ESC European Society of Cardiology

European Journal of Heart Failure (2022) **24**, 1869–1878
doi:10.1002/ejhf.2607

RESEARCH ARTICLE

Biomarker-driven prognostic models in chronic heart failure with preserved ejection fraction: the EMPEROR–Preserved trial

Stuart J. Pocock[1]*, João Pedro Ferreira[2,3], Milton Packer[4,5], Faiez Zannad[3], Gerasimos Filippatos[6], Toru Kondo[7,8], John J.V. McMurray[8], Scott D. Solomon[9], James L. Januzzi[10], Tomoko Iwata[11], Afshin Salsali[11], Javed Butler[12,13], and Stefan D. Anker[14]

[1]Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK; [2]UnIC@Rise, Department of Surgery and Physiology, Faculty of Medicine, Cardiovascular Research and Development Center, University of Porto, Porto, Portugal; [3]Inserm, Centre d'Investigations Cliniques Plurithématique 1433, and Inserm U1116, CHRU, F-CRIN INI-CRCT (Cardiovascular and Renal Clinical Trialists), Université de Lorraine, Nancy, France; [4]Baylor Heart and Vascular Hospital, Baylor University Medical Center, Dallas, TX, USA; [5]Imperial College, London, UK; [6]National and Kapodistrian University of Athens, School of Medicine, Department of Cardiology, Attikon University Hospital, Athens, Greece; [7]Department of Cardiology, Nagoya University Graduate School of Medicine, Nagoya, Japan; [8]British Heart Foundation Cardiovascular Research Centre, University of Glasgow, Glasgow, UK; [9]Cardiovascular Division, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; [10]Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; [11]Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach, Germany; [12]Baylor Scott and White Research Institute, Dallas, TX, USA; [13]Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA; and [14]Department of Cardiology, and Berlin Institute of Health Center for Regenerative Therapies, German Centre for Cardiovascular Research Partner Site Berlin, Charité Universitätsmedizin, Berlin, Germany

Received 4 March 2022; revised 1 July 2022; accepted 2 July 2022; online publish-ahead-of-print 24 July 2022

Aims Biomarker-driven prognostic models incorporating N-terminal pro-B-type natriuretic peptide (NT-proBNP) and high-sensitivity cardiac troponin T (hs-cTnT) in heart failure (HF) with preserved ejection fraction (HFpEF) are lacking. We aimed to generate a biomarker-driven prognostic tool for patients with chronic HFpEF enrolled in EMPEROR-Preserved.

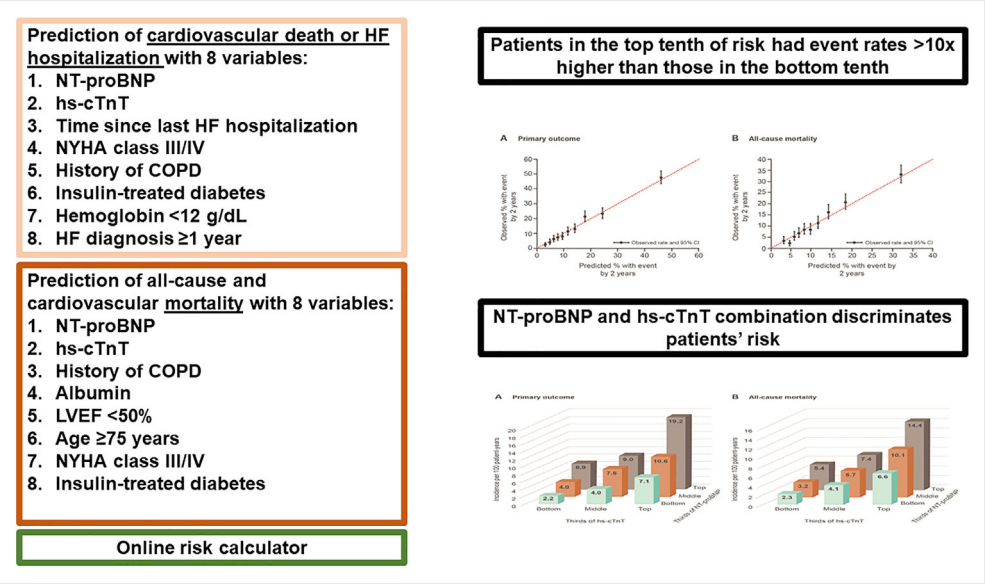
Methods and results Multivariable Cox regression models were created for (i) the primary composite outcome of HF hospitalization or cardiovascular death, (ii) all-cause death, (iii) cardiovascular death, and (iv) HF hospitalization. PARAGON-HF was used as a validation cohort. NT-proBNP and hs-cTnT were the dominant predictors of the primary outcome, and in addition, a shorter time since last hospitalization, New York Heart Association (NYHA) class III or IV, history of chronic obstructive pulmonary disease (COPD), insulin-treated diabetes, low haemoglobin, and a longer time since HF diagnosis were key predictors (eight variables, all $p < 0.001$). The consequent primary outcome risk score discriminated well (*c*-statistic = 0.75) with patients in the top 10th of risk having an event rate >22× higher than those in the bottom 10th. A model for HF hospitalization alone had even better discrimination ($c = 0.79$). Empagliflozin reduced the risk of cardiovascular death or hospitalization for HF in patients across all risk levels. NT-proBNP and hs-cTnT were also the dominant predictors of all-cause and cardiovascular mortality followed by history of COPD, low albumin, older age, left ventricular ejection fraction ≥50%, NYHA class III or IV and insulin-treated diabetes (eight variables, all $p < 0.001$). The mortality risk model had similar discrimination for all-cause and cardiovascular mortality (*c*-statistic = 0.72 for both). External validation provided *c*-statistics of 0.71, 0.71, 0.72, and 0.72 for the primary outcome, HF hospitalization alone, all-cause death, and cardiovascular death, respectively.

Conclusions The combination of NT-proBNP and hs-cTnT along with a few readily available clinical variables provides effective risk discrimination both for morbidity and mortality in patients with HFpEF. A predictive tool-kit facilitates the ready implementation of these risk models in routine clinical practice.

*Corresponding author. London School of Hygiene and Tropical Medicine, Keppel St, London WC1E 7HT, UK. Email: stuart.pocock@lshtm.ac.uk

© 2022 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Graphical Abstract



Biomarker-driven prognostic models for heart failure (HF) with preserved ejection fraction. COPD, chronic obstructive pulmonary disease; hs-cTn, high-sensitivity cardiac troponin T; LVEF, left ventricular ejection fraction; NT-proBNP, N-terminal pro-B-type natriuretic peptide; NYHA, New York Heart Association.

Keywords Risk model • Biomarkers • Heart failure with preserved ejection fraction

Introduction

Patients with chronic heart failure (HF) and preserved ejection fraction (HFpEF) typically have a poor prognosis. An accurate prediction of individual patient prognosis may be important to define tailored care strategies (e.g. frequency of clinic visits, home follow-up by health care professionals, or prognosis-related discussions with the patient and his/her family members). Although risk models have been developed for patients with HF and a reduced ejection fraction (HFrEF), there is a lack of established prognostic models for patients with HFpEF. The CHARM risk score was developed for patients with both HFrEF and HFpEF, but it is now over 15 years old and did not include well-established prognostic biomarkers such as N-terminal pro-B-type natriuretic peptide (NT-proBNP) and high-sensitivity cardiac troponin T (hs-cTnT).[1] The MAGGIC meta-analysis of 30 cohort studies across the whole spectrum of chronic HF, included a sub-analysis in 407 patients with HFpEF.[2] These models did not include natriuretic peptides, cardiac troponins or their combination. The I-Preserve risk models,[3] and the 3A3B score,[4] included NT-proBNP but did not include cardiac troponins.

We have recently developed biomarker-driven prognostic models for patients with HFrEF. In these models, NT-proBNP and hs-cTnT were the strongest prognostic predictors, followed by a few readily available clinical variables, which facilitated the application of this prognostic tool-kit in routine practice.[5] This experience motivated us to use data from Empagliflozin Outcome Trial in Patients with Chronic Heart Failure and a Preserved Ejection Fraction (EMPEROR-Preserved) on NT-proBNP and hs-cTnT to develop simple and ready-to-implement models for predicting the individual patient incidence of the composite of HF hospitalization or cardiovascular death, all-cause death, and cardiovascular death in patients with HFpEF. We go on to validate these models using the PARAGON-HF trial. Available risk scores for HFpEF do not include the combined use of high-sensitivity troponin and NT-proBNP, which are biomarkers with strong prognostic value in HF. We hypothesize that a novel biomarker-driven risk score will outperform the existing risk scores for HFpEF allowing a better stratification of risk in these patients.

Methods

Study design

The design and primary results of the EMPEROR-Preserved trial have been described previously.[6,7] In brief, participants had chronic HF with New York Heart Association (NYHA) functional class II to IV symptoms and a left ventricular ejection fraction (LVEF) >40% with no

© 2022 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

prior measurement ≤40%. Patients were also required to have an elevated NT-proBNP level (>900 pg/ml or >300 pg/ml in patients with and without atrial fibrillation, respectively) and a documented hospitalization for HF or evidence of structural heart disease within 12 months before enrolment. A total of 5988 patients were randomized during March 2017 to April 2020 to receive either empagliflozin 10 mg or placebo daily over a median follow-up of 26.2 months. The primary outcome was time-to-first event in a composite of HF hospitalization or cardiovascular death. The mode of death and hospitalizations for HF were independently adjudicated by a blinded committee based on pre-specified criteria.

Blood was collected for measurement of NT-proBNP (expressed in pg/ml) and hs-cTnT (expressed in ng/L) at baseline and measured in a central laboratory (Roche Diagnostics, Risch-Rotkreutz, Switzerland) using a Roche® Cobas analyzer. The Institutional Review Board of each study site approved all study procedures and all patients provided informed consent.

External validation was performed in a subset of patients from the PARAGON-HF trial. During July 2014 to December 2016, the PARAGON-HF trial randomized 4796 patients with HFpEF to sacubitril/valsartan or valsartan; its methods and primary results have been published previously.[8] A subset of 1251 patients from PARAGON-HF who had available hs-cTnT, NT-proBNP (using the same assays), and the remaining variables required to validate the EMPEROR-Preserved models was used for external validation.[9]

Statistical analysis

For the primary composite outcome of HF hospitalization or cardiovascular death, HF hospitalization alone, all-cause mortality, and cardiovascular death, multivariable Cox proportional hazard models were used to study the relation of patient variables at baseline to outcome incidence. First, 35 variables were selected by the EMPEROR-Preserved Executive Committee based on their availability, clinical significance and potential prognostic importance; these 35 candidate predictor variables are listed in online supplementary *Table S1*. Second, we used stepwise forward variable selection with $p < 0.001$ as a criterion for inclusion (to keep the model parsimonious) with log-transformed NT-proBNP and hs-cTnT to achieve a good linear fit. Third, we evaluated model discrimination (using Harrell's *c*-statistics) and calibration (by plotting the observed vs. predicted 2-year risk by deciles of risk). Missing data were rare, with all candidate covariates available in >90% of patients. We used single value imputation to impute missing data using the median for continuous variables or the mode for categorical variables.

The EMPEROR-Preserved risk models' coefficients were applied to the PARAGON-HF subset of 1251 patients to obtain their predictive capacity expressed by *c*-statistics and model calibration shown in calibration plots.

Analyses were performed using SAS, version 9.4 (SAS Institute Inc., Cary, NC, USA) and STATA, version 17.0 (StataCorp 2021, College Station, TX, USA).

Results

Primary composite outcome

The primary composite outcome of HF hospitalization or cardiovascular death occurred in 415 of 2997 patients (13.8%) in the empagliflozin group and in 511 of 2991 patients (17.1%) in the placebo group (hazard ratio [HR] 0.79, 95% confidence interval [CI] 0.69–0.90, $p < 0.001$).

After stepwise variable selection, the prognostic model for the primary outcome included log-transformed NT-proBNP and hs-cTnT as the most powerful predictors, based on the chi-squared statistic for inclusion. These two biomarkers were followed by (1) shorter time since most recent HF hospitalization, (2) NYHA functional class III/IV, (3) history of chronic obstructive pulmonary disease (COPD), (4) insulin-treated diabetes, (5) haemoglobin <12 g/dl, and (6) time since HF diagnosis of 1 year or greater. The randomized treatment (empagliflozin vs. placebo) remained highly predictive after adjustment for these eight baseline predictors (*Table 1*).

The strength of prediction for the primary outcome is captured by the *c*-statistic of 0.748 (95% CI 0.732–0.764). The goodness of fit and strength of prediction for this model, based on eight baseline predictors plus randomized treatment, are displayed in *Figure 1A*.

A risk score based on the coefficients in *Table 1* has its distribution divided into 10 equal-sized groups. In each decile there is good agreement between the observed and model-predicted patient risk, both expressed as the percentage having a primary event in 2 years. Comparing top and bottom deciles of risk, the observed 2-year event rates are 47.8% and 2.1%, respectively. We considered the option of extending the model to include other predictors that each achieved $p < 0.01$ rather than the more stringent $p < 0.001$: this would have added total bilirubin and history of left bundle branch block. Also, both albuminuria determined by urinary albumin-to-creatinine ratio (UACR), and Kansas City Cardiomyopathy Questionnaire overall summary score (KCCQ-OSS) were highly significant independent predictors (each $p < 0.0001$), but they were excluded from our primary model because these variables are not readily available for clinical use, and also UACR was not available in our validation cohort. The addition of these two variables would increase model complexity but with only a slight gain in strength of prediction (*c* becomes 0.756). The consequent model is shown in online supplementary *Table S3A*.

To evaluate the effect of empagliflozin at different levels of patient risk, we calculated a risk score using coefficients for the eight significant predictors in *Table 1* (excluding the coefficient for randomized treatment). We then stratified patients into equal-sized thirds of risk each containing around 2000 patients per risk tertile. *Figure 2* (upper panel) shows the HR and 95% CIs for empagliflozin versus placebo by tertiles of risk. There is a consistency of relative risk reduction in all three risk groups (p for trend = 0.68). *Figure 2* (lower panel) shows the absolute difference in risk for the same tertiles, expressed as the treatment difference in primary event rate per 100 patient-years. On this absolute scale there is a significant trend in treatment effect by risk groups: for low, medium, and high-risk groups the rate differences are −0.64, −1.86 and −3.93 primary events per 100 patient-years, respectively (p for trend = 0.026). Kaplan–Meier plots of the primary outcome over 24 months by risk tertiles and by treatment groups confirm these patterns of treatment effect by patient risk (online supplementary *Figure S1*).

© 2022 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

Table 1 **The EMPEROR-Preserved risk model for the primary outcome (heart failure hospitalization or cardiovascular death)**

Variables	Hazard ratio (95% CI)	Chi-squared statistic	Coefficient (SE)*	*p*-value
Log NT-proBNP (pg/ml)	1.61 (1.49–1.74)	144.7	0.48 (0.04)	<0.0001
Log hs-cTnT (ng/L)	1.59 (1.44–1.75)	90.5	0.46 (0.05)	<0.0001
Time since most recent HHF				
>6 months	1.00 (reference)			<0.0001
3–6 months	1.71 (1.36–2.15)	21.4	0.54 (0.12)	
<3 months	1.95 (1.64–2.31)	57.8	0.67 (0.09)	
NYHA class III/IV	1.67 (1.45–1.92)	49.8	0.51 (0.07)	<0.0001
History of COPD	1.70 (1.45–1.99)	43.3	0.53 (0.08)	<0.0001
Use of insulin in DM patients				
Non-DM	1.00 (reference)			<0.0001
DM + no insulin	1.14 (0.98–1.32)	2.8	0.13 (0.08)	
DM + insulin	1.61 (1.35–1.92)	28.9	0.48 (0.09)	
Baseline haemoglobin				
≥12 g/dl	1.00 (reference)			<0.0001
<12 g/dl	1.42 (1.23–1.64)	22.0	0.35 (0.07)	
Time since HF diagnosis				
3 months–1 year	1.00 (reference)			0.0002
≥1 year	1.41 (1.20–1.66)	17.1	0.34 (0.08)	
Randomized to empagliflozin	0.75 (0.66–0.86)	18.6	−0.29 (0.07)	<0.0001

CI, confidence interval; COPD, chronic obstructive pulmonary disease; DM, diabetes mellitus; HF, heart failure; HHF, hospitalization for heart failure; hs-cTnT, high-sensitivity cardiac troponin T; NT-proBNP, N-terminal pro-B-type natriuretic peptide; NYHA, New York Heart Association; SE, standard error.

[a]Coefficient (SE) are the log hazard ratio and its standard error.

An estimate of each individual's 2-year risk can be calculated as follows: $1 - [0.99932^{exp}$ $(0.46 \times \log$ hs-cTnT $+ 0.48 \times \log$ NT-proBNP $+ 0.51 \times$ NYHA class $+ 0.54 \times$ recent HHF1 $+ 0.67 \times$ recent HHF2 $+ 0.53 \times$ COPD $+ 0.48 \times$ DM (insulin) $+ 0.13 \times$ DM (without insulin) $+ 0.34 \times$ HF diagnosis $+ 0.35 \times$ haemoglobin $- 0.29 \times$ empagliflozin)], where 'recent HHF1' and 'recent HHF2' are indicator variables for whether the most recent HHF was within 3–6 or <3 months, respectively. NYHA class is an indicator variable for whether the patient's NYHA class is III or IV. COPD and empagliflozin are indicator variables for whether the patient has COPD or is to be treated with empagliflozin, respectively. DM (insulin) and DM (without insulin) are indicator variables for whether the patient has diabetes and use of insulin at baseline or has diabetes without use of insulin. HF diagnosis is an indicator variable for whether time since HF diagnosis is ≥1 year. Haemoglobin is an indicator variable for whether baseline haemoglobin is <12 g/dl. *C*-statistic (95% CI): 0.748 (0.732–0.764).

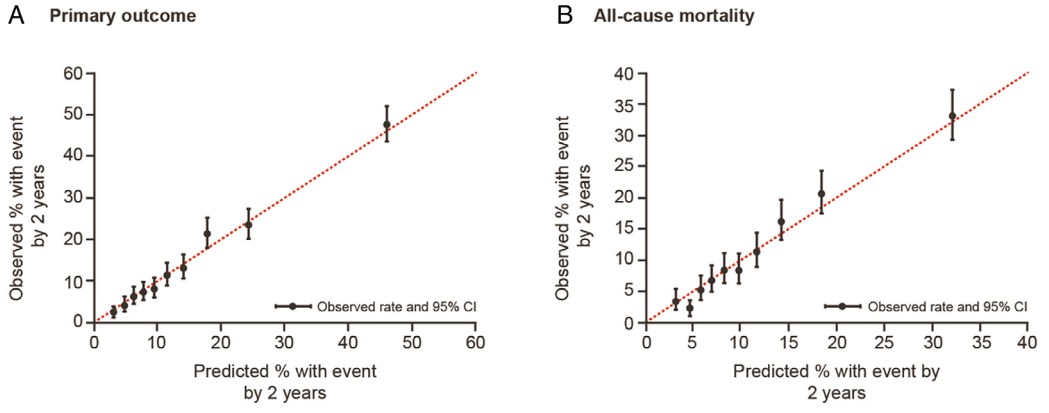


Figure 1 Observed versus predicted events by tenths of the risk score distribution. Hosmer–Lemeshow goodness-of-fit test $p = 0.198$ for the primary outcome (*A*) and $p = 0.039$ for all-cause mortality (*B*), indicating adequate calibration. CI, confidence interval.

© 2022 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

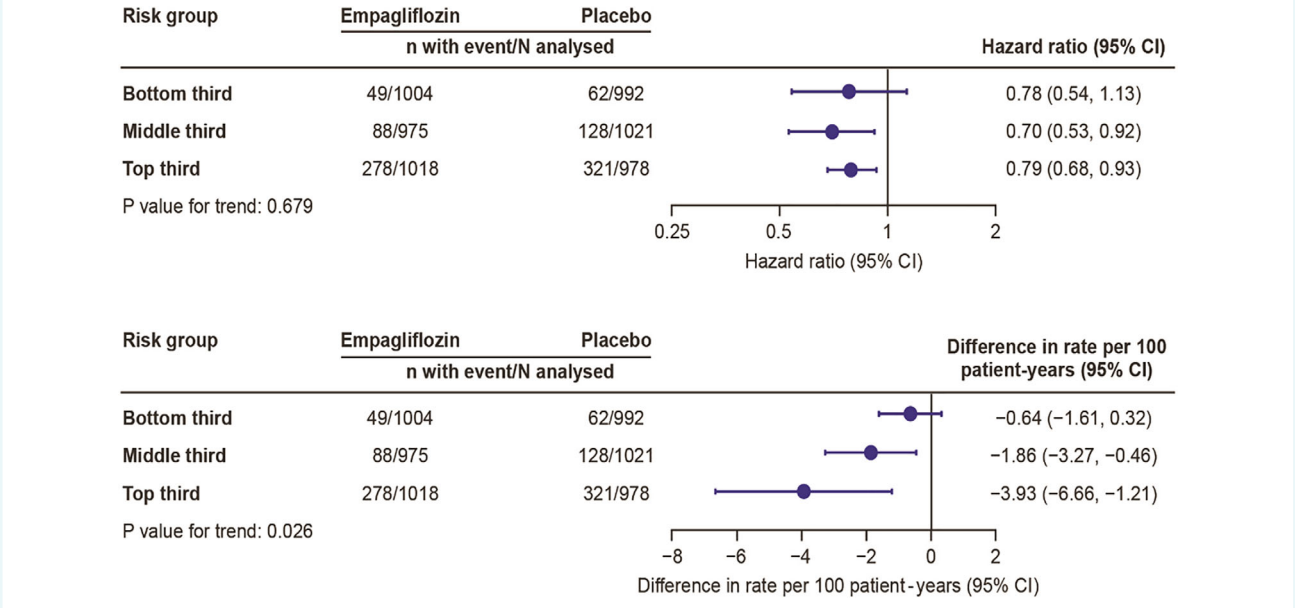


Figure 2 Hazard ratios and rate differences for empagliflozin versus placebo for the primary outcome according to thirds of the risk score distribution. CI, confidence interval.

Heart failure hospitalization only

There is interest in also determining risk models for incidence of HF hospitalization only. The same set of predictors are identified as for the primary endpoint, with the exception that diabetes is no longer such a powerful predictor (online supplementary *Table S4A*). The *c*-statistic is 0.787 which is higher than for the primary endpoint, and the goodness of it and strength of prediction for this model are shown in online supplementary *Figure S2*. In addition to NT-proBNP and hs-cTnT remaining powerful predictors, a recent prior HF hospitalization predicts particularly strongly.

Adding in UACR and KCCQ-OSS slightly enhances prediction (*c* becomes 0.793), and this fuller model is in online supplementary *Table S4B*. Further inclusions of urea nitrogen and total bilirubin as significant predictors increase *c* to 0.798 but of course add to model complexity.

All-cause and cardiovascular mortality

Death from any cause occurred in 849 (14.2%) of 5988 patients, 422 in the empagliflozin group and 427 in the placebo group (HR 1.00, 95% CI 0.87–1.15). A new risk model for mortality, using the same set of candidate predictors as above, yielded eight variables that achieved the inclusion criterion of $p < 0.001$ (*Table 2*). NT-proBNP and hs-cTnT remained the most dominant risk factors for mortality, followed by history of COPD, lower albumin, older age (especially over 75 years), LVEF <50%, NYHA class III/IV and insulin-treated diabetes. The distributions of all five quantitative risk factors are presented in online supplementary *Figure S2*.

The *c*-statistic for this prognostic model for all-cause death is 0.715 (95% CI 0.697–0.733) and *Figure 1B* shows the model's goodness of fit and strength of prediction. Comparing top and bottom deciles of risk, the observed 2-year mortality rates are 33.2% and 3.3%, respectively. A model that also included UACR and KCCQ-OSS slightly enhanced prediction (*c* becomes 0.721) and is shown in online supplementary *Table S3B*.

A cardiovascular cause accounted for 55% of all deaths ($n = 463$): 219 in empagliflozin and 244 in placebo (HR 0.91, 95% CI 0.76–1.09). Using the same eight variables as in the all-cause death model, a risk model for cardiovascular mortality yielded a very similar strength of prediction (*c*-statistic 0.718; 95% CI 0.694–0.741) (online supplementary *Table S2*).

A summary of our predictive models is presented in the *Graphical Abstract*.

Biomarker combination for risk prediction

Given the key roles of NT-proBNP and hs-cTnT in determining patient risk of the primary outcome and all-cause death, we explored how risk was affected by both markers. In *Figure 3* and online supplementary *Table S5*, we simultaneously stratified both NT-proBNP and hs-cTnT into thirds of their distribution and then showed how event rates of the primary outcome and all-cause death varied according to these nine patient groups. Patients with both the lowest NT-proBNP and lowest hs-cTnT had a primary event rate of 2.2 per 100 patient-years compared to 19.2 per 100 patient-years in those with highest NT-proBNP and hs-cTnT, with a rate ratio of 8.7. For all-cause death, a similar pattern emerged with a rate ratio of 6.3 for comparing the two extremes.

© 2022 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

Table 2 The EMPEROR-Preserved risk model for all-cause mortality

Variables	Hazard ratio (95% CI)	Chi-squared statistic	Coefficient (SE)[a]	*p*-value
Log hs-cTnT (ng/L)	1.52 (1.38–1.68)	68.6	0.42 (0.05)	<0.0001
Log NT-proBNP (pg/ml)	1.40 (1.30–1.52)	67.6	0.34 (0.04)	<0.0001
History of COPD	1.71 (1.45–2.01)	40.8	0.54 (0.08)	<0.0001
Baseline albumin (per 0.1 decrease below 4.5) (g/dl)	1.07 (1.05–1.10)	39.3	0.07 (0.01)	<0.0001
LVEF				<0.0001
≥50%	1.00 (reference)			
<50%	1.37 (1.19–1.57)	18.8	0.31 (0.07)	
Age				
<65 years	1.00 (reference)			<0.0001
65 to <75 years	1.20 (0.97–1.49)	2.9	0.18 (0.11)	
≥75 years	1.57 (1.28–1.92)	18.6	0.45 (0.10)	
NYHA class III/IV	1.39 (1.20–1.62)	18.1	0.33 (0.08)	<0.0001
Use of insulin in DM patients				0.0006
Non-DM	1.00 (reference)			
DM + no insulin	1.02 (0.88–1.20)	0.08	0.02 (0.08)	
DM + insulin	1.42 (1.18–1.71)	13.6	0.35 (0.10)	

CI, confidence interval; COPD, chronic obstructive pulmonary disease; DM, diabetes mellitus; HF, heart failure; HHF, hospitalization for heart failure; hs-cTnT, high-sensitivity cardiac troponin T; LVEF, left ventricular ejection fraction; NT-proBNP, N-terminal pro-B-type natriuretic peptide; NYHA, New York Heart Association; SE, standard error.

[a]Coefficient (SE) are the log hazard ratio and its standard error.

An estimate of each individual's 2-year risk can be calculated as follows: $1 - [0.99859^{\exp}$ $(0.42 \times \log$ hs-cTnT $+ 0.34 \times \log$ NT-proBNP $\times 0.33 \times$ NYHA class $+ 0.54 \times$ COPD $+ 0.07 \times (4.5 -$ Albumin$)/0.1) + 0.18 \times$ Age1 $+ 0.45 \times$ Age2 $+ 0.31 \times$ LVEF $+ 0.35 \times$ DM (insulin) $+ 0.02 \times$ DM (without insulin)], where 'Age1' and 'Age2' are indicator variables for whether the patient's age is 65 to <75 or ≥75 years, respectively. NYHA class is an indicator variable for whether the patient's NYHA class is III or IV. COPD is an indicator variable for whether the patient has COPD. LVEF is an indicator variable for whether the patient has LVEF <50%. DM (insulin) and DM (without insulin) are indicator variables for whether the patient has diabetes and use of insulin at baseline or has diabetes without use of insulin. *C*-statistic (95% CI): 0.715 (0.697–0.733).

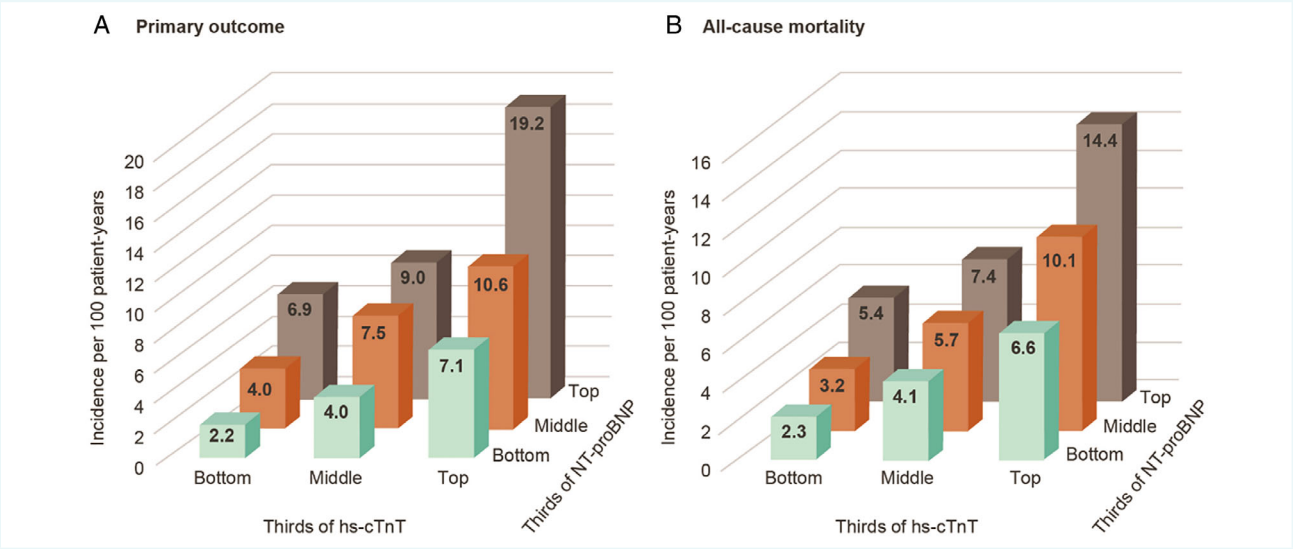


Figure 3 Incidence rates of (*A*) the primary outcome and (*B*) all-cause death for patients simultaneously grouped in thirds of N-terminal pro-B-type natriuretic peptide (NT-proBNP) and high-sensitivity cardiac troponin T (hs-cTnT).

There are strikingly consistent monotonic trends in risk for both biomarkers simultaneously for both outcomes. It is noteworthy that the two biomarkers are positively associated with a Pearson correlation of $r = 0.34$.

It is worth exploring how risk prediction models based on these two biomarkers only (i.e. ignoring the other predictors in *Tables 1* and *2*) would perform: the *c* statistics are 0.703 (for the primary endpoint) and 0.679 (for death) which are reasonable, but still

© 2022 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

substantially less than the corresponding 0.748 and 0.715 for the full models.

Models by ejection fraction subgroups

In online supplementary *Tables S6* and *S7* we present separately for patients with LVEF <50% and LVEF ≥50%, risk models for the primary outcome and all-cause mortality. There is a notable consistency of findings, whereby all factors show similar strengths of risk prediction for both LVEF categories, i.e. <50% and ≥50%.

External validation

External validation was performed in a subset of 1251 patients of the PARAGON-HF trial with available hs-cTnT and NT-proBNP at randomization.[9] A comparison of the relevant baseline variables in EMPEROR-Preserved and PARAGON-HF subset is shown in online supplementary *Table S8*. Patient characteristics were similar between the two trials. Online supplementary *Table S8* also compares the PARAGON-HF subset with other PARAGON-HF patients, and shows a broad similarity. In this PARAGON-HF subset over a median follow-up of 24 months, 223 (17.8%) patients had a composite of cardiovascular death or HF hospitalization, 143 (11.4%) patients died, of which 82 were cardiovascular deaths.

Applying the EMPEROR-Preserved risk models to the PARAGON-HF population gave *c*-statistics of 0.711 (95% CI 0.672–0.749) for the composite of cardiovascular death or HF hospitalization, 0.719 (95% CI 0.665–0.772) for all-cause death, and 0.718 (95% CI 0.652–0.784) for cardiovascular death. A monotonic increase in events was observed across quintiles of all three risk score distributions, and the models presented good calibration for all studied outcomes, despite that for all-cause death the observed deaths were lower than expected from the model (*Figure 4*).

We also applied the EMPEROR-Preserved risk model for HF hospitalization in online supplementary *Table S4A* to the PARAGON-HF subset and achieved $c = 0.712$ (0.668–0.757). Model goodness-of-fit and calibration is shown in online supplementary *Figure S4*.

Discussion

The biomarker-driven risk models that we present here provide effective risk discrimination for patients with HFpEF both for HF-related morbidity and mortality, as summarized in the *Graphical Abstract*. Their practical value is enhanced by the fact that these models require only two widely available biomarkers and a handful of clinical variables. The implementation of these prognostic models is furtherly facilitated by our online calculator (see online supplementary material), providing risk estimates for each individual patient embedded in routine clinical practice.

We have previously reported on the key roles of NT-proBNP and hs-cTnT in the prognosis of patients with HFrEF.[5] The present work demonstrates that these biomarkers are similarly important in the prediction of risk of patients with HFpEF. There is a dearth of risk models for HFpEF patients, particularly contemporary models that incorporate biomarkers with strong prognostic value.

Elevated levels of NT-proBNP have been associated with poor prognosis in patients with HFpEF. In the PARAGON-HF and I-Preserve trials, NT-proBNP was a robust predictor of cardiovascular death and HF hospitalizations.[3,10] The prognostic value of cardiac troponins in HFpEF is less well established than that of NT-proBNP, but some studies support the strong prognostic value of cardiac troponins in patients with HFpEF.[11–13] Combining NT-proBNP and hs-cTnT adds complementary and independent prognostic information.

Other highly significant predictors of the primary outcome were a recent HF hospitalization, NYHA class III or IV, history of COPD, insulin-treated diabetes, low haemoglobin, and a longer HF duration. Patients with HFpEF, a recent HF hospitalization and worse symptoms have a high risk of subsequent rehospitalizations and death.[14,15] Patients with HFpEF and COPD have a shared clinical presentation including signs and symptoms (e.g. peripheral oedema and breathlessness) and elevated natriuretic peptides. The presence of both conditions confers a worse prognosis than either disease alone.[16,17] Diabetes mellitus is a frequent comorbidity among HFpEF patients which is associated with a poor prognosis, particularly when requiring insulin for lowering glycaemia which occurs more often in long-standing diabetes with difficult glycaemic control and concomitant cardiovascular complications.[18] Low haemoglobin levels and anaemia are also frequent among patients with HFpEF and associated with worse symptoms and a poor prognosis.[19] A longer duration of HFpEF has been associated with a poor prognosis due to the cumulative organ damage caused by the disease over time.[20]

The all-cause and cardiovascular death risk models had some overlap with the primary outcome model (that also included NT-proBNP and hs-cTnT as the most important predictors, plus COPD, NYHA class III or IV, and insulin-treated diabetes). However, three other highly significant independent predictors (i.e. lower albumin level, older age and LVEF <50%) were retained in the mortality prediction models instead of time since last HF hospitalization, HF duration and low haemoglobin. It has been proposed that low albumin is a useful marker of patient nutritional status, liver and renal dysfunction, and frailty.[21,22] Older age predicts mortality better than it predicts HF hospitalizations; while the former is an inexorable event with ageing, the latter is less dependent on age. It is also noteworthy that across the spectrum of HFpEF patients, those with a moderately lower LVEF in the range 41% to 49% do carry a 37% excess mortality risk compared to those with an LVEF ≥50%. However, when risk models were separately produced for patients with LVEF below and above 50%, the contributions of the other risk factors were broadly similar.

To the best of our knowledge, only four risk models have been developed for HFpEF: CHARM,[1] MAGGIC,[2] I-Preserve,[3] and 3A3B score.[4] The CHARM and MAGGIC risk scores did not include B-type natriuretic peptide (BNP) or NT-proBNP which are the single most important biomarkers for prognostic assessment in HF. Still, the levels of BNP were added to the MAGGIC score *a posteriori* showing a strong association with outcomes.[2] I-Preserve included NT-proBNP, but the models were rather complex with

© 2022 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

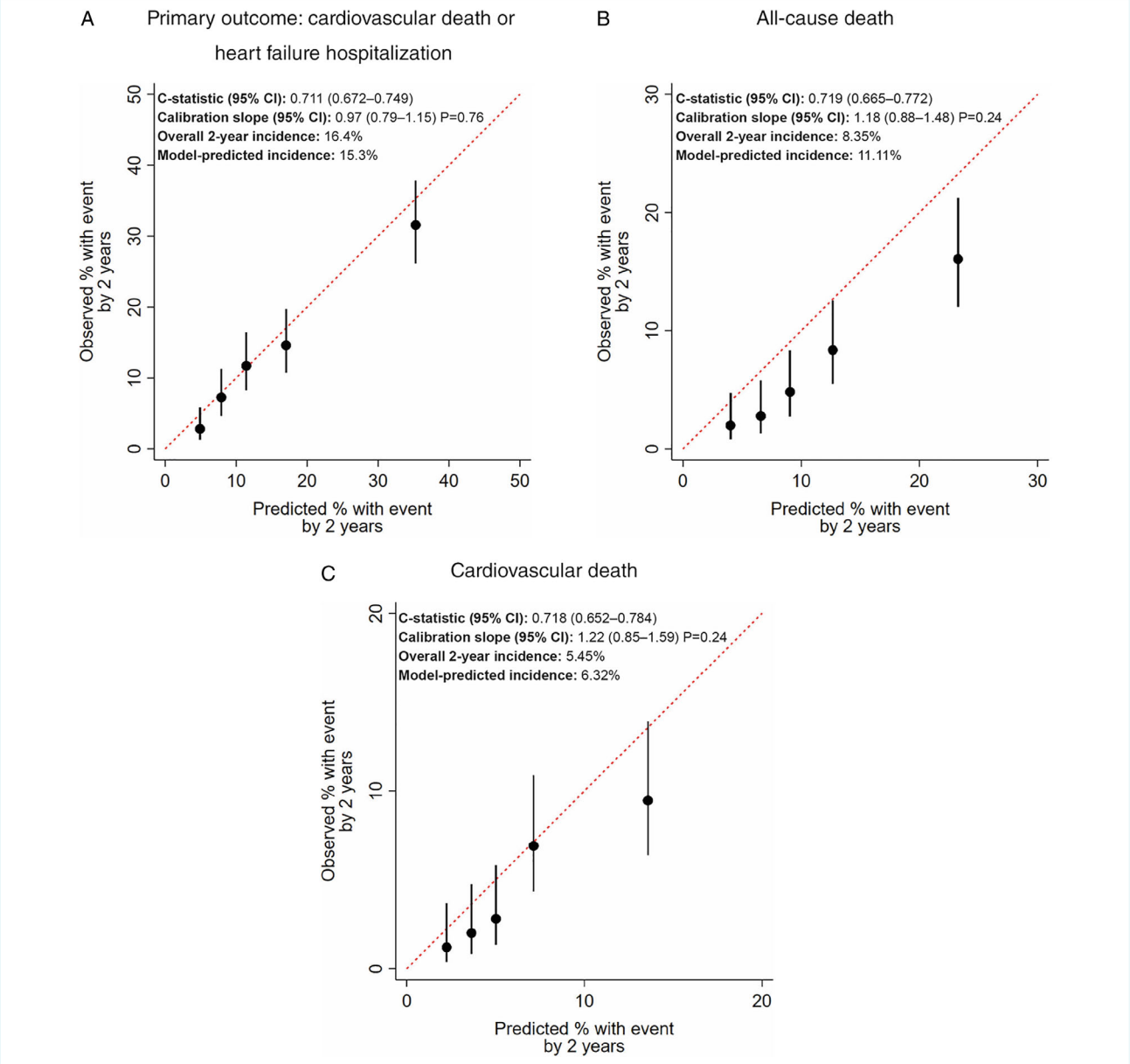


Figure 4 Observed versus predicted events by fifths of the risk score distribution in PARAGON-HF, along with metrics for discrimination and calibration. (*A*) Primary outcome: cardiovascular death or heart failure hospitalization. (*B*) All-cause death. (*C*) Cardiovascular death. CI, confidence interval.

12 variables, and they included quality of life scores not routinely performed in clinical practice.[3] The 3A3B score was tested in Japanese patients with LVEF ≥50% from the CHART-2 registry to predict mortality from any cause and included natriuretic peptide levels.[4] Its risk scores did not include cardiac troponins or the combination of natriuretic peptides with cardiac troponins.

It would have been useful to assess the predictive strength of the CHARM, I-Preserve and 3A3B risk models in our two patient cohorts, but this is not possible because some variables are not available. For the MAGGIC risk score we obtained relatively poor risk discrimination for both the primary endpoint and death in both EMPEROR-Preserved and PARAGON-HF cohorts (all *c*-statistics under 0.64), reflecting that the MAGGIC score was not specifically aimed at HFpEF, and was also developed before key biomarkers were routinely recorded.

The prognostic capacity of our risk models was useful with *c*-statistics ranging from 0.71 to 0.75 for the prediction of mortality and the composite of HF hospitalizations or cardiovascular death, respectively. These models may allow clinicians to accurately predict the event probability in each individual patient in routine clinical practice (using our online calculator), potentially allowing

© 2022 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

to incorporate each patient's risk to better establish follow-up and care plans with patients and their families.

We have also presented risk models for predicting HF hospitalizations only, with a higher *c*-statistic of 0.79. They have a methodological complication in that all deaths that happened without a prior HF hospitalization are handled as censorings, so interpretation of these models (like any model of a non-fatal event accompanied by the competing risk of death) is somewhat challenging.

The external validation replicating the good prognostic capacity of these models in a subset of 1251 patients from the PARAGON-HF trial, reinforces the capacity of our models to be applied to different HFpEF populations.

Importantly, our biomarker-driven risk models are simple to use since they rely on only a few readily available clinical variables plus NT-proBNP and hs-cTnT, which can be easily obtained in most clinical settings. Our findings support the joint assessment of NT-proBNP and hs-cTnT in the comprehensive risk assessment of all HF patients, both HFpEF (as reported here) and HFrEF (as previously reported).[5]

In addition to the variables included in our models, health status assessed by the KCCQ-OSS and albuminuria had highly significant associations with the studied outcomes; however, we have decided not to incorporate these variables in our final models because they are not routinely measured in clinics and their addition only slightly improved the *c*-statistics of our models, that were largely driven by NT-proBNP and hs-cTnT.

Despite its prognostic utility, the risk score does not allow to make decisions about whom to treat with empagliflozin because all patients benefit similarly in terms of relative risk reduction; still, patients with a higher baseline risk may experience a greater absolute benefit.

Limitations

Despite the good performance and external validation of our models, both EMPEROR-Preserved and PARAGON-HF are clinical trials with specific eligibility criteria, hence one cannot assume that these findings apply to all HFpEF patients.

Specifically, EMPEROR-Preserved excluded (i) patients with *de novo* HFpEF within 3 months of diagnosis, (ii) patients with acute decompensated HF within the past week, and (iii) patients with recent myocardial infarction within 3 months, and hence our risk models are not applicable in any such patients. The latter ensures that a troponin elevation relates to chronic HFpEF and not any recent acute coronary syndrome event.

Our risk models' predictive strength are reasonable but there is inevitable room for improvement. Individual patient risk of death and cardiovascular events can never be fully captured by biomarkers and clinical variables: factors such frailty, socio-economic status, diet and health care system delivery all play a part.

Conclusions

The combination of NT-proBNP and hs-cTnT with a few readily available clinical variables provide effective risk discrimination for both morbidity and mortality in patients with HFpEF. A predictive tool-kit is provided in the online supplementary material to facilitate the ready implementation of our novel risk models in routine clinical practice.

Supplementary Information

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

[Correction added on 16 September 2022, after first online publication: Supplementary files have been updated in this version.]

Funding

The EMPEROR-Reduced trial was supported by Boehringer Ingelheim and Eli Lilly and Company. Dr. Januzzi is supported in part by the Hutter Family Professorship at Harvard Medical School.

Conflict of interest: S.J.P. is a consultant to Boehringer Ingelheim. J.P.F. is a consultant for Boehringer Ingelheim. M.P. reports personal fees from Boehringer Ingelheim, during the conduct of the study; personal fees from Abbvie, Akcea, Amarin, AstraZeneca, Amgen, Boehringer Ingelheim, Daiichi Sankyo, Johnson & Johnson, Lilly, Novartis, ParatusRx, Pfizer, Relypsa, Sanofi, Synthetic Biologics and Theravance outside the submitted work. F.Z. has recently received steering committee or advisory board fees from Amgen, AstraZeneca, Bayer, Boehringer Ingelheim, Boston Scientific, Cardior, CVRx, Janssen, Livanova, Merck, Mundipharma, Novartis, Novo Nordisk, and Vifor Fresenius. G.F. reports lectures and/or Committee Member contributions in trials sponsored by Medtronic, Vifor, Servier, Novartis, Bayer, Amgen, and Boehringer Ingelheim. J.L.J. is a Trustee of the American College of Cardiology, a Board member of Imbria Pharmaceuticals, has received grant support from Applied Therapeutics, Innolife, Novartis Pharmaceuticals and Abbott Diagnostics, consulting income from Abbott, Janssen, Novartis, and Roche Diagnostics, and participates in clinical endpoint committees/data safety monitoring boards for Abbott, AbbVie, Amgen, Bayer, CVRx, Janssen, MyoKardia and Takeda. J.B. is a consultant for Abbott, Adrenomed, Amgen, Array, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol Myers Squibb, CVRx, G3 Pharmaceutical, Impulse Dynamics, Innolife, Janssen, LivaNova, Luitpold, Medtronic, Merck, Novartis, NovoNordisk, Relypsa, Roche, V-Wave Limited, and Vifor. S.D.A has received fees from Abbott, Bayer, Boehringer Ingelheim, Cardiac Dimension, Cordio, Impulse Dynamics, Novartis, Occlutech, Servier, and Vifor Pharma, and grant support from Abbott and Vifor Pharma. All other authors have nothing to disclose.

References

1. Pocock SJ, Wang D, Pfeffer MA, Yusuf S, McMurray JJ, Swedberg KB, et al. Predictors of mortality and morbidity in patients with chronic heart failure. *Eur Heart J*. 2006;**27**:65–75.
2. Rich JD, Burns J, Freed BH, Maurer MS, Burkhoff D, Shah SJ. Meta-Analysis Global Group in Chronic (MAGGIC) Heart Failure risk score: validation of a simple tool for the prediction of morbidity and mortality in heart failure with preserved ejection fraction. *J Am Heart Assoc*. 2018;**7**:e009594.
3. Komajda M, Carson PE, Hetzel S, McKelvie R, McMurray J, Ptaszynska A, et al. Factors associated with outcome in heart failure with preserved ejection fraction: findings from the Irbesartan in Heart Failure with Preserved Ejection Fraction study (I-PRESERVE). *Circ Heart Fail*. 2011;**4**:27–35.

© 2022 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

4. Kasahara S, Sakata Y, Nochioka K, Tay WT, Claggett BL, Abe R, et al. The 3A3B score: the simple risk score for heart failure with preserved ejection fraction – a report from the CHART-2 study. *Int J Cardiol*. 2019; **284**:42–9.
5. Pocock SJ, Ferreira JP, Gregson J, Anker SD, Butler J, Filippatos G, et al. Novel biomarker-driven prognostic models to predict morbidity and mortality in chronic heart failure: the EMPEROR-Reduced trial. *Eur Heart J*. 2021;**42**: 4455–64.
6. Anker SD, Butler J, Filippatos G, Ferreira JP, Bocchi E, Böhm M, et al.; EMPEROR-Preserved Trial Investigators. Empagliflozin in heart failure with a preserved ejection fraction. *N Engl J Med*. 2021;**385**:1451–61.
7. Anker SD, Butler J, Filippatos GS, Jamal W, Salsali A, Schnee J, et al. Evaluation of the effects of sodium-glucose co-transporter 2 inhibition with empagliflozin on morbidity and mortality in patients with chronic heart failure and a preserved ejection fraction: rationale for and design of the EMPEROR-Preserved trial. *Eur J Heart Fail*. 2019;**21**:1279–87.
8. Solomon SD, McMurray JJV, Anand IS, Ge J, Lam CSP, Maggioni AP, et al.; PARAGON-HF Investigators and Committees. Angiotensin-neprilysin inhibition in heart failure with preserved ejection fraction. *N Engl J Med*. 2019;**381**: 1609–20.
9. Gori M, Senni M, Claggett B, Liu J, Maggioni AP, Zile M, et al. Integrating high-sensitivity troponin T and sacubitril/valsartan treatment in HFpEF: the PARAGON-HF trial. *JACC Heart Fail*. 2021;**9**:627–35.
10. Cunningham JW, Vaduganathan M, Claggett BL, Zile MR, Anand IS, Packer M, et al. Effects of sacubitril/valsartan on N-terminal pro-B-type natriuretic peptide in heart failure with preserved ejection fraction. *JACC Heart Fail.* 2020;**8**: 372–81.
11. Myhre PL, O'Meara E, Claggett BL, de Denus S, Jarolim P, Anand IS, et al. Cardiac troponin I and risk of cardiac events in patients with heart failure and preserved ejection fraction. *Circ Heart Fail*. 2018;**11**:e005312.
12. Gohar A, Chong JPC, Liew OW, den Ruijter H, de Kleijn DPV, Sim D, et al. The prognostic value of highly sensitive cardiac troponin assays for adverse events in men and women with stable heart failure and a preserved vs. reduced ejection fraction. *Eur J Heart Fail*. 2017;**19**:1638–47.
13. Suzuki S, Motoki H, Minamisawa M, Okuma Y, Shoin W, Okano T, et al. Prognostic significance of high-sensitivity cardiac troponin in patients with heart failure with preserved ejection fraction. *Heart Vessels*. 2019;**34**:1650–6.
14. Vaduganathan M, Claggett BL, Desai AS, Anker SD, Perrone SV, Janssens S, et al. Prior heart failure hospitalization, clinical outcomes, and response to sacubitril/valsartan compared with valsartan in HFpEF. *J Am Coll Cardiol*. 2020;**75**:245–54.
15. Dalos D, Mascherbauer J, Zotter-Tufaro C, Duca F, Kammerlander AA, Aschauer S, et al. Functional status, pulmonary artery pressure, and clinical outcomes in heart failure with preserved ejection fraction. *J Am Coll Cardiol*. 2016;**68**:189–99.
16. Marcun R, Stankovic I, Vidakovic R, Farkas J, Kadivec S, Putnikovic B, et al. Prognostic implications of heart failure with preserved ejection fraction in patients with an exacerbation of chronic obstructive pulmonary disease. *Intern Emerg Med*. 2016;**11**:519–27.
17. Mooney L, Hawkins NM, Jhund PS, Redfield MM, Vaduganathan M, Desai AS, et al. Impact of chronic obstructive pulmonary disease in patients with heart failure with preserved ejection fraction: insights from PARAGON-HF. *J Am Heart Assoc*. 2021;**10**:e021494.
18. Shen L, Rørth R, Cosmi D, Kristensen SL, Petrie MC, Cosmi F, et al. Insulin treatment and clinical outcomes in patients with diabetes and heart failure with preserved ejection fraction. *Eur J Heart Fail*. 2019;**21**:974–84.
19. Gupta K, Kalra R, Rajapreyar I, Joly JM, Pate M, Cribbs MG, et al. Anemia, mortality, and hospitalizations in heart failure with a preserved ejection fraction (from the TOPCAT trial). *Am J Cardiol*. 2020;**125**:1347–54.
20. Owan TE, Hodge DO, Herges RM, Jacobsen SJ, Roger VL, Redfield MM. Trends in prevalence and outcome of heart failure with preserved ejection fraction. *N Engl J Med*. 2006;**355**:251–9.
21. Liu M, Chan CP, Yan BP, Zhang Q, Lam YY, Li RJ, et al. Albumin levels predict survival in patients with heart failure and preserved ejection fraction. *Eur J Heart Fail*. 2012;**14**:39–44.
22. Prenner SB, Kumar A, Zhao L, Cvijic ME, Basso M, Spires T, et al. Effect of serum albumin levels in patients with heart failure with preserved ejection fraction (from the TOPCAT trial). *Am J Cardiol*. 2020;**125**:575–82.

© 2022 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.