Title

Incorporation of natriuretic peptides with clinical risk scores to predict heart failure among individuals with dysglycaemia.

Journal

European Journal of Heart Failure, 24(1)

Authors

Segar, Matthew
Khan, Muhammad
Patel, Kershaw
et al.

Publication Date

2022

DOI

10.1002/ejhf.2375

Peer reviewed

HHS Public Access

Author manuscript

Eur J Heart Fail. Author manuscript; available in PMC 2023 September 28.

Published in final edited form as:

Eur J Heart Fail. 2022 January ; 24(1): 169–180. doi:10.1002/ejhf.2375.

Incorporation of natriuretic peptides with clinical risk scores to predict heart failure among individuals with dysglycaemia

Matthew W. Segar[1,2], **Muhammad Shahzeb Khan**[3], **Kershaw V. Patel**[4], **Muthiah Vaduganathan**[5], **Vaishnavi Kannan**[6], **Duwayne Willett**[1,6], **Eric Peterson**[1], **W.H. Wilson Tang**[7], **Javed Butler**[3], **Brendan M. Everett**[5], **Gregg C. Fonarow**[8], **Thomas J. Wang**[1], **Darren K. McGuire**[1], **Ambarish Pandey**[1,*]

[1]Division of Cardiology, Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, TX, USA

[2]Department of Cardiology, Texas Heart Institute, Houston, TX, USA

[3]Division of Cardiology, Duke University School of Medicine, Durham, NC, USA

[4]Department of Cardiology, Houston Methodist DeBakey Heart and Vascular Center, Houston, TX, USA

[5]Brigham and Women's Hospital Heart and Vascular Center, Department of Medicine, Harvard Medical School, Boston, MA, USA

[6]Clinical Informatics Center, University of Texas Southwestern Medical Center, Dallas, TX, USA

[7]Department of Cardiovascular Medicine, Cleveland Clinic, Cleveland, OH, USA

[8]Division of Cardiology, Ronald Reagan UCLA Medical Center, Ahmanson-UCLA Cardiomyopathy Center, Los Angeles, CA, USA

[*]Corresponding author. Division of Cardiology, Department of Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA. Tel: +1 214 645-7521, Fax: +1 214 648-9639, ambarish.pandey@utsouthwestern.edu.

Conflict of interest:

M.W.S. has received non-financial support from Pfizer and Merck. M.V. has received research grant support or served on advisory boards for American Regent, Amgen, AstraZeneca, Bayer AG, Baxter Healthcare, Boehringer Ingelheim, Cytokinetics, Lexicon Pharmaceuticals, Roche Diagnostics, and Relypsa, and participates in clinical endpoint committees for studies sponsored by Galmed, Novartis, and participates in speaking engagements for Novartis and Roche Diagnostics. W.H.W.T. has consulted for Sequana Medical AG, Relypsa, PreCardia, CardiolRx Inc, and Owkin Inc., and served as exam writing committee member for American Board of Internal Medicine. B.M.E. has consulted for Amgen, Amarin, Gilead, Johnson & Johnson, Lily, Provention, Merck, Roche Diagnostics, American Heart Association, and UpToDate, and has grant support from PCORI, NHLBI, and Novo Nordisk. G.C.F. has consulted for Abbott, Amgen, Bayer, Cytokinetics, Edwards, Janssen, Medtronic, Merck, and Novartis. D.K.M. has had leadership roles in clinical trials for AstraZeneca, Boehringer Ingelheim, Eisai, Esperion, GlaxoSmithKline, Janssen, Lexicon, Merck & Co., Inc., Novo Nordisk, CSL Behring, and Sanofi USA; and has received consultancy fees from AstraZeneca, Bayer AG, Boehringer Ingelheim, Lilly USA, Merck & Co., Inc., Pfizer, Novo Nordisk, Metavant, Afimmune, and Sanofi. A.P. has served on the advisory board of Roche Diagnostics, has received non-financial support from Pfizer and Merck, and has received research support from the Texas Health Resources Clinical Scholarship, the Gilead Sciences Research Scholar Program, the National Institute on Aging GEMSSTAR Grant (1R03AG067960-01), and Applied Therapeutics. J.B reports honoraria from Abbott, Adrenomed, Array, Amgen, Applied Therapeutics, Astra Zeneca, Bayer, Boehringer Ingelheim, CVRx, Eli Lilly, G3 Pharma, Imbria, Impulse Dynamics, Innolife, Janssen, LivaNova, Luitpold, Medtronic, Merck, Novartis, Novo Nordisk, Relypsa, Sequana Medical, V-Wave Limited, Vifor. All other authors have nothing to disclose.

Supplementary Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Abstract

Aims—To evaluate the performance of the WATCH-DM risk score, a clinical risk score for heart failure (HF), in patients with dysglycaemia and in combination with natriuretic peptides (NPs).

Methods and results—Adults with diabetes/pre-diabetes free of HF at baseline from four cohort studies (ARIC, CHS, FHS, and MESA) were included. The machine learning-[WATCH-DM(ml)] and integer-based [WATCH-DM(i)] scores were used to estimate the 5-year risk of incident HF. Discrimination was assessed by Harrell's concordance index (C-index) and calibration by the Greenwood–Nam–D'Agostino (GND) statistic. Improvement in model performance with the addition of NP levels was assessed by C-index and continuous net reclassification improvement (NRI). Of the 8938 participants included, 3554 (39.8%) had diabetes and 432 (4.8%) developed HF within 5 years. The WATCH-DM(ml) and WATCH-DM(i) scores demonstrated high discrimination for predicting HF risk among individuals with dysglycaemia (C-indices = 0.80 and 0.71, respectively), with no evidence of miscalibration (GND $P \geq 0.10$). The C-index of elevated NP levels alone for predicting incident HF among individuals with dysglycaemia was significantly higher among participants with low/intermediate (<13) vs. high (≥13) WATCH-DM(i) scores [0.71 (95% confidence interval 0.68–0.74) vs. 0.64 (95% confidence interval 0.61–0.66)]. When NP levels were combined with the WATCH-DM(i) score, HF risk discrimination improvement and NRI varied across the spectrum of risk with greater improvement observed at low/intermediate risk [WATCH-DM(i) <13] vs. high risk [WATCH-DM(i) ≥13] (C-index = 0.73 vs. 0.71; NRI = 0.45 vs. 0.17).

Conclusion—The WATCH-DM risk score can accurately predict incident HF risk in community-based individuals with dysglycaemia. The addition of NP levels is associated with greater improvement in the HF risk prediction performance among individuals with low/intermediate risk than those with high risk.

Keywords

Diabetes; Pre-diabetes; Biomarkers; Risk stratification; Risk prediction; Heart failure

Introduction

Diabetes affects approximately 60 million adults in Europe and confers a more than two-fold increased risk for developing heart failure (HF).[1–3] Besides diabetes, elevated blood glucose level below the diagnostic thresholds for diabetes — commonly referred to as pre-diabetes — is also common and associated with greater HF risk.[4] Moreover, up to 25% of individuals with pre-diabetes may progress to overt type 2 diabetes within 3–5 years.[5]

Sodium–glucose co-transporter 2 inhibitors and intensive lifestyle modifications have recently been identified as effective therapies for preventing cardiovascular disease (CVD), particularly HF, in individuals with diabetes.[6,7] Implementation of such treatments in individuals with dysglycaemia, who have early stages of cardiometabolic dysregulation, could be an effective population-level approach to mitigate the risk of HF. Thus, risk stratification strategies that can efficiently and accurately identify the highest risk individuals with pre-diabetes or diabetes are crucial for population-based HF prevention approaches.

Several risk scores have been developed and validated for HF risk prediction among community-dwelling individuals.[4,8–11] Recently, risk prediction models specific to diabetes, such as the WATCH-DM risk score that incorporates readily available clinical data to predict the 5-year risk of incident HF, have been developed and validated.[12] However, WATCH-DM has not been evaluated in community-based cohorts of patients with pre-diabetes and does not incorporate prognostic cardiac biomarkers. This is particularly relevant considering the well-established role of abnormally elevated cardiac biomarkers, such as natriuretic peptides (NPs), as predictors of HF risk among individuals with dysglycaemia independent of clinical risk factors.[13,14] Thus, this study aimed to validate the WATCH-DM risk score among a community-based cohort of adults with diabetes and pre-diabetes. Additionally, we assessed the incremental value of adding NP levels to the WATCH-DM risk score for HF risk prediction.

Methods

Study population

Deidentified participant-level data from four community-based cohort studies obtained from the National Institute of Health Biologic Specimen and Data Repository Coordinating Center (BioLINCC) were pooled: Atherosclerosis Risk in Communities (ARIC), Cardiovascular Health Study (CHS), Framingham Offspring Study (FHS), and Multi-Ethnic Study of Atherosclerosis (MESA). The design of each study has been previously reported and described in the Methods in online supplementary Appendix S1.[15–19] ARIC visit 5, CHS baseline exam, FHS exam 6, and MESA baseline exam were used as baseline encounters. We included participants with pre-diabetes or diabetes free of HF. Participants missing either >20% of the WATCH-DM risk score covariates (n = 107), outcome data (n = 15), or biomarkers (n = 1172) were excluded (online supplementary Figure S1). Each participant provided written informed consent. Study approval was obtained at the coordinating centre for each cohort. The present analysis was considered exempt from Institutional Review Board approval at the University of Texas Southwestern Medical Center, Dallas, Texas.

Definition of diabetes and pre-diabetes and other clinical covariates

Diabetes and pre-diabetes status was defined at the baseline visit of each cohort using established criteria as reported previously and detailed in the Methods in online supplementary Appendix S1.[15–17,19] Baseline covariates were assessed using standardized protocols as previously described and harmonized across cohorts (Methods in online supplementary Appendix S1).[15–17,19] Missing data were imputed using random forest imputation after excluding participants with >20% missingness.[20]

Measurement of cardiac biomarkers

Details of NP level, high-sensitivity cardiac troponin (hs-cTn), high-sensitivity C-reactive protein (hs-CRP), and left ventricular hypertrophy (LVH) on electrocardiogram (ECG), and biomarker assessment are described in the Methods in online supplementary Appendix S1. Elevated NP levels were body mass index (BMI)-specific as follows: N-terminal pro-B-type natriuretic peptide (NT-proBNP) ≥125 pg/ml or B-type natriuretic peptide (BNP) ≥40 pg/ml

for BMI ≤30 kg/m^2; NT-proBNP ≥100 pg/ml or BNP ≥30 pg/ml for BMI >30 kg/m^2.[21] Consistent with prior studies, the following cutoffs were used to identify abnormal elevation in biomarkers: hs-cTn (hs-cTnT for ARIC and MESA and hs-cTnI in FOS) ≥6 ng/L and hs-CRP ≥3 mg/L.[4,22,23] Presence of LVH on ECG was determined using Sokolow–Lyon criteria.[24]

Outcome of interest

The primary outcome of interest was incident HF, an outcome adjudicated by an independent panel of investigators for each study. In ARIC, potential HF hospitalizations were first identified from hospital discharge records or death certificates that indicated HF with International Classification of Diseases (ICD), 9/10th Revision codes 428x. HF events were subsequently adjudicated based on clinical history, imaging, and medication use at the time of hospitalization.[25,26] In CHS, potential HF events were identified at semi-annual visits. HF events were confirmed by reviewing medical records for medication use, physical exam, or imaging findings of HF.[27,28] In FHS, medical records were obtained from all hospitalizations and physician visits related to cardiovascular disease. Events were assessed using the FHS criteria as previously described.[29] In MESA, participants were contacted semi-annually to inquire about HF diagnoses with subsequent review of medical records to confirm HF events.[30] HF events were initially adjudicated by two paired physicians with disagreements rectified by a review of the full committee. HF subtype data were available in participants from the ARIC and MESA cohorts. HF with reduced ejection fraction (HFrEF) was defined as an ejection fraction <50% at the time of diagnosis and HF events with an ejection fraction ≥50% identified HF with preserved ejection fraction (HFpEF). Similar to our approach, several prior studies have pooled data from these cohorts to evaluate the epidemiology of incident HF.[8,31–33]

Clinical risk scores of interest

WATCH-DM risk score—The derivation of the WATCH-DM risk score has been previously described.[12] Briefly, the WATCH-DM risk score was initially developed using data from the Action to Control Cardiovascular Risk in Diabetes (ACCORD) and externally validated among participants with diabetes from the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT).[34,35] The WATCH-DM risk score includes 10 clinical, laboratory, and ECG variables [age, BMI, systolic blood pressure (BP), diastolic BP, serum creatinine, fasting plasma glucose, high-density lipoprotein cholesterol, QRS duration, history of myocardial infarction, history of coronary artery bypass graft]. The risk score includes machine learning- and integer-based models [referred to as WATCH-DM(ml) and WATCH-DM(i), respectively]. Both the WATCH-DM(ml) and WATCH-DM(i) models use the same set of covariates and differ only in the modelling strategy to estimate HF risk. The WATCH-DM(ml) model has been updated by pooling data from multiple cohorts (ACCORD, ALLHAT diabetes cohort, Look AHEAD cohort, and electronic health record based diabetes registry) since its original derivation.[34–36] The WATCH-DM(i) model used in this study is the same as that which was derived in the ACCORD cohort.

PCP-HF score—The PCP-HF score incorporates nine variables (age, systolic BP, BMI, total cholesterol, high-density lipoprotein cholesterol, QRS duration, smoking status, use

of antihypertensive medication, use of diabetes medication).[8] The score calculates an estimated 10-year risk of incident HF in the general population free of baseline CVD using Cox proportional hazard regression. Separate equations are provided for each race-sex group. Since WATCH-DM estimates the 5-year risk of incident HF, the PCP-HF equation coefficients were recalculated in the overall cohort to predict 5-year HF risk.

Statistical analysis

Performance of the WATCH-DM risk score—The performance of the WATCH-DM risk score was assessed separately among participants with overall dysglycaemia (diabetes/pre-diabetes), diabetes, and pre-diabetes. HF outcomes were censored at 5 years. Discrimination was assessed by Harrell's C-index and calibration by the Brier score and the Greenwood–Nam–D'Agostino (GND) chi-square statistic.[37–39] Participants in the overall cohort were categorized into 'low/intermediate' and 'high' WATCH-DM risk score groups (below and above the optimal cutoff respectively) determined by the maximum value of Youden's index on receiver operating characteristic analysis.[40] The cumulative risk of incident HF was assessed across low/intermediate vs. high WATCH-DM categories.

Improvement in heart failure risk prediction with NT-proBNP levels—Analyses were conducted with categorical measures of biomarker levels (above vs. below cutoff). First, separate plots of the change in C-index (discrimination performance) and Brier score (calibration performance) with the addition of NP levels to the WATCH-DM(i) score were generated in overall, diabetes, and pre-diabetes cohorts. The mean and 95% confidence interval (CI) were plotted as a continuous variable with locally estimated scatterplot smoothing (LOESS) curves across WATCH-DM(i) scores using bootstrapping with 1000 replicates. Second, the clinical net benefit of adding NP levels across increasing risk thresholds was assessed using decision curves — a measure of the number of true positive cases identified without an increase in the false-positive rate.[41] Third, the discrimination and calibration performance of NP levels alone was assessed separately in high vs. low/intermediate WATCH-DM(i) categories. Fourth, the risk prediction indices for discrimination (C-index), calibration (Brier score), net reclassification improvement (NRI) [categorical (<2%, 2%–5%, 5.1%–15%, and >15% risk groups) and continuous], and integrated discrimination improvement (IDI) were also compared across the following models: WATCH-DM(i) only vs. WATCH-DM(i) + NP level assessment among participants with low/intermediate WATCH-DM risk scores only vs. WATCH-DM(i) + NP level assessment among participants with high WATCH-DM risk scores only. The improvement in C-indices was compared using the DeLong test.[42,43] Differences in NRI and IDI estimates were compared using Welch's *t*-test comparison of means. Similar analysis was also performed to assess the improvement in model performance indices for the WATCH-DM(ml) model with incorporation of NP levels at different risk thresholds (low/intermediate vs. high risk).

Sensitivity analysis—Several sensitivity and subgroup analyses were performed in the overall cohort to assess the robustness of our study findings. First, the performance of the WATCH-DM models was also compared with other established HF clinical risk scores to predict HF — the PCP-HF risk score, ARIC-HF risk score, and MESA-HF risk score in the

overall cohort.[8,10,44] Second, the performance of the WATCH-DM risk score was assessed separately in sex- (men vs. women) and race-based (Black vs. non-Black) subgroups in the overall cohort. Third, the performance of the WATCH-DM risk score to predict incident HFpEF and HFrEF outcomes was also assessed in the subset with available data on HF subtype adjudication in BioLINCC (MESA and ARIC). Fourth, sensitivity analysis was also performed to evaluate if incorporation of sex and substitution of creatinine with estimated glomerular filtration rate modifies the performance of WATCH-DM models. Fifth, the performance of the WATCH-DM models in the absence of ECG-based covariates that may not be commonly available in routine clinical setting was assessed. Sixth, the change in HF risk prediction performance among individuals with dysglycaemia with the incorporation of NP levels to WATCH-DM risk score was also evaluated among men and women separately. Seventh, the change in HF risk prediction performance among individuals with dysglycaemia was also evaluated with the incorporation of NP levels to another HF risk score — the PCP-HF score.[8] Since PCP-HF was developed in a primary prevention cohort, patients with CVD at baseline were excluded. High and low/intermediate PCP-HF scores were defined using Youden's index. Finally, change in performance of the WATCH-DM risk score was also assessed with incorporation of other relevant biomarkers such as hs-cTn (elevated vs. not elevated), hs-CRP (elevated vs. not elevated), and presence of LVH on ECG. Analyses were performed using R 3.6.3 (R Foundation, Vienna, Austria) with a p-value of <0.05 indicating significance.

Results

Among 8938 participants in the overall cohort, 3554 (39.8%) had diabetes and 5384 (60.2%) had pre-diabetes at baseline (online supplementary Figure S1). Baseline characteristics of participants with dysglycaemia who were free of HF and were included vs. excluded from the analysis are shown in online supplementary Table S1. Participants excluded vs. included from the present study had a higher burden of risk factors and CVD. Over 5-year follow-up, 245 (6.9%) participants with diabetes and 187 (3.5%) participants with pre-diabetes developed incident HF (incident rate of HF = 15.7 and 7.5 events 1000 person-years, respectively).

Validation of the WATCH-DM risk score among participants with diabetes or pre-diabetes

Among participants with diabetes, the WATCH-DM(ml) demonstrated the highest discrimination with C-index of 0.76 (95% CI 0.73–0.80) and adequate calibration (p = 0.12) (Table 1 and online supplementary Figure S2). Similar performance of WATCH-DM(ml) was observed among participants with pre-diabetes (C-index = 0.83, 95% CI 0.80–0.85; GND p = 0.14), and in the dysglycaemia cohort pooling participants with diabetes and pre-diabetes together (C-index = 0.80, 95% CI 0.78–0.81; GND p = 0.10) (Table 1 and online supplementary Figure S2).

The median WATCH-DM(i) score among participants with diabetes and pre-diabetes was 12 [interquartile range (IQR) 10–14] and 11 (IQR 9–12), respectively, with an observed range of 5–25. The discriminative ability of WATCH-DM(i) among participants with diabetes, pre-diabetes, and in the overall cohort was adequate with C-indices of 0.69 (95% CI 0.67–0.71),

0.72 (95% CI 0.69–0.74), and 0.71 (95% CI 0.68–0.73), respectively (Table 1 and online supplementary Figure S2). WATCH-DM(i) also demonstrated no evidence of miscalibration across glycaemic groups (GND $p > 0.10$ for all).

Risk prediction performance of natriuretic peptide levels across WATCH-DM risk strata

Considering the robust performance of the WATCH-DM risk score in diabetes and pre-diabetes subgroups, the contribution of NP levels to HF risk prediction across WATCH-DM(i) risk strata was assessed in the dysglycaemia cohort of participants. The optimal cut-point of WATCH-DM(i) for predicting HF, as determined by the Youden index, was 13. Of the 8938 participants in the overall cohort, 32% had high WATCH-DM(i) scores (≥13), of which 8.4% developed incident HF on follow-up. In contrast, the 5-year incidence of HF in the low/intermediate WATCH-DM(i) score category was 3.2% (Figure 1A).

Baseline characteristics of the overall cohort stratified by low/intermediate vs. high WATCH-DM(i) score and NP levels are shown in online supplementary Table S2. The C-index of elevated NP levels alone for predicting incident HF in the overall cohort was 0.69 (95% CI 0.67–0.72) with a significantly higher discrimination performance noted among participants with low/intermediate vs. high WATCH-DM(i) scores [C-index = 0.71 (95% CI 0.68–0.74) vs. 0.64 (95% CI 0.61–0.66); DeLong $p < 0.001$] (online supplementary Table S3). Elevated NP levels also demonstrated better risk stratification among individuals with low/intermediate vs. high WATCH-DM(i) scores. Among participants with low/intermediate WATCH-DM(i) scores, there was a near six-fold gradient in the 5-year HF incidence between those with elevated (8.2%) vs. not elevated NP levels (1.4%) (Figure 1B). In contrast, among participants with high WATCH-DM(i), there was only a three-fold gradient in HF incidence across the NP level strata (4.7% vs. 14.1% in low vs. high NP groups). A similar pattern of results was observed with greater HF risk discrimination and stratification by NP levels among participants with low/intermediate vs. high WATCH-DM(i) scores in subgroups of participants with diabetes and pre-diabetes analysed separately (online supplementary Table S3).

Improvement in WATCH-DM(i) model performance with addition of natriuretic peptide levels

Addition of NP levels to the WATCH-DM(i) score significantly improved the overall C-index for predicting incident HF to 0.76 (95% CI 0.74–0.78; DeLong $p < 0.001$) with no worsening in calibration (GND $p = 0.41$). Notably, the improvement in HF risk discrimination with NP levels varied across the spectrum of the WATCH-DM(i) score. A greater improvement in the C-index was observed with the addition of NP levels among participants with low/intermediate WATCH-DM(i) scores with degradation in risk discrimination at higher WATCH-DM(i) scores (Figure 2A). Specifically, the addition of NP levels significantly increased the C-index for HF by 0.05 at a WATCH-DM(i) score of 7 compared with no improvement in the C-index at a WATCH-DM(i) score of 20. Calibration (Brier score) was also improved with the addition of NP levels at lower vs. higher WATCH-DM(i) scores (Figure 2B). Similar results were observed in subgroup analysis limited to participants with diabetes or pre-diabetes (online supplementary Figure S3).

In decision curve analysis, the addition of elevated NP levels to the WATCH-DM(i) score increased the number of HF events detected by 4 per 1000 participants compared with the WATCH-DM(i) score alone at an event rate of 8% (Figure 3A). However, no benefit was observed with the addition of elevated NP levels to the WATCH-DM(i) scores above a 5-year HF risk threshold of 11%. Similar results were observed in subgroup analysis among participants with diabetes and pre-diabetes (online supplementary Figure S4).

Optimal risk-based approach of incorporating natriuretic peptide levels with WATCH-DM(i)

The model performance indices for risk prediction approaches using WATCH-DM(i) only vs. WATCH-DM(i) + NP levels among participants with low/intermediate risk vs. WATCH-DM(i) + NP levels among participants with high risk are shown in Table 2. Among individuals with dysglycaemia, compared with WATCH-DM(i) alone (C-index = 0.71, 95% CI 0.68–0.73); Brier score: 5.1), a risk prediction model that combined WATCH-DM(i) with NP level assessment only in low/intermediate risk participants (WATCH-DM(i) score <13) had significantly improved discrimination [C-index = 0.73, 95% CI 0.71–0.75; DeLong $p < 0.001$) and model calibration (Brier score: 2.2). In contrast, a risk prediction model combining WATCH-DM(i) with NP level assessment only among participants with high risk (scores ≥13) demonstrated no improvement in discrimination (C-index = 0.71, 95% CI 0.69–0.73; DeLong $p = 0.34$) and reduced model calibration (Brier score: 24.7) (Table 2). Furthermore, a significantly greater improvement in the NRI (both continuous and categorical) was observed by combining the WATCH-DM(i) score with selective NP level assessment in participants with low/intermediate (score <13) vs. high risk (score ≥13) (mean categorical NRI = 0.36 vs. 0.08; $p < 0.001$; mean continuous NRI = 0.45 vs. 0.17; $p < 0.001$) (Table 3). Similar improvement in model performance was observed with use of risk prediction approaches combining the WATCH-DM(i) score with NP level assessment among participants with low/intermediate risk and not high risk in subgroup analysis of participants with diabetes and pre-diabetes separately (Tables 2 and 3).

Sensitivity analyses

Comparison of WATCH-DM models to other heart failure risk scores—Among participants with dysglycaemia, the WATCH-DM(ml) demonstrated superior performance in discrimination, calibration, and clinical utility than other well established risk scores (vs. PCP-HF, MESA-HF, and ARIC-HF) (Figure 3B and online supplementary Table S4). In decision curve analysis, compared to established risk scores, WATCH-DM(ml) identified an additional 18–23 HF events per 1000 patients. WATCH-DM(i) had comparable discrimination, but superior calibration and clinical utility compared with the other scores. In decision curve analysis, WATCH-DM(i) identified an additional 8–13 HF events per 1000 patients.

Performance of WATCH-DM models in different settings—Multiple subgroup and sensitivity analyses were performed to assess the robustness of the WATCH-DM model performance in the overall cohort. In subgroup analysis stratified by sex and race, WATCH-DM models demonstrated good discrimination and no evidence of miscalibration among sex- and race-based subgroups (online supplementary Table S5). Among HF subtypes, the WATCH-DM models demonstrated better performance in predicting HFrEF vs. HFpEF

(online supplementary Table S6). Additionally, no change in the performance of WATCH-DM models was noted in the setting of missing ECG covariate, or with inclusion of sex as a covariate. Substituting estimated glomerular filtration rate for serum creatinine led to a modest decline in the model performance (online supplementary Table S7).

Incorporation of other biomarkers to the WATCH-DM(i) model—The addition of elevated hs-cTn, hs-CRP, or LVH to WATCH-DM(i) did not improve discrimination or calibration compared with the WATCH-DM(i) model alone (online supplementary Table S8). In contrast, incorporation of elevated NP level to WATCH-DM(i) improved discrimination (C-index = 0.76, 95% CI 0.74–0.78; DeLong $p < 0.001$], calibration (Brier score: 4.9×10^{-4}), and reclassification (NRI = 0.81, 95% CI 0.74–0.87). Discrimination and calibration were similar for the addition of all biomarkers (elevated NP, hs-cTn, CRP, LVH) or NP alone to WATCH-DM(i).

Incorporation of natriuretic peptide levels to WATCH-DM(ml) model—As reported previously, the WATCH-DM(ml) model demonstrated very high discrimination and calibration in participants with dysglycaemia. The optimal cut-point of WATCH-DM(ml) for predicting HF, as determined by the Youden index, was 0.082.

A significant improvement in categorical and continuous NRI was also observed with the incorporation of NP levels at low/intermediate vs. high WATCH-DM(ml) strata (online supplementary Table S9).

Incorporation of natriuretic peptide levels to WATCH DM model across sex-based subgroups—Addition of elevated NP level status to the WATCH-DM score significantly improved incident HF risk discrimination in both women (C-index = 0.74) and men (C-index = 0.77; DeLong $p < 0.001$ for both). Among both women and men, a significant improvement in risk discrimination was observed by addition of elevated NP levels to the WATCH-DM risk score in the low/intermediate risk participants (WATCH-DM(i) <13) but not high-risk (WATCH-DM(i) ≥13) participants (online supplementary Table S10). Similar results were also observed in reclassification (both categorical and continuous NRI) with elevated NP levels at low/intermediate vs. high WATCH-DM scores among both women and men (online supplementary Table S11).

Incorporation of natriuretic peptide levels to the PCP-HF score—The generalizability of our findings was further evaluated by estimating HF risk using the PCP-HF score. Among 8938 participants with dysglycaemia in the overall cohort, 8221 (92.0%) were free of CVD at baseline, and 374 (4.5%) developed incident HF (10.5 incident HF events per 1000 person-years). The optimal cut-point of PCP-HF for predicting incident HF, as determined by the Youden index, was 0.056. Of the 8221 participants in the primary prevention cohort, 2570 (31.3%) were categorized as high-risk (score ≥0.056). The C-index of the PCP-HF score and for elevated NP levels alone was 0.68 (95% CI 0.66–0.70) and 0.65 (95% CI 0.63–0.68), respectively. Addition of elevated NP levels to the PCP-HF score significantly improved incident HF risk discrimination (C-index = 0.77; DeLong $p < 0.001$). Similar to the WATCH-DM risk score, we observed less increase in C-index with the addition of elevated NP levels across increasing PCP-HF risk estimates (online

supplementary Figure S5A). Specifically, compared with the PCP-HF score alone, addition of elevated NP levels among participants with low/intermediate (PCP-HF score <0.056) but not those with high risk (PCP-HF score ≥0.056) resulted in significant improvement in discrimination (low/intermediate risk: C-index = 0.74, DeLong $p < 0.001$; high risk: C-index = 0.70; DeLong $p = 0.07$) (online supplementary Table S12). Similar results were observed in reclassification (both categorical and continuous NRI) with elevated NP levels at low/intermediate vs. high PCP-HF scores (online supplementary Table S13). Finally, in decision curve analysis, no benefit was observed with the addition of elevated NP levels to the PCP-HF score above a 5-year HF risk threshold of 11% (online supplementary Figure S5B).

Discussion

In this study, we report several key findings. First, the WATCH-DM risk score was validated in a pooled cohort of participants with diabetes from four community-based studies. We observed good discrimination and calibration for predicting 5-year risk of incident HF with the highest performance observed in the machine learning-based [WATCH-DM(ml)] method. Second, in addition to predicting HF risk in individuals with diabetes, the WATCH-DM risk scores performed well in a cohort of participants with pre-diabetes. Third, addition of NP levels to the WATCH-DM risk score resulted in significant improvement in HF risk discrimination with the highest benefit observed at the lower WATCH-DM risk scores (low/intermediate HF risk). Reclassification of participants by the presence or absence of elevated NP levels was significantly higher in participants with low/intermediate vs. high WATCH-DM risk scores. Finally, we were able to generalize our findings by showing improvement in HF risk prediction at low/intermediate HF risk with addition of NP levels by using an alternative clinical HF risk score (PCP-HF). Taken together, the WATCH-DM risk score can accurately stratify HF risk in community-dwelling adults with diabetes and pre-diabetes as well as inform selective use of cardiac biomarker testing for HF risk stratification by up-classifying otherwise low/intermediate risk patients. The WATCH-DM risk score to predict 5-year incident HF risk is publicly available at www.cvriskscores.com.

Several novel aspects of our study are noteworthy. First, it validates the performance of the WATCH-DM risk score in a community-based cohort of participants with diabetes. The performance of the WATCH-DM(ml) score was superior to other established HF risk scores. Second, the WATCH-DM models performed well even with missingness in covariates which may not be routinely available (such as ECG data), highlighting its generalizability in real-world clinical settings. Unlike traditional Cox models, which require all variables to be available for risk prediction, machine learning-based models such as the WATCH-DM(ml) score can better handle missing data and are more flexible in its application across cohorts. Third, our study also establishes WATCH-DM as a clinical risk screening tool among individuals with pre-diabetes. This is particularly relevant as individuals with pre-diabetes despite having abnormalities in cardiometabolic physiology and elevated HF risk, are not targeted for and may benefit from HF prevention therapies. Finally, our data demonstrate the value of biomarkers to identify residual risk of HF in low/intermediate risk patients with diabetes and pre-diabetes. As such, these findings could translate to selective increased

NP-based screening of HF among individuals with dysglycaemia who are deemed to have a low/intermediate risk based on the clinical risk factor profile.

Various studies have demonstrated the utility of cardiac biomarkers — especially NP levels — in informing HF risk prediction among community-dwelling adults. In pooled analyses, multiple studies have demonstrated the association between elevated NT-proBNP and an increased risk of incident HF independent of other risk factors.[31,45] Similar findings have also been demonstrated among individuals with pre-diabetes and diabetes.[46,47] In a recent pooled analysis, a biomarker-based risk score incorporating four biomarkers demonstrated good risk prediction performance comparable to the best fit clinical risk model.[4] In the present study, biomarkers were added to clinical information used to calculate the WATCH-DM risk score and this combined approach led to superior risk prediction compared with a model only incorporating risk factors. Recent studies also evaluated the improvement in risk prediction performance with incorporation of NT-proBNP to clinical risk models. In the ARIC cohort, the addition of NT-proBNP improved the discriminatory performance of established clinical risk prediction models.[48] De Boer *et al.*[31] demonstrated modest improvement in risk prediction with addition of NP levels to a clinical model (delta C-index = 0.005 to 0.02) in a pooled cohort that included community-dwelling participants from America and Europe. Inclusion of NT-proBNP into the clinical risk models has also been shown to improve NRI for predicting HF risk.[44,49] The present study builds on the existing literature by evaluating the optimal approach to combining clinical risk assessment with biomarker testing in individuals with dysglycaemia.

The prognostic utility of biomarkers in risk prediction has also been evaluated across baseline cardiovascular risk strata. In the PREVEND study cohort, Brouwers *et al.*[46] demonstrated that addition of NP levels to clinical parameters improved discrimination of HF risk comparably in high- and low-risk participants as determined by the Framingham risk score. The PREVEND analysis focused on individuals with baseline kidney disease and analysed participants based on history of CVD for baseline risk assessment. In contrast, we observed greater improvement in risk discrimination and calibration with incorporation of NT-proBNP testing among individuals with dysglycaemia who were at low/intermediate risk vs. high risk of HF. The differences in observations by Brouwers *et al.* and our study may be related to differences in the study population (all comers vs. diabetes/pre-diabetes), use of a HF-specific risk score in the present study (vs. only presence of CVD history). Furthermore, our study findings are consistent with prior observations from the coronary artery calcium scoring and atherosclerotic CVD literature.[50]

Clinical trials of NP level testing, such as STOP-HF (St. Vincent's Screening TO Prevent Heart Failure) and PONTIAC (NT-proBNP Selected PreventiOn of cardiac eveNts in a populaTion of diabetic patients without A history of Cardiac disease), have demonstrated significant gains for preventing HF.[13,14] Based on these data, the HF guidelines recommend NP screening in patients at increased risk for developing HF but there is limited guidance regarding the optimal approach to incorporating NP screening with clinical risk assessment.[51] While all patients with diabetes or pre-diabetes may be considered high-risk for developing HF (stage A), 32% of those had a WATCH-DM risk score ≥13 with an 8.4% 5-year risk of developing HF irrespective of NP status. In this subset of high-risk

individuals based on clinical risk factors alone, NP testing did not notably improve HF risk discrimination beyond the clinical HF risk score, and thus may not improve risk prediction. However, more than two-thirds of adults with dysglycaemia had a low WATCH-DM risk score and use of NP level for risk stratification improved HF risk assessment in this subgroup. Nearly one-third of adults with a WATCH-DM risk score <13 had elevated NP levels with a 5-year HF risk of approximately 8%.

The two-step approach to HF risk assessment using a clinical risk model followed by selective biomarker testing in low/intermediate risk individuals has several advantages (Figure 4). Since biomarker testing is associated with notable costs,[52] our study findings provide a pragmatic and cost-efficient approach to risk stratification. Pharmacotherapies such as sodium–glucose co-transporter 2 inhibitors and glucagon-like peptide-1 receptor agonists, intensive lifestyle interventions with diet and exercise, intensive BP control and aggressive weight loss therapies such as bariatric surgery are some of the well-established effective therapies for prevention of HF in patients with diabetes.[7,53,54] However, these therapies are substantially underused in patients with diabetes owing to challenges of high cost and their resource intensive nature.[55] A pragmatic risk-based approach of targeting effective HF prevention strategies to patients at high risk for developing HF may improve appropriate use as these individuals are most likely to derive the greatest absolute benefits from these effective preventive therapies. The two-step approach to risk stratification combining the WATCH-DM risk score with selective biomarker testing can help match high-risk patients with dysglycaemia to more intensive risk-lowering therapies. It is noteworthy that while the benefits of biomarker testing for HF risk prediction are greater among individuals considered low risk for developing HF, there may also be benefits among those at high risk and should be considered using shared decision-making. An electronic medical record-based, clinical decision support-driven randomized controlled trial combining clinical risk assessment (WATCH-DM risk score) with selective biomarker testing is currently underway to evaluate if such an approach can guide effective allocation of preventive therapies for HF among individuals with diabetes (NCT04791826).[56]

Our study is not without limitations. First, this analysis used data from community-based cohort studies that primarily began enrolment prior to 1990. Such studies may not represent a contemporary cohort with access to current glucose-lowering therapies. Second, our study was limited to only individuals with available biomarker testing and data on hs-cTn were not available in participants from CHS. However, analyses in over 6500 participants from ARIC, FHS, and MESA did show similar improvement in HF risk prediction in low-risk individuals with the addition of cardiac biomarkers. Third, similar to blood-based cardiac biomarkers, imaging-based measures of cardiac structure and function may also inform HF risk prediction. However, cardiac imaging data were not consistently available using the same modality across the three cohorts. Thus, we could not assess the complementary role of cardiac imaging in addition to clinical risk score for HF risk prediction. Finally, HF event adjudication differed slightly across each of the four cohorts and HF subtype data were only available in two of the cohorts. However, each of the cohorts used pre-specified and well-established criteria for HF event ascertainment and have been pooled together for prior analyses.[31–33,45,57] Performance of WATCH-DM was better for predicting risk of HFrEF compared with HFpEF. This is consistent with our observation from the initial

external validation analysis in the ALLHAT diabetes cohort.[12] The underperformance of WATCH-DM risk score in predicting risk of HFpEF may be related to greater heterogeneity and misclassification in diagnosis of HFpEF compared with HFrEF as has been reported in several clinical trials.[58]

In conclusion, the WATCH-DM risk score was able to effectively risk stratify and predict 5-year risk of incident HF among adults with diabetes and pre-diabetes in community-dwelling individuals. Furthermore, clinical risk prediction models such as the WATCH-DM score in combination with selective cardiac biomarker testing among individuals at the lower end of the clinical risk score spectrum may be an effective and efficient HF risk stratification strategy among individuals with dysglycaemia.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

This study was supported by the Texas Health Resources Clinical Scholarship to Dr. Pandey.

References

1. Cosentino F, Grant PJ, Aboyans V, Bailey CJ, Ceriello A, Delgado V, et al. 2019 ESC guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD. Eur Heart J. 2020;41(2):255–323. [PubMed: 31497854]
2. Maack C, Lehrke M, Backs J, Heinzel FR, Hulot JS, Marx N, et al. Heart failure and diabetes: metabolic alterations and therapeutic interventions: a state-of-the-art review from the Translational Research Committee of the Heart Failure Association-European Society of Cardiology. Eur Heart J. 2018;39(48):4243–54. [PubMed: 30295797]
3. Seferovic PM, Petrie MC, Filippatos GS, Anker SD, Rosano G, Bauersachs J, et al. Type 2 diabetes mellitus and heart failure: a position statement from the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail. 2018;20(5):853–72. [PubMed: 29520964]
4. Pandey A, Vaduganathan M, Patel KV, Ayers C, Ballantyne CM, Kosiborod MN, et al. Biomarker-based risk prediction of incident heart failure in pre-diabetes and diabetes. JACC Heart Fail. 2021;9(3):215–23. [PubMed: 33422434]
5. Tabak AG, Herder C, Rathmann W, Brunner EJ, Kivimaki M. Prediabetes: a high-risk state for diabetes development. Lancet. 2012;379(9833):2279–90. [PubMed: 22683128]
6. Zelniker TA, Wiviott SD, Raz I, Im K, Goodrich EL, Bonaca MP, et al. SGLT2 inhibitors for primary and secondary prevention of cardiovascular and renal outcomes in type 2 diabetes: a systematic review and meta-analysis of cardiovascular outcome trials. Lancet. 2019;393(10166):31–9. [PubMed: 30424892]
7. Pandey A, Patel KV, Bahnson JL, Gaussoin SA, Martin CK, Balasubramanyam A, et al. ; Look AHEAD Research Group. Association of intensive lifestyle intervention, fitness, and body mass index with risk of heart failure in overweight or obese adults with type 2 diabetes mellitus: an analysis from the Look AHEAD trial. Circulation. 2020;141(16):1295–306. [PubMed: 32134326]
8. Khan SS, Ning H, Shah SJ, Yancy CW, Carnethon M, Berry JD, et al. 10-year risk equations for incident heart failure in the general population. J Am Coll Cardiol. 2019;73(19):2388–97. [PubMed: 31097157]
9. Butler J, Kalogeropoulos A, Georgiopoulou V, Belue R, Rodondi N, Garcia M, et al. ; Health ABC Study. Incident heart failure prediction in the elderly: the Health ABC heart failure score. Circ Heart Fail. 2008;1(2):125–33. [PubMed: 19777072]

10. Agarwal SK, Chambless LE, Ballantyne CM, Astor B, Bertoni AG, Chang PP, et al. Prediction of incident heart failure in general practice: the Atherosclerosis Risk in Communities (ARIC) study. Circ Heart Fail. 2012;5(4):422–9. [PubMed: 22589298]
11. Segar MW, Jaeger BC, Patel KV, Nambi V, Ndumele CE, Correa A, et al. Development and validation of machine learning-based race-specific models to predict 10-year risk of heart failure: a multicohort analysis. Circulation. 2021;143(24):2370–83. [PubMed: 33845593]
12. Segar MW, Vaduganathan M, Patel KV, McGuire DK, Butler J, Fonarow GC, et al. Machine learning to predict the risk of incident heart failure hospitalization among patients with diabetes: the WATCH-DM risk score. Diabetes Care. 2019;42(12):2298–306. [PubMed: 31519694]
13. Ledwidge M, Gallagher J, Conlon C, Tallon E, O'Connell E, Dawkins I, et al. Natriuretic peptide-based screening and collaborative care for heart failure: the STOP-HF randomized trial. JAMA. 2013;310(1):66–74. [PubMed: 23821090]
14. Huelsmann M, Neuhold S, Resl M, Strunk G, Brath H, Francesconi C, et al. PONTIAC (NT-proBNP selected prevention of cardiac events in a population of diabetic patients without a history of cardiac disease): a prospective randomized controlled trial. J Am Coll Cardiol. 2013;62(15):1365–72. [PubMed: 23810874]
15. Investigators ARIC. The Atherosclerosis Risk in Communities (ARIC) study: design and objectives. Am J Epidemiol. 1989;129(4):687–702. [PubMed: 2646917]
16. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, et al. The Cardiovascular Health Study: design and rationale. Ann Epidemiol. 1991;1(3):263–76. [PubMed: 1669507]
17. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham offspring study. Design and preliminary data. Prev Med. 1975;4(4):518–25. [PubMed: 1208363]
18. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. The Framingham Offspring Study. Am J Epidemiol. 1979;110(3):281–90. [PubMed: 474565]
19. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, et al. Multi-Ethnic Study of Atherosclerosis: objectives and design. Am J Epidemiol. 2002;156(9):871–81. [PubMed: 12397006]
20. Stekhoven DJ, Buhlmann P. MissForest–non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012;28(1):112–8. [PubMed: 22039212]
21. Madamanchi C, Alhosaini H, Sumida A, Runge MS. Obesity and natriuretic peptides, BNP and NT-proBNP: mechanisms and diagnostic implications for heart failure. Int J Cardiol. 2014;176(3):611–7. [PubMed: 25156856]
22. Pandey A, Keshvani N, Ayers C, Correa A, Drazner MH, Lewis A, et al. Association of cardiac injury and malignant left ventricular hypertrophy with risk of heart failure in African Americans: the Jackson Heart Study. JAMA Cardiol. 2019;4(1):51–8. [PubMed: 30566191]
23. Patel KV, Simek S, Ayers C, Neeland IJ, de Filippi C, Seliger SL, et al. Physical activity, subclinical myocardial injury, and risk of heart failure subtypes in black adults. JACC Heart Fail. 2021;9(7):484–93. [PubMed: 34119468]
24. Sokolow M, Lyon TP. The ventricular complex in left ventricular hypertrophy as obtained by unipolar precordial and limb leads. Am Heart J. 1949;37(2):161–86. [PubMed: 18107386]
25. Loehr LR, Rosamond WD, Chang PP, Folsom AR, Chambless LE. Heart failure incidence and survival (from the Atherosclerosis Risk in Communities study). Am J Cardiol. 2008;101(7):1016–22. [PubMed: 18359324]
26. Rosamond WD, Chang PP, Baggett C, Johnson A, Bertoni AG, Shahar E, et al. Classification of heart failure in the Atherosclerosis Risk in Communities (ARIC) study: a comparison of diagnostic criteria. Circ Heart Fail. 2012;5(2):152–9. [PubMed: 22271752]
27. Gottdiener JS, Arnold AM, Aurigemma GP, Polak JF, Tracy RP, Kitzman DW, et al. Predictors of congestive heart failure in the elderly: the cardiovascular health study. J Am Coll Cardiol. 2000;35(6):1628–37. [PubMed: 10807470]
28. Ives DG, Fitzpatrick AL, Bild DE, Psaty BM, Kuller LH, Crowley PM, et al. Surveillance and ascertainment of cardiovascular events: the Cardiovascular Health Study. Ann Epidemiol. 1995;5(4):278–85. [PubMed: 8520709]

29. McKee PA, Castelli WP, McNamara PM, Kannel WB. The natural history of congestive heart failure: the Framingham study. N Engl J Med. 1971;285(26):1441–6. [PubMed: 5122894]
30. de Lemos JA, Ayers CR, Levine BD, de Filippi CR, Wang TJ, Hundley WG, et al. Multimodality strategy for cardiovascular risk assessment: performance in 2 population-based cohorts. Circulation. 2017;135(22):2119–32. [PubMed: 28360032]
31. de Boer RA, Nayor M, de Filippi CR, Enserro D, Bhambhani V, Kizer JR, et al. Association of cardiovascular biomarkers with incident heart failure with preserved and reduced ejection fraction. JAMA Cardiol. 2018;3(3):215–24. [PubMed: 29322198]
32. Ho JE, Enserro D, Brouwers FP, Kizer JR, Shah SJ, Psaty BM, et al. Predicting heart failure with preserved and reduced ejection fraction: the international collaboration on heart failure subtypes. Circ Heart Fail. 2016;9(6):e003116. [PubMed: 27266854]
33. Pandey A, LaMonte M, Klein L, Ayers C, Psaty BM, Eaton CB, et al. Relationship between physical activity, body mass index, and risk of heart failure. J Am Coll Cardiol. 2017;69(9):1129–42. [PubMed: 28254175]
34. Action to Control Cardiovascular Risk in Diabetes Study Group. Effects of intensive glucose lowering in type 2 diabetes. N Engl J Med. 2008;358(24):2545–59. [PubMed: 18539917]
35. ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial. Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). JAMA. 2002;288(23):2981–97. [PubMed: 12479763]
36. Look AHEAD Research Group. Cardiovascular effects of intensive lifestyle intervention in type 2 diabetes. N Engl J Med. 2013;369(2):145–54. [PubMed: 23796131]
37. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA. 1982;247(18):2543–6. [PubMed: 7069920]
38. D'Agostino RB, Nam BH. Evaluation of the performance of survival analysis models: discrimination and calibration measures. Handbook of statistics. Volume 23. Elsevier; 2003. p. 1–25.
39. Rufibach K Use of Brier score to assess binary predictions. J Clin Epidemiol. 2010;63(8):938–9; author reply 939. [PubMed: 20189763]
40. Youden WJ. Index for rating diagnostic tests. Cancer. 1950;3(1):32–5. [PubMed: 15405679]
41. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26(6):565–74. [PubMed: 17099194]
42. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837–45. [PubMed: 3203132]
43. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med. 2011;30(1):11–21. [PubMed: 21204120]
44. Chahal H, Bluemke DA, Wu CO, McClelland R, Liu K, Shea SJ, et al. Heart failure risk prediction in the Multi-Ethnic Study of Atherosclerosis. Heart. 2015;101(1):58–64. [PubMed: 25381326]
45. Suthahar N, Lau ES, Blaha MJ, Paniagua SM, Larson MG, Psaty BM, et al. Sex-specific associations of cardiovascular risk factors and biomarkers with incident heart failure. J Am Coll Cardiol. 2020;76(12):1455–65. [PubMed: 32943164]
46. Brouwers FP, van Gilst WH, Damman K, van den Berg MP, Gansevoort RT, Bakker SJL, et al. Clinical risk stratification optimizes value of biomarkers to predict new-onset heart failure in a community-based cohort. Circ Heart Fail. 2014;7(5):723–31. [PubMed: 25057112]
47. Nguyen K, Fan W, Bertoni A, Budoff MJ, Defilippi C, Lombardo D, et al. N-terminal pro B-type natriuretic peptide and high-sensitivity cardiac troponin as markers for heart failure and cardiovascular disease risks according to glucose status (from the Multi-Ethnic Study of Atherosclerosis [MESA]). Am J Cardiol. 2020;125(8):1194–201. [PubMed: 32106929]
48. Nambi V, Liu X, Chambless LE, de Lemos JA, Virani SS, Agarwal S, et al. Troponin T and N-terminal pro-B-type natriuretic peptide: a biomarker approach to predict heart failure

risk-the Atherosclerosis Risk in Communities study. Clin Chem. 2013;59(12):1802–10. [PubMed: 24036936]

49. Smith JG, Newton-Cheh C, Almgren P, Struck J, Morgenthaler NG, Bergmann A, et al. Assessment of conventional cardiovascular risk factors and multiple biomarkers for the prediction of incident heart failure and atrial fibrillation. J Am Coll Cardiol. 2010;56(21):1712–9. [PubMed: 21070922]
50. Nasir K, Bittencourt MS, Blaha MJ, Blankstein R, Agatson AS, Rivera JJ, et al. Implications of coronary artery calcium testing among statin candidates according to American College of Cardiology/American Heart Association Cholesterol Management Guidelines: MESA (Multi-Ethnic Study of Atherosclerosis). J Am Coll Cardiol. 2015;66(15):1657–68. [PubMed: 26449135]
51. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr, Colvin MM, et al. 2017 ACC/AHA/HFSA focused update of the 2013 ACCF/AHA guideline for the Management of Heart Failure: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. Circulation. 2017;136(6):e137–61. [PubMed: 28455343]
52. Ledwidge MT, O'Connell E, Gallagher J, Tilson L, James S, Voon V, et al. Cost-effectiveness of natriuretic peptide-based screening and collaborative care: a report from the STOP-HF (St Vincent's Screening TO Prevent Heart Failure) study. Eur J Heart Fail. 2015;17(7):672–9. [PubMed: 26139583]
53. McGuire DK, Shih WJ, Cosentino F, Charbonnel B, DZI C, Dagogo-Jack S, et al. Association of SGLT2 inhibitors with cardiovascular and kidney outcomes in patients with type 2 diabetes: a meta-analysis. JAMA Cardiol. 2021;6(2):148–58. [PubMed: 33031522]
54. Sattar N, Lee MMY, Kristensen SL, Branch KRH, del Prato S, Khurmi NS, et al. Cardiovascular, mortality, and kidney outcomes with GLP-1 receptor agonists in patients with type 2 diabetes: a systematic review and meta-analysis of randomised trials. Lancet Diabetes Endocrinol. 2021;9(10):653–62. [PubMed: 34425083]
55. Newman JD, Berger JS, Ladapo JA. Underuse of medications and lifestyle counseling to prevent cardiovascular disease in patients with diabetes. Diabetes Care. 2019;42(5):e75–6. [PubMed: 30862654]
56. ClinicalTrials.gov. Electronic Alerts for Heart Failure Prevention in Diabetes. https://ClinicalTrials.gov/show/NCT04791826. Accessed 8 November 2021.
57. Segar MW, Khan MS, Patel KV, et al. Prevalence and prognostic implications of diabetes with cardiomyopathy in community-dwelling adults. J Am Coll Cardiol. 2021;78(16):1587–98. [PubMed: 34649696]
58. Segar MW, Patel KV, Ayers C, Basit M, Tang WHW, Willett D, et al. Phenomapping of patients with heart failure with preserved ejection fraction using machine learning-based unsupervised cluster analysis. Eur J Heart Fail. 2020;22(1):148–58. [PubMed: 31637815]

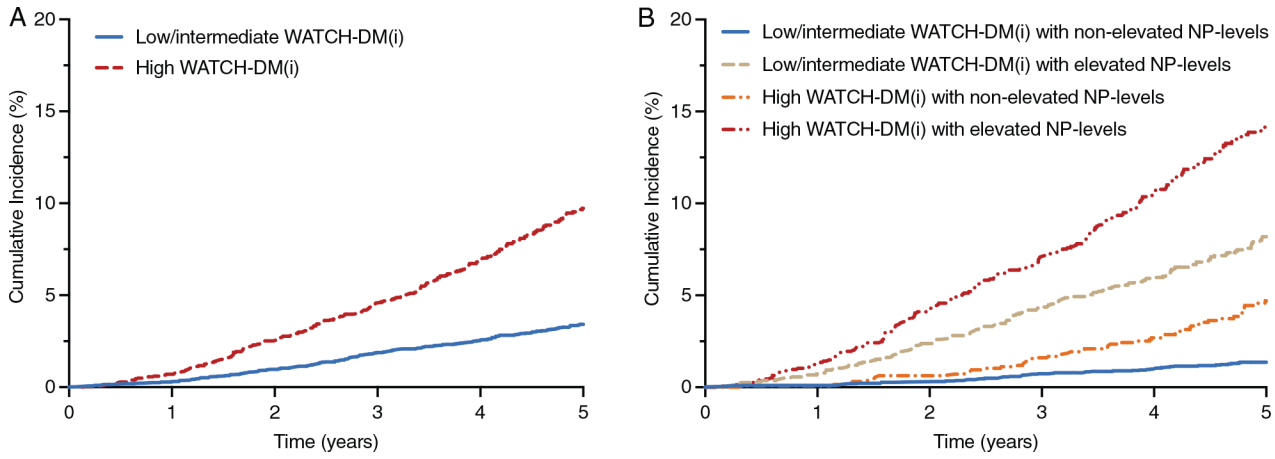

Figure 1.

Cumulative incidence of heart failures events (*A*) across high (scores≥13) and low/intermediate (<13) WATCH-DM(i) scores and (*B*) additionally reclassifying participants by presence or absence of elevated natriuretic peptide (NP) levels among individuals with dysglycaemia. Elevated NP levels were able to identify individuals with high risk of heart failure in the low/intermediate WATCH-DM risk score strata.

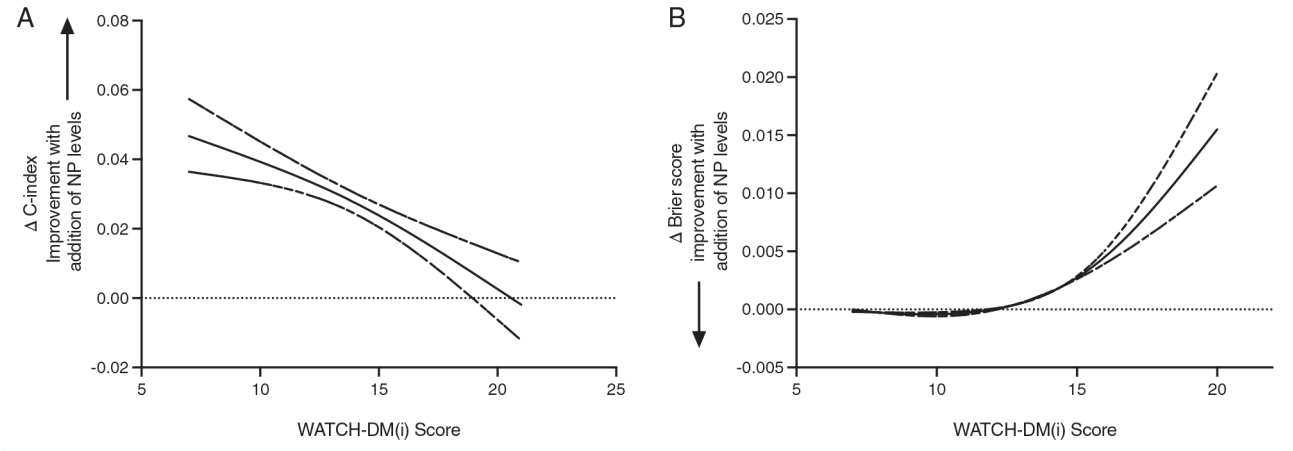

Figure 2.

Mean improvement in (*A*) discrimination (C-index) and (*B*) calibration (Brier score) of the WATCH-DM(i) score with addition of natriuretic peptide (NP) levels across WATCH-DM(i) scores in the overall cohort. The dashed lines represent the 95% confidence intervals obtained from bootstrapping with 1000 replicates. The improvement in C-index and Brier score with incorporation of NP levels was observed only at low/intermediate risk levels of the WATCH-DM risk score but not at high levels.

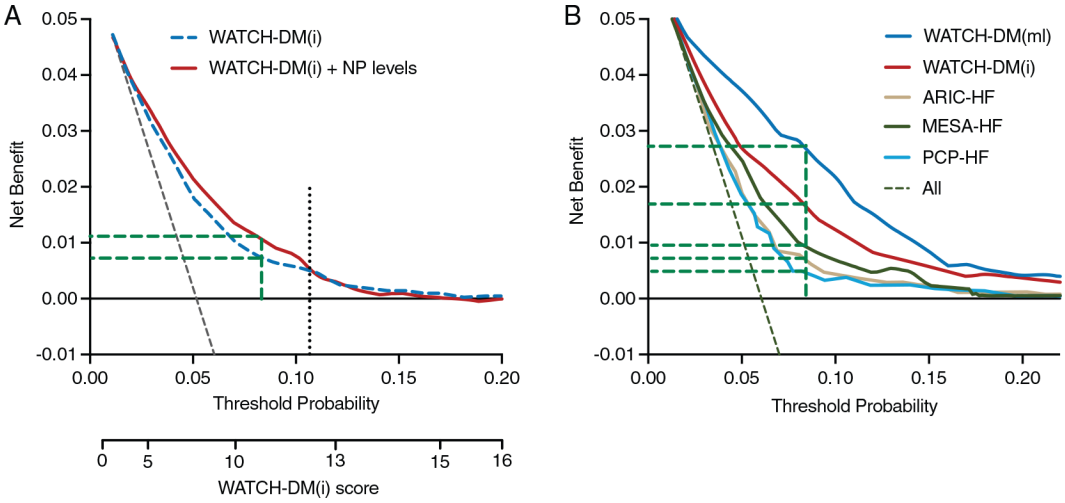

Figure 3.

Decision curve analysis for comparison of (*A*) WATCH-DM(i) score with and without elevated natriuretic peptide (NP) levels in the overall cohort and (*B*) different heart failure risk prediction models. The grey dashed line indicates the net benefit of all patients. The WATCH-DM(i) score detected 8 true positive events per 1000 patients compared with 12 per 1000 patients for the WATCH-DM(i) + elevated NP level model. Thus, addition of NP levels to the WATCH-DM(i) score increased the number of heart failure events detected by 4 per 1000 patients compared with the WATCH-DM(i) score alone. Above a threshold probability of 10.9%, addition of NP levels to the WATCH-DM(i) score confers no added net clinical benefit compared with the WATCH-DM risk score alone. (*B*) At an event rate of 8%, the WATCH-DM(ml), WATCH-DM(i), ARIC-HF, MESA-HF, and PCP-HF risk scores detected 28, 18, 8, 10, and 5 true positive events per 1000 patients, respectively (horizontal dashed green lines).

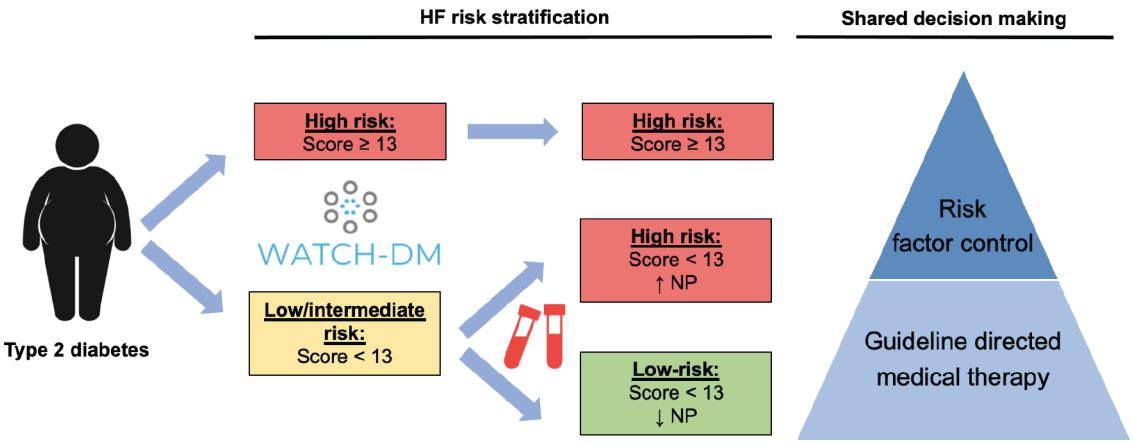

Figure 4.

A pragmatic approach to heart failure risk stratification in diabetes. The study findings highlight the complementary role of natriuretic peptide assessment in patients with low/intermediate risk of HF as determined by a clinical risk score such as the WATCH-DM score. High-risk patients as identified by an elevated clinical risk score or by elevated biomarker levels in the setting of low/intermediate risk can be targeted with guideline-recommended effective heart failure prevention therapies.

Table 1

Discrimination (C-index) and calibration (Greenwood-Nam-D'Agostino statistic) of the machine learning- and integer-based WATCH-DM risk score among participants with dysglycaemia, diabetes, and pre-diabetes

Risk score	C-index (95% CI)	Brier score ($\times 10^{-4}$)	GND statistic (*p*-value)
Overall			
WATCH-DM(ml)	0.80 (0.78–0.81)	9.0	1 3.3 (0.1 0)
WATCH-DM(i)	0.71 (0.68–0.73)	5.1	6.2 (0.63)
Diabetes			
WATCH-DM(ml)	0.76 (0.73–0.80)	1 4.0	1 2.7 (0.1 2)
WATCH-DM(i)	0.69 (0.67–0.71)	1 0.4	6.1 (0.64)
Pre-diabetes			
WATCH-DM(ml)	0.83 (0.80–0.85)	5.8	1 2.3 (0.1 4)
WATCH-DM(i)	0.72 (0.69–0.74)	1.8	11.5 (0.18)

CI, confidence interval; GND, Greenwood–Nam–D'Agostino.

A GND $p > 0.05$ indicates adequate calibration. GND is the chi-square statistic (*p*-value).

Table 2

Comparison of risk discrimination (C-index) and calibration (Brier score) between WATCH-DM(i) score alone vs. WATCH-DM(i) + natriuretic peptide level assessment (elevated vs. not elevated) among participants with low/intermediate clinical risk [WATCH-DM(i) <13] vs. WATCH-DM(i) + natriuretic peptide level assessment among participants with high clinical risk [WATCH-DM(i) ≥13]

Risk models	Discrimination		Calibration
	C-index (95% CI)	Comparison *p*-value	Brier score ($\times 10^{-4}$)
Overall			
WATCH-DM(i) only	0.71 (0.68–0.73)	Ref.	5.1
WATCH-DM(i) + NP levels in low/intermediate-risk participants	0.73 (0.71 –0.75)	<0.001	2.2
WATCH-DM(i) + NP levels in high-risk participants	0.71 (0.69–0.73)	0.34	24.7
Diabetes			
WATCH-DM(i) only	0.69 (0.67–0.71)	Ref.	1 0.4
WATCH-DM(i) + NP levels in low/intermediate-risk participants	0.71 (0.68–0.73)	<0.001	6.4
WATCH-DM(i) + NP levels in high-risk participants	0.69 (0.66–0.71)	0.57	1 2.1
Pre-diabetes			
WATCH-DM(i) only	0.72 (0.69–0.74)	Ref.	1.8
WATCH-DM(i) + NP levels in low/intermediate-risk participants	0.77 (0.74–0.80)	<0.001	1.3
WATCH-DM(i) + NP levels in high-risk participants	0.71 (0.67–0.75)	0.1 9	1 3.7

BMI, body mass index; BNP B-type natriuretic peptide; CI, confidence interval; NP, natriuretic peptide; NT-proBNP; N-terminal pro-B-type natriuretic peptide.

Separate analyses were performed in the overall, diabetes, and pre-diabetes cohorts. Elevated NP levels were BMI-specific as follows: NT-proBNP ≥125pg/ml or BNP ≥40pg/ml for BMI ≤30 kg/m^2; NT-proBNP ≥100pg/ml or BNP ≥30pg/ml for BMI ≥30kg/m^2. The DeLong test was used to compare C-indices across risk models.

Table 3

Comparison of reclassification (categorical and continuous net reclassification improvement and integrated discrimination improvement) between WATCH-DM(i) score alone vs. WATCH-DM(i) + natriuretic peptide level assessment (elevated vs. not elevated) among participants with low/intermediate clinical risk [WATCH-DM(i) <13] vs. WATCH-DM(i) + natriuretic peptide level assessment among participants with high clinical risk [WATCH-DM(i) ≥ 13]

Risk models	Categorical NRI		Continuous NRI		IDI	
	Estimate (95% CI)	Comparison *p*-value	Estimate (95% CI)	Comparison *p*-value	Estimate (95% CI)	Comparison *p*-value
Overall						
WATCH-DM(i) + NP-levels in low/intermediate-risk participants	0.36 (0.32–0.40)	<0.001	0.45 (0.33–0.56)	<0.001	0.04 (0.02–0.07)	0.06
WATCH-DM(i) + NPlevels in high-risk participants	0.08 (0.02–0.11)		0.17 (0.09–0.27)		0.02 (0.01 –0.03)	
Diabetes						
WATCH-DM(i) + NPlevels in low/intermediate-risk participants	0.27 (0.24–0.33)	0.02	0.43 (0.28–0.56)	0.09	0.05 (0.01 –0.05)	0.68
WATCH-DM(i) + NPlevels in high-risk participants	0.15 (0.07–0.27)		0.25 (0.10–0.37)		0.02 (0.01 –0.05)	
Pre-diabetes						
WATCH-DM(i) + NPlevels in low/intermediate-risk participants	0.48 (0.37–0.61)	<0.001	0.66 (0.53–0.81)	<0.001	0.08 (0.04–0.1 4)	0.03
WATCH-DM(i) + NP-levels in high-risk participants	0.02 (−0.07–0.09)		0.02 (−0.13–0.15)		0.02 (0.004–0.04)	

BMI, body mass index; BNP, B-type natriuretic peptide; CI, confidence interval; IDI, integrated discrimination improvement; NP, natriuretic peptide; NRI, net reclassification improvement; NT-proBNP; N-terminal pro-B-type natriuretic peptide.

Separate analyses were performed in the overall, diabetes, and pre-diabetes cohorts. Elevated NP levels were BMI-specific as follows: NT-proBNP ≥125pg/ml or BNP ≥40pg/ml for BMI ≤30kg/m^2; NT-proBNP ≥100pg/ml or BNP ≥30pg/ml for BMI >30kg/m^2. Categorical NRI was classified into <2%, 2%–5% 5.1%–15%, and >15% risk. The Welch's t-test was used to compare NRI values of low/intermediate and high WATCH-DM(i) strata.