difference. A lower CI score indicates a more fair model, suggesting that the model does not significantly favor one subgroup over another.

Results

Description of Study Population

Table 1 summarizes our training (N=19,036) and testing datasets (N=4,759).

Table 1. Characteristics of study participants (n=23,795) with type 2 Diabetes from All of Us dataset. Categorical features are represented as n (%) and continuous features are represented as mean (SD).

Characteristics	Training data (n=19,036)	Testing data (n=4,759)
Demographics		
Age, years		
40-49	2,097 (11.77%)	501 (11.25%)
50-59	4,767 (26.76%)	1,171(26.29%)
60-69	6,098 (34.23%)	1,556 (34.93%)
> 69	4,854 (27.25%)	1,227 (27.54%)
Sex		
Female	10,757 (56.51%)	2,657 (55.83%)
Male	7,836 (41.16%)	1,980 (41.61%)
Other	443 (3.00%)	122 (2.56%)
Race and Ethnicity		
White	9,081 (47.70%)	2,317 (48.69%)
Black	5,167 (27.14%)	1,256 (26.39%)
Hispanic	3,548 (18.64%)	896 (18.83%)
Other	1,240 (6.51%)	290 (6.09%)
Socio-economic factors		
Education Level: College graduate or advanced degree	5,692 (29.90%)	1,456 (30.59%)

Employment Status, employed	5,043 (26.49%)	1,240 (26.06%)
Health Insurance, current	17,910 (94.08%)	4,490 (94.34%)
Living Situation: Stable House Concern	3,093 (16.25%)	799 (16.79%)
Income, annual income		
Less than 10k	3,046 (16.00%)	791 (16.62%)
Between 10k and 100k	9,808 (51.52%)	2,414 (50.72%)
More than 100k	2,064 (10.84%)	531 (11.16%)
Other	3,758 (19.74%)	1,023 (21.50%)
Housing Type		
Apartment	86 (0.45%)	17 (0.36%)
Single Family	1,791 (9.42%)	474 (9.96%)
Mixed		215 (4.52%)
Townhouse	858 (4.51%)	130 (2.73%)
No Information	485 (2.55%)	3,943 (82.85%)
	15,816 (83.08%)	
Psychological features		
Religious Practice: Frequent	927 (4.87%)	222 (4.67%)
Neighborhood Trust: High	17,325 (91.01%)	4,360 (91.62%)
Neighborhood Drug Use Concern: High	15,107 (79.36%)	139 (0.03%)
Speaks Non-English Language	576 (3.03%)	
Clinical features		
Tobacco Usage, 100 in lifetime	8,719 (45.80%)	2,188 (45.08%)
History of Cardiovascular Disease	4,699 (24.68%)	1108 (23.28%)
History of Kidney Disease	4,198 (22.05%)	1,044 (21.94%)
BMI, kg/m ²	33.52 (6.09)	33.53 (6.90)
Blood pressure		
Systolic, mm Hg	131.23 (14.95)	131.09 (14.46)
Diastolic, mm Hg	76.10 (10.83)	76.11 (12.75)
Heart rate, bpm	78.43 (12.25)	78.27 (12.33)
Medication history within 1 year of survey date		

Aspirin	5,677 (29.82%)	1,407 (29.57%)
Cholesterol Regulation Medications	3,944 (20.98%)	1,046 (21.98%)
Diabetes Medications		
Category 1 (non-insulin oral med + GLP1)	7,639 (40.13%)	
Category 2 (Category 1 + 1-shot BASAL)	1,348 (7.08%)	2,027 (42.59%) 298 (6.26%)
Category 3 (multiple daily injections)	2,554 (13.42%)	608 (12.78%)
Category 4 (no diabetes medications)	7,495 (39.37%)	1,826 (38.37%)
Statin	11,626 (61.07%)	2,893 (60.79%)
Anti-Hypertension Medications	14,376 (75.52%)	3,579 (75.20%)
Laboratory measurements		
Calcium (mg/dL)	9.5 (130.65)	9.5 (132.41)
Cholesterol in HDL (mg/dL)	47.91 (11.12)	47.91 (11.12)
Creatine (mmol/L)	58.11 (126.11)	59.82 (127.59)
Magnesium (mg/dL)	1.94 (0.15)	1.94 (0.15)
Potassium (mmol/L)	63.06 (128.11)	63.88 (129.67)
Triglyceride (mg/dL)	156.38 (87.91)	158.50 (100.16)
Total Cholesterol (mg/dL)	167.60 (39.01)	168.12 (39.65)
Hemoglobin A1c (%)	7.16 (1.48)	7.13 (1.45)
Non-HDL Cholesterol (mg/dL)	119.69 (37.61)	120.22 (38.16)
Glomerular filtration rate (mL/min/1.73 m ²)	78.62 (9.61)	78.94 (9.00)
Duration of CVD event		
< 1 year	6,090 (32.00%)	1,477 (31.03%)
1 year - 2 years	3,064 (16.10%)	799 (16.37%)
2 years - 3 years	5,506 (28.92%)	1,370 (28.79%)
3 years - 4 years	3,581 (18.81%)	943 (19.82%)
> 4 years	795 (4.18%)	190 (3.99%)
Outcomes		

Heart Failure	2,598 (13.65%)	604 (12.69%)
Stroke	1,656 (8.70%)	413 (6.68%)
Heart Attack	1,113 (5.85%)	273 (5.74%)
Total CVD	5,367 (28.19%)	1,290 (27.11%)

Model Explanations

Permutation variable importance plots are shown in Figure 1. The color of the bars represents the positive or negative impact of each predictor on the predicted CVD risk. When evaluated on all patients, CVD history is the most important predictor in our developed model, followed by kidney disease-related predictors (e.g. Creatinine, Calcium, History of Kidney Disease). When evaluated on only patients without CVD history, age appears to be the most important factor. Importantly, for both model types, socio-demographic factors such as employment status, housing types, education level, and sex are among the top predictors.

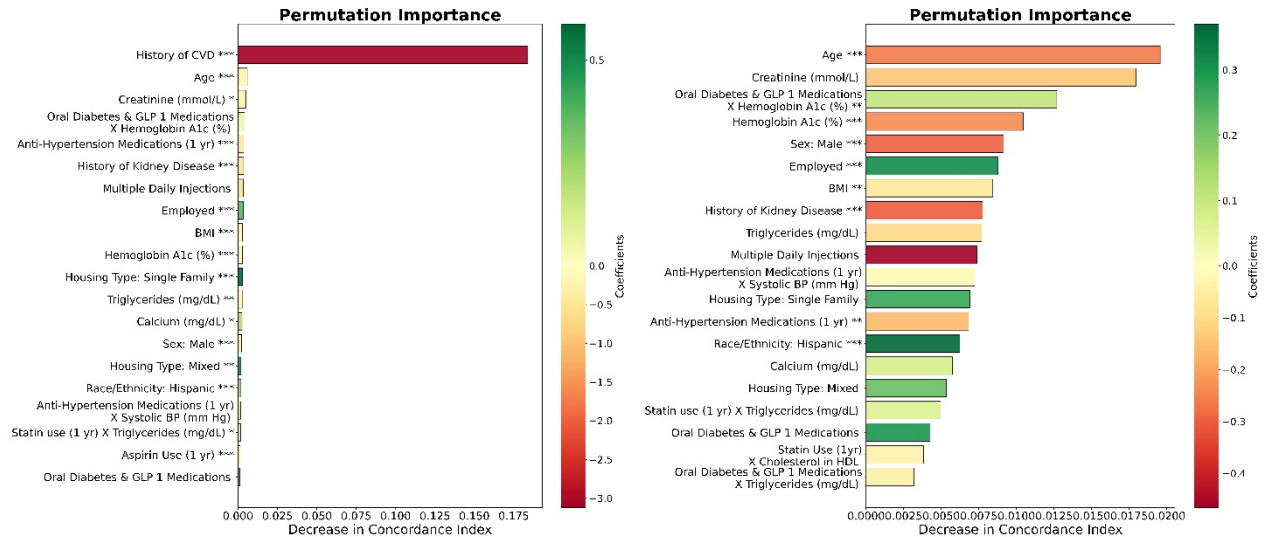


Figure 1. Permutation importance plots. Left: our Weibull AFT model when evaluated on all patients. Right: our Weibull AFT model when evaluated on patients without CVD history. Predictors are color-coded based on their variable coefficients to indicate whether they increase (red) or decrease (green) CVD risk. Note that the color scale is different in these two plots. Statistical significance of coefficients given by: * $P < 0.05$; ** $P < 0.01$; * $P < 0.001$.**

Model Assessment - Risk Calibration

Figure 2 (top) shows risk calibration plots for sex subgroups and race and ethnicity subgroups, indicating strong model calibration across the entire patient cohort. Additionally, our Weibull AFT model appears to have better calibration than PREVENT when testing on patients without CVD history (see Figure 2, bottom). Notably, PREVENT over-estimates risk of CVD events in

both analyses. We provide model coefficients and risk calculation examples for the Weibull AFT model in Supplement 2.

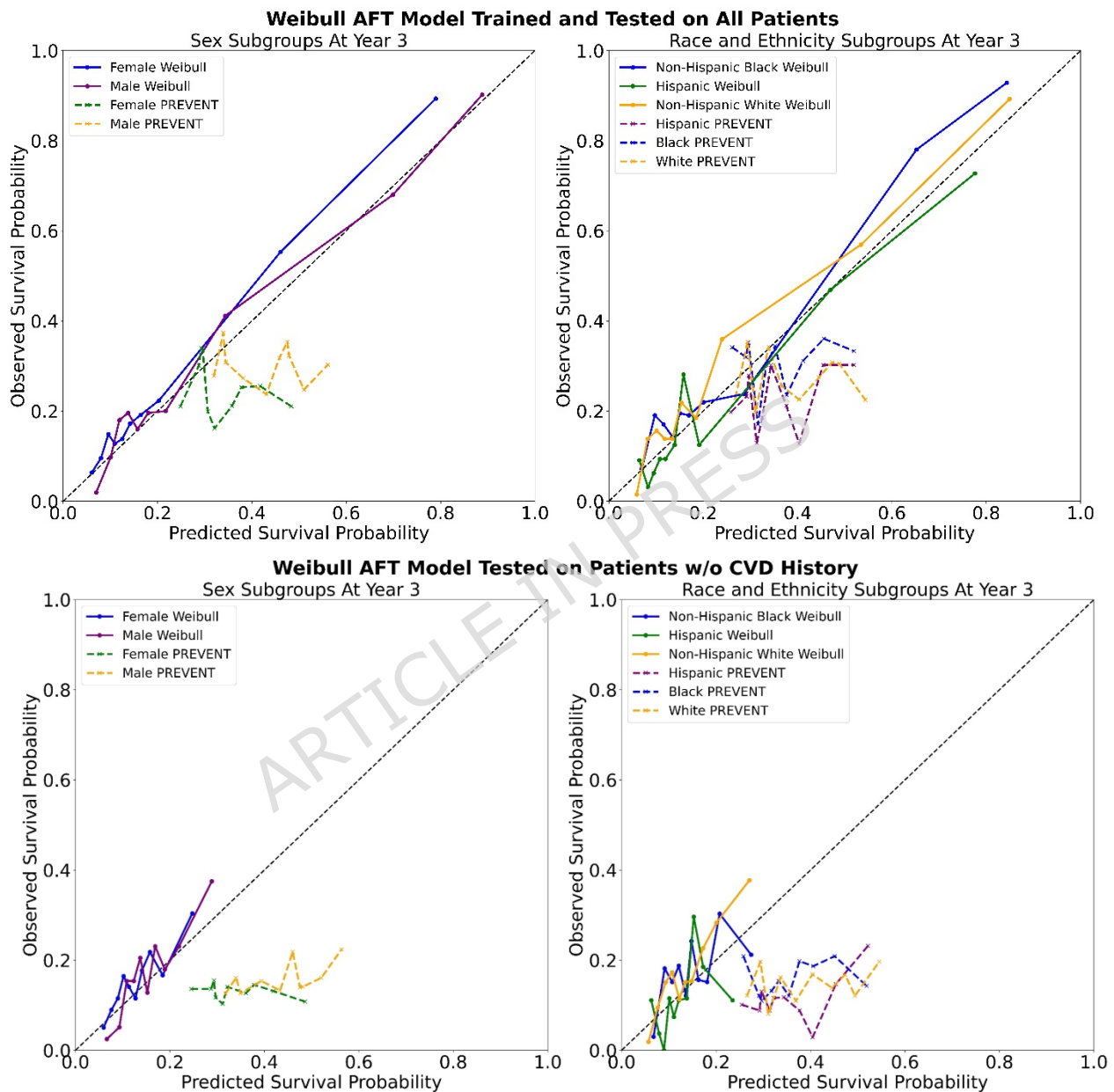


Figure 2. Three-year risk calibration plots for different sex subgroups and race and ethnicity subgroups comparing our

developed Weibull AFT model with PREVENT when (1) testing on all patients (top) and (2) when testing on patients without CVD history (bottom). Ideal calibration is marked with a black dashed line.

Model Assessment - Predictive Accuracy

When applied to the entire patient cohort in our testing set, our Weibull AFT model achieves a high C-index of 0.812, significantly outperforming PREVENT which achieves a C-index of 0.485. When comparing these models on the patients without CVD history in the testing set, the Weibull AFT model achieves a higher C-index of 0.646 compared to PREVENT's 0.465.

Model Assessment - Prediction Fairness

Table 2 summarizes the CF and CI achieved by our Weibull AFT model and the PREVENT equations on different subgroups of data. When applied to all patients in the testing data, the Weibull AFT model achieves CF values ranging from 0.795-0.826 and a CI of 0.006 for the sex subgroups and 0.031 for the race and ethnicity subgroups. These values indicate better fairness than PREVENT, whose CF values range from 0.551-0.592 and CI equaling 0.019 and 0.041 for sex and race and ethnicity subgroups, respectively. Among patients without CVD history, Weibull AFT still achieves greater CF values than PREVENT (0.610-0.674 vs. 0.541-0.600), although the differences between the two models are smaller than when tested on the entire patient

cohort. To this end, PREVENT achieved smaller CI values than the Weibull AFT for both sex (0.002 vs. 0.006) and race and ethnicity (0.058 vs. 0.065) subgroups, though these differences are small.

ARTICLE IN PRESS

Table 2. Comparison of subgroup fairness metrics between our Weibull AFT model trained on all patients and PREVENT which was developed on patients without CVD history for primary prevention when evaluated on all patients and on patients without CVD history. We exclude other ethnicities (e.g., Asian) as they represent only 6% of our studied cohort, which may introduce large variability.

Metric	Testing set	Model	Sex		Race and Ethnicity		
			Female	Male	Non-Hispanic Black	Hispanic	Non-Hispanic White
Concordance Fraction	All patients	Weibull	0.813	0.808	0.795	0.826	0.813
		PREVENT	0.574	0.554	0.551	0.592	0.559
	Patients w/o CVD history	Weibull	0.649	0.643	0.610	0.674	0.657
		PREVENT	0.568	0.569	0.541	0.600	0.572
Concordance Imparity	All patients	Weibull	0.006		0.031		
		PREVENT	0.019		0.041		
	Patients w/o CVD history	Weibull	0.006		0.065		
		PREVENT	0.002		0.058		

Discussion

Using NIH All of Us data, this study sought to investigate whether existing CVD risk prediction models can accurately and fairly estimate risk of CVD among patients with T2D. In particular, we developed a Weibull AFT model to estimate CVD risk among patients with T2D and compared its accuracy and fairness across sex and race and ethnicity subgroups to the PREVENT equations. Chiefly, our Weibull AFT model generally outperformed the established PREVENT equations on a large cohort of adults with T2D, both with and without history of CVD, in terms of calibration, predictive accuracy, and fairness on stratifications of this cohort by sex and race and ethnicity.

This superior performance was upheld even when the Weibull AFT model was trained without the use of race and ethnicity as a predictor (see Appendix C) and when it was trained only on the subset of patients without history of CVD (see eAppendix D).

Our results support the need for risk equations designed specifically for adults with T2D. While many CVD risk factors overlap between our cohort of patients with T2D and the general population, we note that markers of kidney disease (e.g. history of kidney disease, creatinine, and calcium levels) emerged as some of the most important predictors in our model for patients with T2D. This finding is particularly significant because - while isolated kidney disease may not be a strong predictor of CVD in the general population - it does play a significant role for CVD risk prediction among patients with T2D, a disease characterized by neurohormonal activation and chronic subclinical inflammation.²⁴⁻²⁹ Consistent with the current understanding of cardiorenal syndromes, kidney dysfunction in this population may contribute to increased risk of CVD through heightened cardiac remodeling, vascular injury, inflammatory toxicity, and neurohormonal activation, in addition to volume overload.^{25,30} The strong association between kidney disease and CVD risk in patients with T2D underscores the importance of monitoring renal function as part of CVD risk assessment in this population.

In further support of the need for T2D-specific risk equations, we found that the Weibull AFT model is better-calibrated among our cohort of T2D compared to PREVENT; PREVENT systematically over-estimated event risk, especially among lower-risk patients without prior CVD, while the Weibull AFT curves nearly trace the ideal line for both sex and race/ethnicity groups. This strong calibration is crucial for clinical decision support since accurately estimated probabilities may help avoid unnecessary interventions in lower-risk patients and ensure high-risk individuals are identified reliably.

We additionally emphasize that the Weibull AFT outperformed PREVENT in terms of predictive fairness. While the exact reason for this performance gap is difficult to pinpoint, we hypothesize that this result may be driven partly by the fact that the Weibull AFT was trained on NIH All of Us data, which over-represents under-represented racial groups compared to the 25 datasets used to derive PREVENT. Further, All of Us enabled us to include fine-grained socio-economic features (e.g. housing type, income levels, employment status, education backgrounds) in our model. Incorporating these detailed socio-economic factors allows us to capture the social determinants of health more effectively. To this end, our results confirm that these features achieve high feature importance, likely because they better characterize patients' accessibility to medical treatment and environments which support healthy lifestyle choices, supported by findings in the literature.³¹⁻³³ This suggests that socioeconomic factors play a critical role in

CVD risk and should be considered in risk prediction models to improve their accuracy and fairness.

Critically, our Weibull models included a patient's race and ethnicity as a predictor. Indeed, whether including race as a predictor promotes equity remains an open question.³⁴⁻³⁸ One perspective - which was adopted by PREVENT - is to remove race as a predictor since race is a social construct rather than biological predictor.¹⁰ However, this *fairness through unawareness* approach has been shown to be ineffective in some cases since race is often strongly correlated with other features such as income level, education level, and other socio-economic status.^{34,39,40} As such, we opted to investigate whether there was any utility in including race as a predictor and found that race is among the top predictors in terms of ensuring predictive accuracy. To ascertain these findings, we also performed a supplementary analysis (see eAppendix C) to determine if a Weibull AFT model which did not include race and ethnicity could perform comparably to our Weibull AFT model that did include race and ethnicity. We found that excluding race and ethnicity resulted in lower C-indices overall and lower CF values in nearly all patient subgroups compared to our Weibull AFT model that included race and ethnicity. In other words, our model including race and ethnicity is both more accurate overall and fairer for most patient subgroups included in the analysis. While these results suggest that it may be beneficial to include race and ethnicity as a predictor in CVD risk prediction models broadly, we caution

that these results are specific to our study data and modeling approach; additional investigation is warranted to ascertain how these findings extend to cohorts beyond our study.

Additionally, we found that when our model was evaluated on all patients, history of CVD events was one of the most important factors for ensuring accurate CVD risk prediction, far outweighing other variables. Our findings are supported by existing literature which suggest a prior CVD event increases the likelihood of future CVD events.⁴¹⁻⁴⁵ When our model was evaluated on patients without CVD history, age was the most important factor. This finding is consistent with previous literature suggesting that older individuals are at higher CVD risk.^{45,46}

Finally, we argue that CVD risk estimation models for people with T2D - including our own Weibull AFT model - have the potential to provide valuable clinical decision support for personalized CVD prevention and risk management.⁴⁷⁻⁵⁰ In particular, these models can quantitatively estimate an individual patient's CVD risk, which can help to characterize the magnitude of risk, support shared decision-making in the clinic, and reinforce adherence to preventive therapies. For example, T2D is often accepted as a CVD equivalent.⁵¹ Thus, for patients with T2D and no prior CVD history, CVD risk estimates can help guide statin intensity (e.g., moderate vs. high-intensity statin therapy), prompt consideration of treatment intensification in higher-

risk individuals, and facilitate discussions around comprehensive risk-factor modifications (e.g., lifestyle interventions, smoking cessation, and adherence to cardioprotective and renoprotective therapies).

While the size and diversity of the NIH All of Us dataset is certainly a strength of our analysis, we acknowledge that the generalizability of our results beyond this cohort may be limited because our model was evaluated using the unseen test split from the same All of Us dataset. In the future, we plan to validate our developed model using external datasets. Additionally, we did not include Apolipoprotein B (ApoB) in our analysis, even though it has been found to be a better predictor of CVD risk than LDL and non-HDL cholesterol.⁵² While approximately 1300 patients have an ApoB measurement in the entire All of Us dataset, our study sample specifically did not include any of these patients. Accordingly, we could not incorporate it in our analysis. Future studies should include ApoB if it is sufficiently available in the dataset. Lastly, our work focuses on time-to-first-event modeling where we censor the data when patients develop competing diseases (e.g. kidney diseases) before any CVD outcomes. We plan to further study how other diseases might affect subsequent CVD risk, such as increased doctor visits and changes in medications.

Conclusion

In conclusion, by leveraging the All of Us dataset, this study developed a Weibull AFT survival model for predicting the CVD risks of patients with T2D. Our model demonstrates better accuracy and fairness than the existing PREVENT equations across different demographic strata and serves as a foundation toward more equitable risk estimation of CVD among patients with T2D, supporting improved clinical decision-making and potentially reducing disparities in CVD outcomes.

Acknowledgments

We gratefully acknowledge *All of Us* participants for their contributions, without whom this research would not have been possible. We also thank the National Institutes of Health's [All of Us Research Program](#) for making available the participant data examined in this study.

Author contributions

YY and TL contributed equally to the data analysis, numerical analysis, tables and figures generation, manuscript writing. GPG conceptualized the study methodology and directed the methodological approach. GPG, EK, CL, SJL provided significant input on the methodology, analysis, technical interpretation of results, and manuscript writing. HS, MBW, and FJ contributed significantly on clinical interpretation and medical insights. All authors reviewed the manuscript.

Funding Declaration

This research is, in part, funded by the National Institutes of Health (NIH) AIM-AHEAD Program Agreement NO. 1OT2OD032581. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the NIH. HS, MBW, and FJ received funding from the Georgia Center for Diabetes and Translation Research under Award No. NIH/NIDDK P30DK111024.

Availability of Data and Materials

The dataset supporting the conclusions of this article is available from the All of Us Research Program. The data are not openly available. Access to the data requires registration, training, and compliance with the All of Us Research Program data usage policies. The underlying code for this study, including those used to extract and process the training/testing datasets and perform the analysis, is not publicly available but may be made available to qualified researchers on reasonable request from the corresponding author.

Declarations

Ethics Approval and Consent to Participate

All research in this manuscript has been performed in accordance with the Declaration of Helsinki and approved by the Institutional Review Board at the

Georgia Institute of Technology (Protocol No. H22333). This study analyzes data from the All of Us Research Program, which uses a centralized, electronic informed consent (eConsent) process. It is modular (Primary Consent, HIPAA Authorization, and Genomics Consent) and uses multimedia tools like short videos to ensure participant understanding. Most importantly, it includes a "teach-back" quiz that participants must pass to confirm they understand the voluntary nature of the study. The details of this consent process are provided on their web page (<https://allofus.nih.gov/article/all-us-consent-process>).

Consent for Publication

This manuscript has been read and its submission approved by all authors.

Competing Interests

FJP reported receiving grants through the institution from Insulet, Tandem Diabetes Care, Ideal Medical Technologies, Novo Nordisk, and Dexcom; receiving consulting fees from Dexcom; receiving consulting fees to the institution from Insulet. All other authors have no competing interests to declare.

References

1. Woodruff RC, Tong X, Khan SS, et al. Trends in Cardiovascular Disease Mortality Rates and Excess Deaths, 2010-2022. *Am J Prev Med*. 2024;66(4):582-589. doi:10.1016/j.amepre.2023.11.009
2. Trends and Disparities in Cardiovascular Mortality Among U.S. Adults With and Without Self-Reported Diabetes, 1988–2015 | Diabetes Care | American Diabetes Association. Accessed November 20, 2024. <https://diabetesjournals.org/care/article/41/11/2306/36539/Trends-and-Disparities-in-Cardiovascular-Mortality>
3. 2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk. doi:10.1161/01.cir.0000437741.48606.98
4. Karmali KN, Goff DC, Ning H, Lloyd-Jones DM. A systematic examination of the 2013 ACC/AHA pooled cohort risk assessment tool for atherosclerotic cardiovascular disease. *J Am Coll Cardiol*. 2014;64(10):959-968. doi:10.1016/j.jacc.2014.06.1136
5. Shah NS, Molsberry R, Rana JS, et al. Heterogeneous trends in burden of heart disease mortality by subtypes in the United States, 1999-2018: observational analysis of vital statistics. *BMJ*. 2020;370:m2688. doi:10.1136/bmj.m2688
6. Tsao CW, Aday AW, Almarzooq ZI, et al. Heart Disease and Stroke Statistics-2023 Update: A Report From the American Heart Association. *Circulation*. 2023;147(8):e93-e621. doi:10.1161/CIR.0000000000001123
7. Diaz CL, Shah NS, Lloyd-Jones DM, Khan SS. State of the Nation's Cardiovascular Health and Targeting Health Equity in the United States: A Narrative Review. *JAMA Cardiol*. 2021;6(8):963-970. doi:10.1001/jamacardio.2021.1137
8. Lloyd-Jones DM, Ning H, Labarthe D, et al. Status of Cardiovascular Health in US Adults and Children Using the American Heart Association's New "Life's Essential 8" Metrics: Prevalence Estimates From the National Health and Nutrition Examination Survey (NHANES), 2013 Through 2018. *Circulation*. 2022;146(11):822-835. doi:10.1161/CIRCULATIONAHA.122.060911
9. Buchholz EM, Rodday AM, Kolor K, Khoury MJ, de Ferranti SD. Prevalence and Predictors of Cholesterol Screening, Awareness, and Statin Treatment Among US Adults With Familial Hypercholesterolemia or Other Forms of Severe Dyslipidemia (1999-2014). *Circulation*. 2018;137(21):2218-2230. doi:10.1161/CIRCULATIONAHA.117.032321

10. Development and Validation of the American Heart Association's PREVENT Equations. doi:10.1161/CIRCULATIONAHA.123.067626
11. Novel Prediction Equations for Absolute Risk Assessment of Total Cardiovascular Disease Incorporating Cardiovascular-Kidney-Metabolic Health: A Scientific Statement From the American Heart Association. doi:10.1161/CIR.0000000000001191
12. New scientific research will test PREVENT risk calculator among diverse groups. American Heart Association. Accessed November 20, 2024. <https://newsroom.heart.org/news/new-scientific-research-will-test-prevent-risk-calculator-among-diverse-groups>
13. All of Us Research Hub. Accessed August 28, 2024. <https://www.researchallofus.org/>
14. Data Snapshots - All of Us Research Hub. Accessed August 28, 2024. <https://www.researchallofus.org/data-tools/data-snapshots/>
15. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk. *Circulation*. 2014;129(25_suppl_2):S49-S73. doi:10.1161/01.cir.0000437741.48606.98
16. Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. *Circulation*. 2016;133(6):601-609. doi:10.1161/CIRCULATIONAHA.115.017719
17. Prinja S, Gupta N, Verma R. Censoring in Clinical Trials: Review of Survival Analysis Techniques. *Indian J Community Med Off Publ Indian Assoc Prev Soc Med*. 2010;35(2):217-221. doi:10.4103/0970-0218.66859
18. Saikia R, Barman MP. A Review on Accelerated Failure Time Models.
19. Wei LJ. The accelerated failure time model: A useful alternative to the cox regression model in survival analysis. *Stat Med*. 1992;11(14-15):1871-1879. doi:10.1002/sim.4780111409
20. Fisher A, Rudin C, Dominici F. All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously.
21. Demler OV, Paynter NP, Cook NR. Tests of Calibration and Goodness of Fit in the Survival Setting. *Stat Med*. 2015;34(10):1659-1680. doi:10.1002/sim.6428
22. Zhang W, Weiss JC. Longitudinal Fairness with Censorship. *Proc AAAI Conf Artif Intell*. 2022;36(11):12235-12243. doi:10.1609/aaai.v36i11.21484

23. Zhang W. Fairness with Censorship: Bridging the Gap between Fairness Research and Real-World Deployment. *Proc AAAI Conf Artif Intell.* 2024;38(20):20. doi:10.1609/aaai.v38i20.30301
24. Coyle M, Flaherty G, Jennings C. A critical review of chronic kidney disease as a risk factor for coronary artery disease. *IJC Heart Vasc.* 2021;35:100822. doi:10.1016/j.ijcha.2021.100822
25. Jankowski J, Floege J, Fliser D, Böhm M, Marx N. Cardiovascular Disease in Chronic Kidney Disease. *Circulation.* Published online March 16, 2021. doi:10.1161/CIRCULATIONAHA.120.050686
26. Said S, Hernandez GT. The link between chronic kidney disease and cardiovascular disease. *J Nephrol Pathol.* 2014;3(3):99-104. doi:10.12860/jnp.2014.19
27. Saeed D, Reza T, Shahzad MW, et al. Navigating the Crossroads: Understanding the Link Between Chronic Kidney Disease and Cardiovascular Health. *Cureus.* 15(12):e51362. doi:10.7759/cureus.51362
28. Kotwal SS, Perkovic V. Kidney Disease as a Cardiovascular Disease Priority. *Circulation.* Published online September 24, 2024. doi:10.1161/CIRCULATIONAHA.124.068242
29. Daly C. Is early chronic kidney disease an important risk factor for cardiovascular disease?: A Background Paper prepared for the UK Consensus Conference on Early Chronic Kidney Disease. *Nephrol Dial Transplant.* 2007;22(suppl_9):ix19-ix25. doi:10.1093/ndt/gfm445
30. Quiroga B, Ortiz A, Navarro-González JF, Santamaría R, De Sequera P, Díez J. From cardiorenal syndromes to cardioneurology: a reflection by nephrologists on renocardiac syndromes. *Clin Kidney J.* 2023;16(1):19-29. doi:10.1093/ckj/sfac113
31. Schultz WM, Kelli HM, Lisko JC, et al. Socioeconomic Status and Cardiovascular Outcomes: Challenges and Interventions. *Circulation.* 2018;137(20):2166-2178. doi:10.1161/CIRCULATIONAHA.117.029652
32. Socioeconomic Deprivation: An Important, Largely Unrecognized Risk Factor in Primary Prevention of Cardiovascular Disease | *Circulation.* Accessed September 18, 2024. <https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.122.060042>
33. Yeager R, Riggs DW, DeJarnett N, et al. Association Between Residential Greenness and Cardiovascular Disease Risk. *J Am Heart Assoc.* 2018;7(24):e009117. doi:10.1161/JAHA.118.009117

34. Xu J, Xiao Y, Wang WH, et al. Algorithmic fairness in computational medicine. *eBioMedicine*. 2022;84:104250. doi:10.1016/j.ebiom.2022.104250
35. Chen RJ, Wang JJ, Williamson DFK, et al. Algorithm fairness in artificial intelligence for medicine and healthcare. *Nat Biomed Eng*. 2023;7(6):719-742. doi:10.1038/s41551-023-01056-8
36. Varga TV. Algorithmic fairness in cardiovascular disease risk prediction: overcoming inequalities. *Open Heart*. 2023;10(2):e002395. doi:10.1136/openhrt-2023-002395
37. Paulus JK, Kent DM. Predictably unequal: understanding and addressing concerns that algorithmic clinical prediction may increase health disparities. *Npj Digit Med*. 2020;3(1):1-8. doi:10.1038/s41746-020-0304-9
38. Saghafian S, Kent DM, Goel S. A Framework for Considering the Value of Race and Ethnicity in Estimating Disease Risk. *Ann Intern Med Res Report Methods*.
39. Dwork C, Hardt M, Pitassi T, Reingold O, Zemel R. Fairness through awareness. In: *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*. ITCS '12. Association for Computing Machinery; 2012:214-226. doi:10.1145/2090236.2090255
40. Grgic-Hlac̆a N, Zafar MB, Gummadi KP, Weller A. The Case for Process Fairness in Learning: Feature Selection for Fair Decision Making.
41. Basu S, Sussman JB, Berkowitz SA, Hayward RA, Yudkin JS. Development and validation of Risk Equations for Complications Of type 2 Diabetes (RECODe) using individual participant data from randomised trials. *Lancet Diabetes Endocrinol*. 2017;5(10):788-798. doi:10.1016/S2213-8587(17)30221-8
42. Basu S, Sussman JB, Berkowitz SA, et al. Validation of Risk Equations for Complications of Type 2 Diabetes (RECODe) Using Individual Participant Data From Diverse Longitudinal Cohorts in the U.S. *Diabetes Care*. 2017;41(3):586-595. doi:10.2337/dc17-2002
43. Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association - PubMed. Accessed September 18, 2024. <https://pubmed.ncbi.nlm.nih.gov/28122885/>
44. Bhatt DL, Eagle KA, Ohman EM, et al. Comparative determinants of 4-year cardiovascular event rates in stable outpatients at risk of or with

- atherothrombosis. *JAMA*. 2010;304(12):1350-1357.
doi:10.1001/jama.2010.1322
45. Yang Y, Liao CY, Keyvanshokoh E, et al. A Responsible Framework for Assessing, Selecting, and Explaining Machine Learning Models in Cardiovascular Disease Outcomes Among People With Type 2 Diabetes: Methodology and Validation Study. *JMIR Med Inform*. 2025;13:e66200. doi:10.2196/66200
46. Kim K il. Risk Stratification of Cardiovascular Disease according to Age Groups in New Prevention Guidelines: A Review. *J Lipid Atheroscler*. 2023;12(2):96-105. doi:10.12997/jla.2023.12.2.96
47. Khan SS, Breathett K, Braun LT, et al. Risk-Based Primary Prevention of Heart Failure: A Scientific Statement From the American Heart Association. *Circulation*. 2025;151(20). doi:10.1161/CIR.0000000000001307
48. Garcia GGP, Steimle LN, Marrero WJ, Sussman JB. Interpretable Policies and the Price of Interpretability in Hypertension Treatment Planning. *Manuf Serv Oper Manag*. 2024;26(1):80-94. doi:10.1287/msom.2021.0373
49. Liao CY, Keyvanshokoh E, Garcia GG. Constraint-Aware Self-Improving Large Language Model for Clinical Role Model Generation. *SSRN*. Preprint posted online 2025. doi:10.2139/ssrn.5642250
50. Cao J, Keyvanshokoh E, Liu T. Safe Reinforcement Learning with Contextual Information: Theory and Application to Personalized Comorbidity Management. *SSRN Electron J*. Published online 2023. doi:10.2139/ssrn.4583667
51. Kuusisto J, Laakso M. Update on Type 2 Diabetes as a Cardiovascular Disease Risk Equivalent. *Curr Cardiol Rep*. 2013;15(2):331. doi:10.1007/s11886-012-0331-5
52. Pencina MJ, D'Agostino RB, Zdrojewski T, et al. Apolipoprotein B improves risk assessment of future coronary heart disease in the Framingham Heart Study beyond LDL-C and non-HDL-C. *Eur J Prev Cardiol*. 2015;22(10):1321-1327. doi:10.1177/2047487315569411