

Variation in clinical and patient-reported outcomes among complex heart failure with preserved ejection fraction phenotypes

Kelsey M. Flint^{1*}, Sanjiv J. Shah², Eldrin F. Lewis³ and David P. Kao⁴

¹Division of Cardiology, Rocky Mountain Regional VA Medical Center, Cardiology (111), Building F2, Room 143, Aurora, CO 80045, USA; ²Division of Cardiology, Department of Medicine, Feinberg Cardiovascular Research Institute, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; ³Cardiovascular Division, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; ⁴Division of Cardiology, University of Colorado School of Medicine, Aurora, CO, USA

Abstract

Aims The aim of this study is to use six previously described heart failure with preserved ejection fraction (HFpEF) phenotypes to describe differences in (i) the biological response to spironolactone, (ii) clinical endpoints, and (iii) patient-reported health status by HFpEF phenotype and treatment arm in the Treatment of Preserved Cardiac Function Heart Failure with an Aldosterone Antagonist Trial (TOPCAT).

Methods and results We analysed 1767 patients in TOPCAT from the Americas. Using 11 clinical variables, patients were classified according to six HFpEF phenotypes previously identified in the I-PRESERVE and CHARM-Preserved studies. Kansas City Cardiomyopathy Questionnaire (KCCQ) measured health status. All phenotypes showed increase in potassium with spironolactone, although only three phenotypes showed significant increase in creatinine, and two phenotypes showed significant decrease in systolic blood pressure. Rate of the TOPCAT primary outcome (cardiovascular death, aborted cardiac arrest, or heart failure hospitalization) differed by HFpEF phenotype ($P < 0.001$) but not by treatment arm within each HFpEF phenotype. Baseline KCCQ score differed by HFpEF phenotype ($P < 0.001$), although some phenotypes with poor health status had lower rates of the TOPCAT primary outcome, and some phenotypes with better health status had higher rates of the TOPCAT primary outcome. However, within 3/6 phenotypes, higher baseline KCCQ score was associated with lower risk of the TOPCAT primary outcome. Change in KCCQ scores at 4 and 12 months did not differ among HFpEF phenotypes overall or by treatment arm.

Conclusions Complex, data-driven HFpEF phenotypes differ according to biological response to spironolactone, baseline health status, and clinical endpoints. These differences may inform the design of targeted clinical trials focusing on improvement in outcomes most relevant for specific HFpEF phenotypes.

Keywords HFpEF; Health status; Mortality; Hospitalization; Heart failure

Received: 15 August 2019; Revised: 21 November 2019; Accepted: 6 February 2020

*Correspondence to: Kelsey Flint, MD, MSCS, Division of Cardiology, Rocky Mountain Regional VA Medical Center, Cardiology (111), Building F2, Room 143, Aurora, CO 80045, USA. Email: kelsey.flint@ucdenver.edu
Clinical trial: NCT00094302; clinicaltrials.gov

Introduction

Despite tremendous effort and expense, there are currently no evidence-based treatments for improving mortality or reducing hospitalization in patients with heart failure with preserved ejection fraction (HFpEF).^{1,2} Among the likely reasons for the disappointing results of numerous clinical trials are marked clinical and physiologic heterogeneity among patients

with HFpEF.³ A more nuanced understanding of HFpEF subpopulations may help target new interventions to improving specific outcomes most relevant to each phenotype. For example, in phenotypes with poor health status but low adverse clinical event rates like hospitalization or death, effective treatment response might be defined as improvement in symptom burden rather than improving survival and reducing hospitalization.

Existing frameworks for categorizing patients with HFpEF generally focus on comorbidity burden.⁴ Such classifications are clinically salient but may oversimplify the complex and often overlapping physiology of patients with HFpEF.⁵ To address this issue, our group built upon previously described HFpEF frameworks by defining distinct cluster-based HFpEF phenotypes and examining clinical endpoints using I-PRESERVE and CHARM-Preserved,^{6–8} the two largest HFpEF intervention trials to date. These phenotypes were identified using latent class analysis of 11 demographic, clinical, and laboratory variables widely available in routine clinical practice. These six phenotypes had significantly different rates of hospitalization and mortality in both derivation and validation cohorts even when adjusted for individual cluster component variables. This suggests that the specific cluster of clinical features that define each phenotype is a clinically useful beyond the individual clinical features.⁶

The Treatment of Preserved Cardiac Function Heart Failure with an Aldosterone Antagonist Trial (TOPCAT) trial tested the hypothesis that the aldosterone antagonist spironolactone would improve a composite primary outcome of cardiovascular (CV) mortality, hospitalization for HF, and aborted sudden cardiac death (TOPCAT primary outcome) in patients with HFpEF.⁹ Overall TOPCAT showed no significant reduction in the primary outcome associated with spironolactone vs. placebo, but there was a signal for efficacy of spironolactone in patients enrolled in the Americas.¹⁰ In the current study, we aimed to (i) validate our group's prior observations regarding characteristics and outcomes according to previously identified data-driven HFpEF phenotypes in this distinct HFpEF clinical trial population, (ii) characterize phenotype-specific biological response (i.e. change in blood pressure and serum potassium and creatinine over time) to spironolactone in the treatment vs. placebo arms within each phenotype, and (iii) determine whether HF-specific health status differed by HFpEF phenotype and treatment arm at baseline, over time, and whether the prognostic value of HF-specific health status differed by phenotype. Our hypothesis was that complex HFpEF phenotypes show differences in biologic response, health status, clinical outcomes, and treatment response, which may help target treatments and clinical trials to outcomes that are most relevant to each phenotype.

Methods

Data source

Data from the TOPCAT were obtained from the National Heart, Lung, and Blood Institute's BioLINCC resource.¹¹ The methods and primary results of the trial are published elsewhere.⁹ Briefly, TOPCAT enrolled 3445 patients at 233 sites in six countries (USA, Canada, Brazil, Argentina, Russia, and

Georgia). Enrolment criteria included age ≥ 50 years, left ventricular ejection fraction $\geq 45\%$, and hospitalization for HF (not adjudicated) in the past 12 months or elevated BNP in the past 6 months. The primary outcome was a composite of death from CV causes, aborted sudden cardiac death, or hospitalization for HF. Components of the TOPCAT primary outcome were adjudicated by a clinical endpoints committee according to prespecified criteria. The primary outcome was numerically but not significantly different between the spironolactone and placebo arms of the trial, although spironolactone was associated with reduced risk of hospitalizations for HF.⁹

In post hoc analyses, there appeared to be a difference in the rate of the primary outcome between the spironolactone and placebo arms in patients enrolled from the Americas, but not from Russia and Georgia.¹² On further examination, there were significant concerns over whether patients enrolled in Russia and Georgia truly had HF based upon (i) high enrolment in these countries using the unadjudicated inclusion criterion of HF hospitalization, (ii) very low event rates in patients enrolled in Russia and Georgia compared with the Americas, and (iii) high prevalence of normal BNP values in a substudy mandated by the TOPCAT Data Safety and Monitoring Board.¹² Furthermore, longitudinal analysis of potassium, creatinine, and systolic blood pressure and metabolites of spironolactone suggested poor adherence to the study drug in participants from Russia and Georgia.^{10,13,14} Consequently, the current analysis will focus only on patients from the Americas, similar to other investigators.^{15,16}

Clinical outcomes

The TOPCAT primary outcome served as our primary outcome. Secondary outcomes in the current analysis include the I-PRESERVE composite primary outcome (all-cause mortality and CV hospitalization)⁸ used in the phenotype derivation (I-PRESERVE) and validation (CHARM-Preserved) cohorts to validate previously observed outcome differences between HFpEF phenotypes.⁶ Other secondary outcomes included all-cause, CV, HF, and non-CV hospitalizations.

Measurements

Blood pressure, serum creatinine, and serum potassium were analysed at baseline, 1, 2, 4, 8, 12, and 24 months. HF-specific health status was measured using the Kansas City Cardiomyopathy Questionnaire (KCCQ).¹⁷ The KCCQ is a 23-question HF-specific health status tool that measures responses in five domains: total symptom burden, social limitation, physical limitation, self-efficacy, and quality of life. These domains are incorporated into the KCCQ overall summary score, all of which are expressed as a range from 0 to 100 with lower

scores indicating worse health status.¹⁷ Although the KCCQ was derived and validated in patients with heart failure with reduced ejection fraction, it has similar distribution, internal consistency (Cronbach's alpha 0.96), and validity (correlation with New York Heart Association scale $r = -0.62$, $P < 0.001$) in patients with HFpEF.¹⁸ The KCCQ is also predictive of death and hospitalization in HFpEF patients.¹⁸ KCCQ was collected at baseline, 4, 12, 24, 36, and 48 months.

Valvular heart disease (VHD) was defined as moderate or severe valvular regurgitation or stenosis. Sex and presence of atrial fibrillation, diabetes, coronary artery disease, and hyperlipidaemia were collected at study entry by patient report. Age was calculated at study enrolment based on patient-reported date of birth and was divided into the following categories: 60–70, 71–80, and >80 years old. Haemoglobin was measured via blood test at study enrolment and divided into the following categories: <6.7, 6.8–10.0, 10.1–13.3, 13.4–16.7, and >16.8 g/dL. Creatinine was measured via blood test at study enrolment and used to calculating estimated glomerular filtration rate from the Chronic Kidney Disease Epidemiology Collaboration equation.¹⁹ Patients were categorized into CKD stages 1–5 based on standard definitions.²⁰ Body mass index (BMI) was calculated from height and weight measured at study entry physical exam, and divided into categories based on the World Health Organization Classification of underweight, normal weight, overweight and obese (<http://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi>). Alcohol use was assessed at study entry based on the question 'how many alcoholic drinks has the subject consumed?' We dichotomized the result into any alcohol use vs. none.

Statistical analysis

For a full description of how the HFpEF phenotypes were derived and validated, please see Kao et al and Supporting Information, *Table S1* with that publication.⁶ Briefly, the derivation of the original phenotypes from the I-PRESERVE data was achieved using latent class analysis with the polCA library in the R statistical package (v 2.15.1; R Foundation for Statistical Computing, Vienna, Australia). Latent class definitions were derived using maximum likelihood estimation to identify the most common patterns among 11 variables that are clinically relevant and easy to obtain in routine clinical care. The optimal number of subgroups was determined using the first minima of the Bayesian information criterion and χ^2 statistic. Through this process, each of the 11 clinical variables used to create the phenotypes was assigned a coefficient. These coefficients were applied to each Americas TOPCAT participant, resulting in the probability of a given patient belonging to each class. TOPCAT participants were assigned to the phenotype for which they had the highest probability of membership.

Differences between baseline demographic, clinical, and health status values by HFpEF phenotype were assessed by Kruskal–Wallis and χ^2 tests for continuous and categorical variables, respectively. All survival curves were constructed using the Kaplan–Meier method. After confirming the proportional hazards assumption for phenotype and treatment arm, differences in time-to-event between phenotypes (both unadjusted and adjusted for the clinical variables used to construct the phenotypes) were assessed using Cox proportional hazards models. All analyses were performed using R (v3.3.0; R Foundation for Statistical Computing, Vienna, Austria).

Results

HFpEF phenotypes

Characteristics of patients enrolled in the Americas stratified by the six HFpEF phenotypes are shown in *Table 1*. The proportion of patients falling into each phenotype differed in TOPCAT compared with the previous derivation (I-PRESERVE) and validation (CHARM-Preserved) cohorts, although clinical profiles of each phenotype were similar. Overlap between the phenotypes is low. *Figure S1* shows the probability of HFpEF phenotype membership by HFpEF phenotype and shows that participants assigned to Phenotypes A, C, D, E, and F had a relatively high probability of membership in their assigned phenotype. However, some participants assigned to Phenotype B also had a moderate probability of belonging to Phenotype D, suggesting some overlap between Phenotypes B and D.

As described previously,⁶ Phenotypes A and E are notable for being all male patients, whereas Phenotypes B and D are composed of mostly female patients. Phenotype A had younger male patients (mean age 62) who were obese (mean BMI 35) (i.e. younger, obese men phenotype). Phenotype B had the youngest mean age of the phenotypes (61 years old) and reported the highest number of metabolic equivalents of all the phenotypes (i.e. younger, active women phenotype). Phenotype C is notable for having nearly ubiquitous prevalence of obesity, diabetes, and hypertension (i.e. metabolic syndrome phenotype). Phenotype D had older women (mean age 72) with the second highest proportion of diabetes and VHD among the phenotypes (i.e. older, diabetic women with VHD phenotype). Phenotype E had older men (mean age 76) and reported the highest alcohol use among the phenotypes (i.e. older male alcohol users phenotype). Patients in Phenotype F were mostly female, had the oldest mean age of the phenotypes (83 years old), the highest prevalence of atrial fibrillation, and had the lowest BMI and haemoglobin values (i.e. frail older women phenotype). For further details, see *Tables 1* and *2*.

Table 1 Baseline characteristics by heart failure with preserved ejection fraction phenotype (Americas) using phenotype definition variables. Data are presented as number (%) or mean \pm standard deviation.

Characteristic*	Total N = 1767					
	A	B	C	D	E	F
162 (9)	146 (8)	538 (30)	305 (17)	297 (17)	319 (18)	
Phenotype definition variables						
Age	61.6 \pm 5.4	60.7 \pm 6.0	68.2 \pm 8.1	72.3 \pm 7.0	75.6 \pm 6.3	82.7 \pm 5.6
Female	0 (0)	126 (86)	235 (44)	305 (100)	0 (0)	216 (68)
BMI	35.0 \pm 7.7	36.0 \pm 8.2	38.1 \pm 7.9	33.5 \pm 7.8	30.5 \pm 6.4	28.5 \pm 6.7
Obesity (BMI \geq 30)	118 (73)	108 (74)	496 (89)	201 (66)	127 (43)	114 (36)
Atrial fibrillation	51 (31)	10 (7)	182 (34)	149 (49)	164 (56)	743 (42)
Coronary artery disease	68 (42)	59 (40)	322 (60)	89 (29)	151 (51)	126 (40)
Diabetes mellitus	63 (39)	44 (30)	534 (99)	42 (42)	50 (17)	788 (45)
Hyperlipidaemia	104 (64)	88 (60)	587 (91)	186 (61)	198 (67)	187 (59)
Valvular heart disease ^a	3/57 (5)	1/52 (2)	29/206 (14)	23/97 (24)	10/105 (10)	29/118 (25)
Alcohol ^b	63 (39)	33 (23)	92 (17)	67 (22)	127 (43)	84 (26)
eGFR, mL/min/1.73m ²	80.9 \pm 20.0	90.2 \pm 27.7	57.3 \pm 18.6	63.4 \pm 16.1	66.8 \pm 16.9	55.4 \pm 17.2
Haemoglobin (g/dL)	14.6 \pm 1.6	13.0 \pm 1.3	12.1 \pm 1.5	12.8 \pm 1.8	13.7 \pm 1.7	12.3 \pm 1.5
Variables not included in phenotype definition						
White	134 (83)	98 (67)	391 (73)	234 (77)	259 (87)	268 (84)
History of MI	30 (19)	24 (16)	146 (27)	31 (10)	77 (26)	51 (16)
Hypertension	140 (86)	133 (91)	519 (96)	270 (89)	248 (84)	278 (87)
COPD or asthma	37 (23)	37 (25)	144 (27)	63 (21)	72 (24)	63 (20)
Tobacco use	25 (15)	25 (17)	29 (5)	10 (3)	17 (6)	11 (3)
Metabolic equivalents/week	10.7 \pm 12.9	12.1 \pm 24.6	9.6 \pm 26.4	9.2 \pm 11.6	9.8 \pm 15.2	9.5 \pm 12.8
KCCQ overall score	59.3 \pm 24.6	53.1 \pm 23.2	53.0 \pm 23.5	58.2 \pm 21.1	68.1 \pm 22.8	59.2 \pm 22.3
NYHA Class 3 or 4	44 (27)	44 (30)	235 (44)	86 (28)	72 (24)	139 (44)
Medications:						
Diuretic	128 (80)	119 (82)	513 (95)	268 (88)	262 (89)	283 (89)
ACE-I or ARB	139 (86)	117 (80)	462 (86)	235 (77)	222 (75)	220 (69)
Beta blocker	128 (80)	101 (69)	456 (85)	239 (78)	224 (76)	239 (75)
CCB	50 (31)	46 (32)	239 (44)	119 (39)	99 (33)	129 (40)

Age, BMI, eGFR, and haemoglobin are reported as continuous data (mean \pm SD) for ease of interpretation. Remaining categorical data are presented as N (%). ACE-I, angiotensin converting enzyme inhibitor; ARB, aldosterone receptor antagonist; BMI, body mass index; CCB, calcium channel blocker; COPD, chronic obstructive pulmonary disease; eGFR, estimated glomerular filtration rate; KCCQ, Kansas City Cardiomyopathy Questionnaire; MI, myocardial infarction NYHA, New York Heart Association.

^aEcho substudy only.

^bPercent (number) of patients reporting any alcohol intake.

* $P < 0.001$ for all characteristics except COPD/asthma ($P = 0.19$) and metabolic equivalents per week ($P = 0.052$).

Table 2 Summary of pertinent clinical, health status, and outcome data by heart failure with preserved ejection fraction phenotype. Data presented as mean \pm standard deviation.

Phenotype	Pertinent characteristics
A (younger, obese men)	<ul style="list-style-type: none"> ● Age 61.6 ± 5.4 years ● BMI 35.0 ± 7.7 ● 100% male ● Large improvement in KCCQ from baseline to 12 months ($+12.4 \pm 20.6$) ● KCCQ is mildly lower than expected for prognosis: <ul style="list-style-type: none"> ○ Moderate KCCQ score (mean 59.3 ± 24.6) ○ Good overall prognosis [33 events in 20% of patients; HR 1.0 (reference) based on TOPCAT primary outcome]
B (younger, active women)	<ul style="list-style-type: none"> ● Mean age 60.7 ± 6.0 ● 86% female ● Phenotype with the highest mean METs/week (12.1 ± 24.6) ● Large improvement in KCCQ from baseline to 12 months ($+13.9 \pm 23.9$) ● KCCQ is much lower than expected for prognosis: <ul style="list-style-type: none"> ○ Poor baseline KCCQ with mean 53.1 ± 23.2 ○ Excellent overall prognosis [25 events in 17% of patients; HR 0.85 (0.51–1.45) (Phenotype A reference) based on TOPCAT primary outcome]
C (metabolic syndrome)	<ul style="list-style-type: none"> ● Obesity (89% have BMI > 30) ● Hypertension (96%) ● Hyperlipidaemia (91%) ● 66% male ● Compared with placebo, increased potassium, and creatinine response to spironolactone but inconsistent BP response to spironolactone ● KCCQ and prognosis are concordantly poor: <ul style="list-style-type: none"> ○ Poor baseline KCCQ (mean 53.0 ± 23.5) ○ Poorest prognosis of the phenotypes [215 events in 40% of patients; HR 2.20 (1.52–3.17) (Phenotype A reference) based on TOPCAT primary outcome]
D (older, diabetic women with VHD)	<ul style="list-style-type: none"> ● Mean age 72.3 ± 7.0 ● 100% female ● Diabetes (42%) ● Valvular heart disease (24%) ● Consistently elevated creatinine and potassium and lower blood pressure in the spironolactone vs. placebo arm throughout the study ● KCCQ is mildly lower than expected for prognosis: <ul style="list-style-type: none"> ○ Moderate KCCQ score (mean 58.2 ± 21.1) ○ Good prognosis [64 events in 21% of patients; HR 0.97 (0.64–1.48) (Phenotype A reference) based on TOPCAT primary outcome]
E (older male alcohol users)	<ul style="list-style-type: none"> ● Mean age 75.6 ± 6.3 ● 100% male ● Phenotype with the highest alcohol use (43%) ● KCCQ is mildly higher than expected for prognosis: <ul style="list-style-type: none"> ○ Phenotype with the highest KCCQ score (mean 68.2 ± 22.8) and good prognosis [75 events in 25% of patients; HR 1.26 (0.84–1.90) (Phenotype A reference) based on TOPCAT primary outcome]
F (frail older women)	<ul style="list-style-type: none"> ● Phenotype with the oldest mean age (82.7 ± 5.6) ● 68% female ● Phenotype with the lowest mean BMI (28.5 ± 6.7) and hemoglobin (12.3 ± 1.5) ● KCCQ and prognosis are concordant and moderately reduced: <ul style="list-style-type: none"> ○ Baseline KCCQ mean 59.2 ± 22.3 ○ 110 events, 34% of patients; HR 1.83 (1.24–2.79) (Phenotype A reference) based on TOPCAT primary outcome

BMI, body mass index; BP, blood pressure; HR, hazard ratio; KCCQ, Kansas City Cardiomyopathy Questionnaire; METs, metabolic equivalents; TOPCAT, Treatment of Preserved Cardiac Function Heart Failure with an Aldosterone Antagonist Trial; VHD, valvular heart disease.

Overall KCCQ score and all KCCQ subscores with the exception of the symptom stability score varied significantly by phenotype. Phenotypes B (younger, active women) and C (metabolic syndrome) had the worst baseline health status (overall KCCQ score 53.1 ± 23.2 and 53.0 ± 23.5 , respectively), whereas Phenotype E (older male alcohol users) had the best baseline health status (overall KCCQ score 68.1 ± 22.8). This pattern was consistent through all KCCQ subscores (Figure S2 and Table S1).

Biological response to spironolactone by HFpEF phenotype

Over 2 years of follow-up, potassium levels were consistently higher in the spironolactone vs. placebo group in all six phenotypes (Figure 1A). However, only Phenotypes C (metabolic syndrome), D (older, diabetic women with VHD), and F (frail, older women) had consistently higher creatinine in the spironolactone vs. placebo group over time

A Change in potassium by phenotype and treatment arm

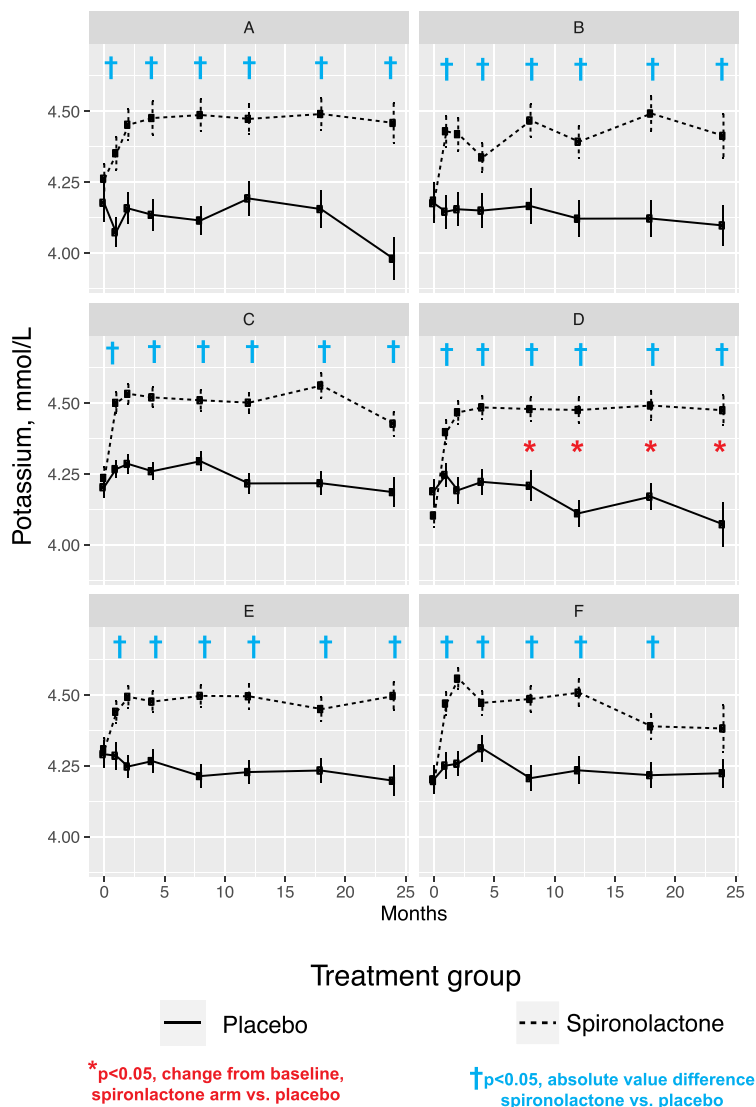


Figure 1 Changes in (A) serum potassium, (B) serum creatinine, and (C) systolic blood pressure (SBP) by HFpEF phenotype and treatment arm.

(Figure 1B). Blood pressure was consistently lower in the spironolactone arm in Phenotypes B (younger, active women) and D (older, diabetic women with VHD) only (Figure 1C).

Mortality and hospitalization outcomes by HFpEF phenotype and treatment arm

Similar to the original derivation and validation cohorts,⁶ the HFpEF phenotypes in TOPCAT differed significantly in risk of the TOPCAT primary outcome and the I-PRESERVE primary outcome (all-cause mortality or CV hospitalization) (Figure

2A, Table S5). Overall, there was a significant reduction in the TOPCAT primary outcome [hazard ratio (HR) 0.82, $P = 0.025$] in the spironolactone vs. placebo arm and a nonsignificant trend toward reduction in the I-PRESERVE primary outcome (HR 0.86, $P = 0.056$). There were no significant reductions in either TOPCAT or I-PRESERVE primary outcomes associated with spironolactone when stratified by HFpEF phenotype (data not shown). Similar results were seen for all-cause mortality (data not shown).

All-cause hospitalization, CV hospitalization, HF hospitalization, and non-CV hospitalization all differed significantly by HFpEF phenotypes (Figure 2B; $P < 0.0001$ for all four hospitalization outcomes). Phenotype B (younger, active females

B
Change in creatinine by phenotype and treatment arm

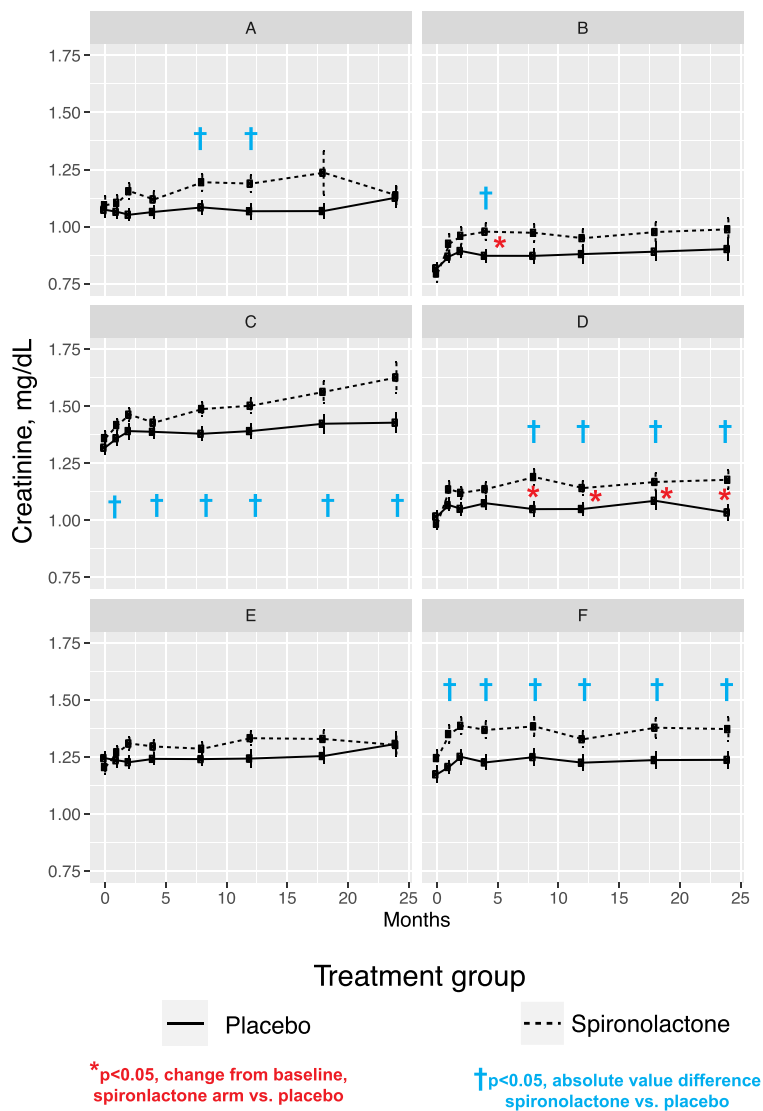


Figure 1 Continued

patients) had the lowest rate of all four hospitalization outcomes (all-cause 45.2%, CV 30.1%, HF 11.0%, and non-CV 28.8%), whereas Phenotype C (metabolic syndrome) had the highest rate of all-cause hospitalization (68.8%), CV (48.7%), and HF hospitalizations (32.7%), and Phenotype F (frail, older women) had the highest rate of non-CV hospitalization (49.2%). The spironolactone arm had significantly fewer HF hospitalizations than the placebo arm (HR 0.81, $P = 0.04$) overall. No other hospitalization outcomes differed by treatment group. Hospitalization outcomes were not

significantly different between spironolactone vs. placebo within each HFpEF phenotype.

Models adjusted for the 11 clinical variables used to construct the HFpEF phenotypes were still significantly predictive of the TOPCAT primary outcome (Table S3). The clinical variables that were strongly associated with TOPCAT primary outcome in these adjusted models were sex, diabetes, haemoglobin, and renal function. When models of the clinical variables used to construct the phenotypes were performed within each phenotype, the models for Phenotypes C

Figure 1.

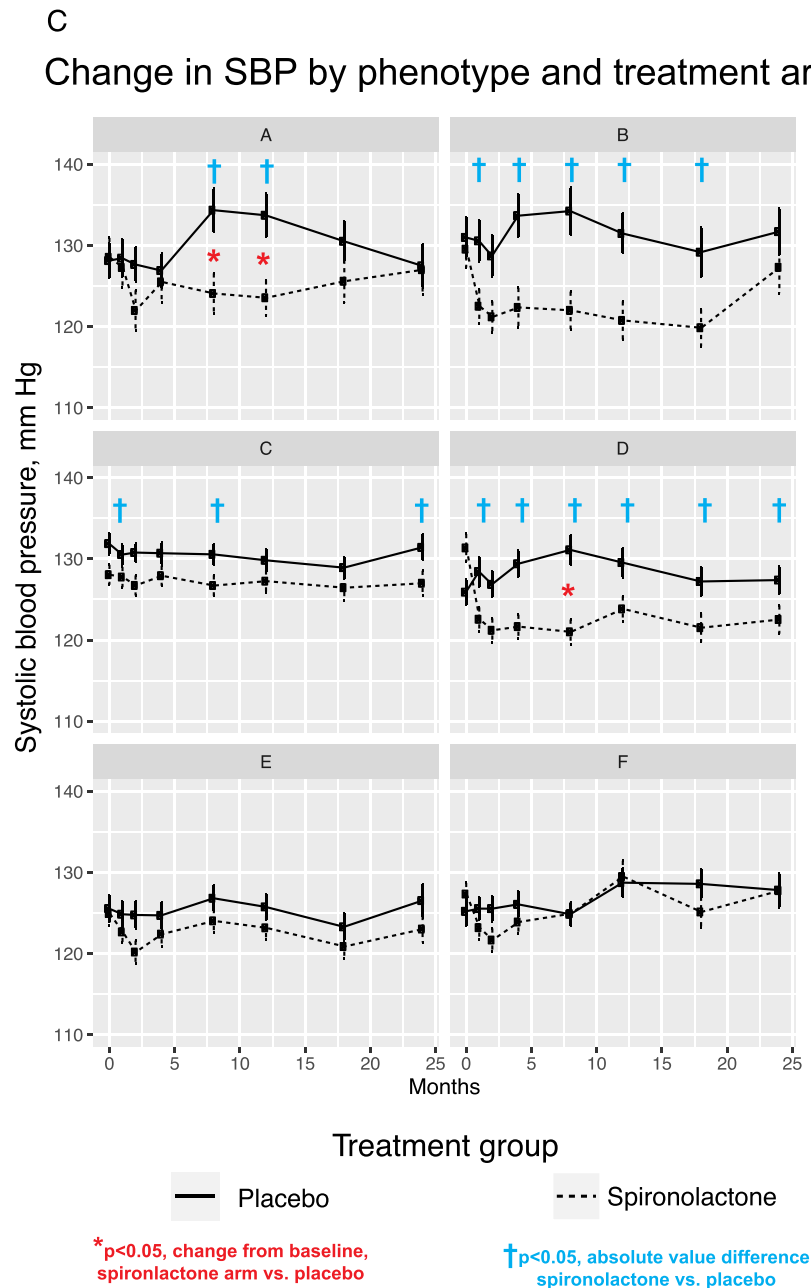


Figure 1 Continued

(metabolic syndrome phenotype) and F (frail older women phenotype) were statistically significant. In general, the most predictive variables across phenotypes were age (Phenotypes B—younger active women and F), diabetes (Phenotypes B and D—older, diabetic women with valvular heart disease), sex (Phenotype C), haemoglobin (Phenotype C, E—older male alcohol users and F), renal function (Phenotype C), coronary

artery disease (Phenotype D), and valvular heart disease (Phenotype F).

HFpEF phenotypes and health status

All phenotypes had a significant improvement in KCCQ overall summary score at 4 and 12 months compared with

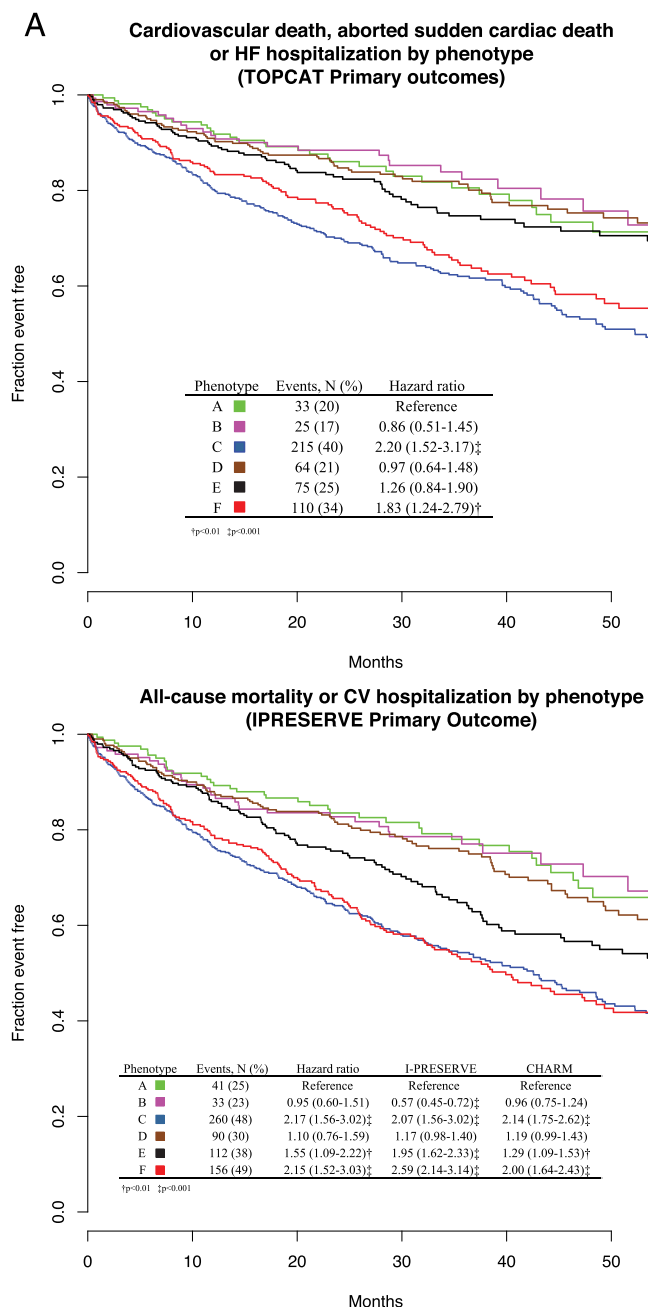


Figure 2 A. Association of HFpEF phenotypes with time to Treatment of Preserved Cardiac Function Heart Failure with an Aldosterone Antagonist Trial (TOPCAT) and I-PRESERVE primary outcomes. All hazard ratios (HR) are calculated with Phenotype A chosen as the reference (i.e. HR 1.0). B. Association of the HFpEF phenotypes in the Treatment of Preserved Cardiac Function Heart Failure with an Aldosterone Antagonist Trial (TOPCAT) with time to all-cause, CV, HF, and non-CV hospitalization. All hazard ratios (HR) are calculated with Phenotype A chosen as the reference (i.e. HR 1.0).

baseline (Figure S4A). This was driven entirely by change from baseline to 4 months, as there were no significant changes in KCCQ score from 4 to 12 months in any phenotype. The phenotype with the most improvement in KCCQ over 4 months was B (younger active women) (KCCQ increase of 11.5 ± 18.9 , $P < 0.001$). In general, Phenotypes

A (younger obese men), B (younger active women), and E (older male alcohol users) had better absolute KCCQ scores over time compared with Phenotypes C (metabolic syndrome), D (older, diabetic women with VHD), and F (frail, older women). This was primarily driven by improvements in the quality of life and social limitation scores in

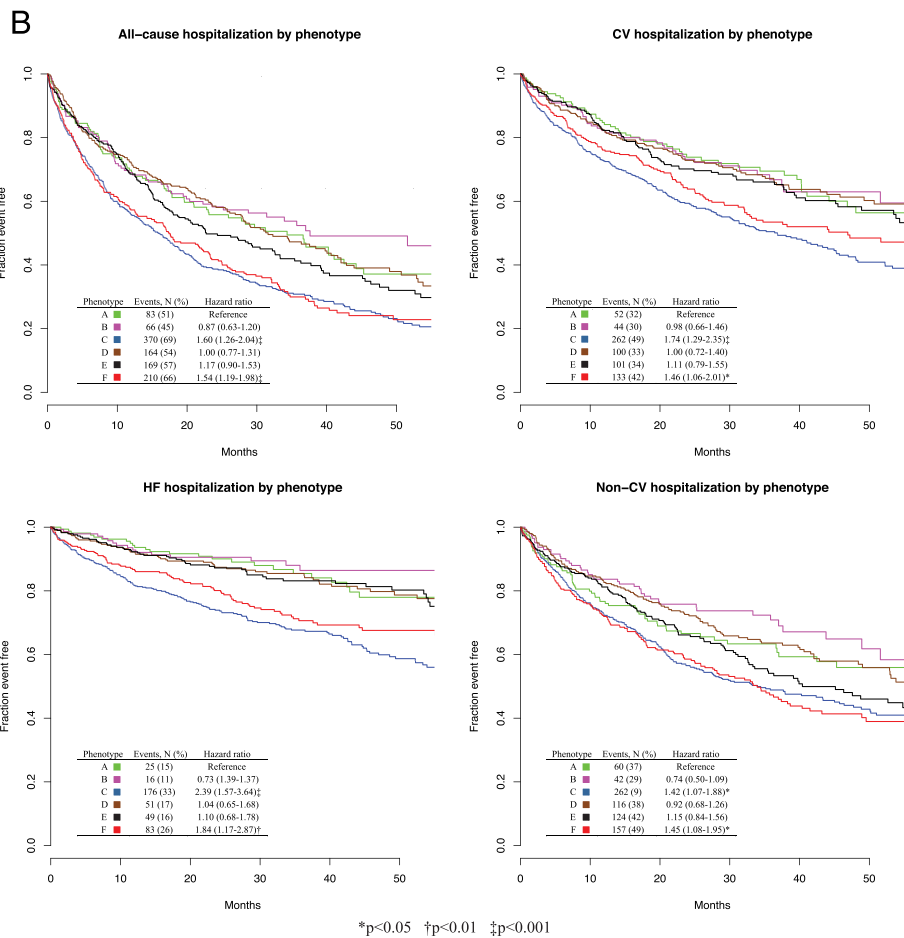


Figure 2.

Phenotypes A (younger, obese men), B (younger active women), and E (older male alcohol users) and worsening of the symptom stability score in Phenotypes C (metabolic syndrome), D (older, diabetic women with VHD), and F (frail, older women) (Figure S4B). Trends in self-efficacy score were similar among all phenotypes.

Within each HFpEF phenotype, all patients who met the primary outcome had numerically lower baseline mean KCCQ (Figure 3). Baseline KCCQ was significantly different in patients with and without the primary outcome in Phenotypes C–F. Lower baseline KCCQ scores were associated with higher univariate risk of the primary outcome in Phenotypes B (younger active women), C (metabolic syndrome), E (older male alcohol users), and F (frail, older women) (Table S2; Figure 3). Change in KCCQ from baseline to 4 months was not associated with the primary outcome overall or in any of the HFpEF phenotypes (data not shown).

Treatment arm and health status among HFpEF phenotypes

Overall the spironolactone arm had significant improvement in KCCQ at 4 months compared with placebo (8.4 ± 19.1 vs. 6.1 ± 19.1 , respectively, $P = 0.03$) but not at 12 months (7.8 ± 20.9 vs. 5.7 ± 20.9 , respectively, $P = 0.065$). All phenotypes had a significant improvement in overall KCCQ at 4 months in both spironolactone and placebo arms (Table 3). In Phenotype A (younger, obese men), there was a significant difference in improvement in KCCQ score from baseline associated with spironolactone at 4 months vs. placebo ($+14.4 \pm 19.0$ vs. $+6.8 \pm 19.7$, respectively, $P = 0.02$) and a nonsignificant trend at 12 months (12.1 ± 20.6 vs. 5.4 ± 23.0 for spironolactone vs. placebo, respectively, $P = 0.07$). There were no other significant differences in KCCQ change associated with spironolactone in other phenotypes.

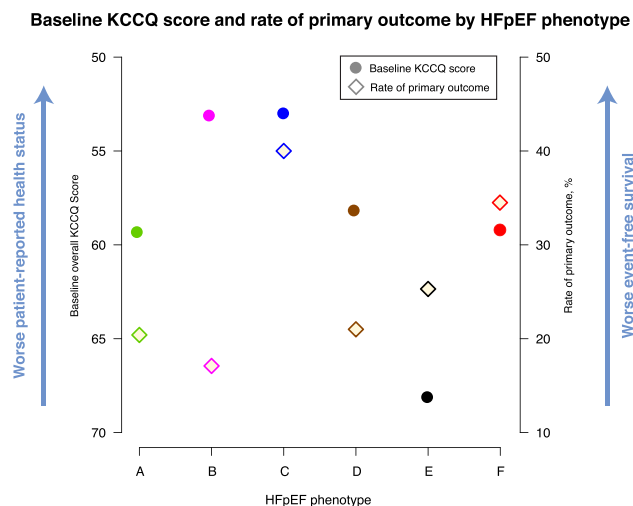


Figure 3 Baseline Kansas City Cardiomyopathy Questionnaire (KCCQ) score by heart failure with preserved ejection fraction (HFpEF) phenotype and primary outcome.

Discussion

The complex HFpEF phenotypes described here provide a data-driven framework for conceptualizing the clinical heterogeneity among patients diagnosed with HFpEF.⁶ This study confirms in a modern HFpEF clinical trial population the relative risks of adverse clinical outcomes observed in the original derivation (I-PRESERVE) and validation (CHARM-Preserved) studies. TOPCAT, I-PRESERVE, and CHARM-Preserved represent the three largest HFpEF clinical trials to date; therefore, the reproducibility of the HFpEF phenotypes across these

three clinical trials suggests that they represent truly distinct populations among patients with HFpEF. The current study extends prior observations to include differences in baseline health status and possible phenotype-specific responses to spironolactone based on physiologic markers of aldosterone blockade and clinical endpoints. Although our study did not detect a phenotype-specific treatment effect of spironolactone for the TOPCAT primary outcome, the findings that some but not all phenotypes had consistent creatinine, blood pressure, and health status responses to spironolactone may help guide future clinical trials designed to target interventions to specific HFpEF phenotypes.

This study also characterizes the baseline and longitudinal health status patterns in each HFpEF phenotype by treatment group, and the association of health status within each phenotype with the TOPCAT primary outcome. The prognostic importance of KCCQ score within each phenotype was comparable in that lower baseline KCCQ score was associated with increased risk of the adverse clinical outcomes among most phenotypes. However, differences in baseline health status between phenotypes were not necessarily concordant with relative differences in prognosis; patients in phenotypes associated with better rate of the TOPCAT primary outcome reported worse baseline health status and vice versa (Table 2, Figure 3). For example, Phenotype E (older male alcohol users) had the highest baseline KCCQ score [68.1 ± 22.8 vs. lowest in phenotype C (metabolic syndrome) 53.0 ± 23.5] yet had the third highest rates of the primary outcome, all-cause hospitalization, and CV hospitalization, and the second highest rate of non-CV hospitalization. Furthermore, Phenotypes B (younger, active women) and C (metabolic syndrome)

Table 3 Change in Kansas City Cardiomyopathy Questionnaire score within each phenotype by treatment arm

	Change baseline to 4 months	Change baseline to 12 months	Change 4 to 12 months
Phenotype A			
Spironolactone	14.4 ± 19.0*†	12.4 ± 20.6*	-1.9 ± 13.5
Placebo	6.8 ± 19.7**	5.4 ± 23.0	-2.7 ± 24.9
Phenotype B			
Spironolactone	13.3 ± 19.8*	13.9 ± 23.9*	-0.2 ± 15.4
Placebo	9.5 ± 17.7*	10.3 ± 19.8*	1.2 ± 15.3
Phenotype C			
Spironolactone	9.6 ± 19.8*	7.2 ± 19.5*	-1.0 ± 19.5
Placebo	7.1 ± 19.8*	6.2 ± 22.4*	-1.5 ± 18.5
Phenotype D			
Spironolactone	6.2 ± 18.6*	8.9 ± 20.1*	2.0 ± 19.7
Placebo	5.4 ± 18.1**	6.2 ± 18.2*	1.2 ± 18.7
Phenotype E			
Spironolactone	4.6 ± 19.0**	4.0 ± 23.2**	-0.8 ± 18.9
Placebo	4.3 ± 17.8**	4.1 ± 21.3***	-1.2 ± 17.0
Phenotype F			
Spironolactone	6.9 ± 17.3*	6.5 ± 19.2**	-0.4 ± -17.7
Placebo	4.7 ± 20.0***	3.9 ± 19.7***	-1.2 ± 16.4

Change between time points:

*P < 0.001.

**P < 0.01.

***P < 0.05.

Spironolactone vs. placebo:

†P < 0.05.

had similarly low baseline KCCQ (53.1 ± 23.2 vs. 53.0 ± 23.5) even though Phenotype B (younger, active women) had the lowest and Phenotype C (metabolic syndrome) had among the top 2 highest rates of the primary outcome (17.1% vs. 40.0%), all-cause hospitalization (45.2% vs. 68.8%), CV hospitalization (30.1% vs. 48.7%), HF hospitalization (11.0% vs. 32.7%), and non-CV hospitalization (28.8% vs. 48.7%). If these findings are confirmed in larger, 'real-world' data sets powered for multiple comparisons—such as pooled data from multiple clinical trials and registries—then they may suggest strategies for phenotype-specific treatment plans. For example, Phenotype B may benefit more from interventions aimed at improving health status, whereas in Phenotype E (older male alcohol users), preventing adverse clinical events may be more important than improving health status. Similarly, goals of therapy for Phenotype C (metabolic syndrome) may be to improve both health status and clinical outcomes, although effective treatment modalities for different types of outcomes may vary based on underlying physiology.

Differences in the patient characteristics and evidence of response to spironolactone vs. placebo summarized in *Table 2* may be exploited in future studies of patients with HFpEF. For example, Phenotype D (older, diabetic women with valvular heart disease) had the most robust biologic response to spironolactone, with consistently elevated potassium and creatinine and lower blood pressure in the intervention group compared with placebo. Therefore, spironolactone may be particularly effective in improving outcomes among Phenotype D, but given their older age and potential susceptibility to adverse events, creatinine and potassium levels should be carefully monitored. Similarly, Phenotype B (younger, active women) patients also showed a consistent biologic response to spironolactone in the treatment arm over time as evidenced by potassium and blood pressure trends over the 2 years of follow-up. Phenotype B patients also had a high prevalence of obesity, despite the most physically active phenotype with a lower prevalence of other comorbid illnesses. These characteristics may have contributed to the low event rate in Phenotype B. On the other hand, Phenotype F (frail older women) had consistently elevated potassium and creatinine over time but did not have lower blood pressure over time in the spironolactone arm. This, along with the high incidence of non-CV hospitalizations in this phenotype suggests that the complexities of ageing, such as frailty and multimorbidity, may make this group highly susceptible to the adverse effects of spironolactone such as elevated creatinine and potassium without conferring much benefit. A better approach to Phenotype F may include comprehensive geriatric assessment and targeted interventions based on the results of such an assessment. Multimorbidity may also play a role in Phenotype C (metabolic syndrome), as these patients also had a consistent potassium and creatinine response but not blood pressure response to spironolactone

in the treatment arm over time. However, the prevalence of obesity, hypertension, hyperlipidaemia, and diabetes in this group suggests that addressing the underlying lifestyle habits in Phenotype C patients may be more effective than aldosterone antagonism.

Strengths and limitations

Strengths of the current study include the completeness of data collection in the TOPCAT. Health status is notoriously difficult to collect in the observational setting; therefore, clinical trials are ideal tools for secondary analyses of health status outcomes. Another strength of the current study is the use of a novel, data-driven approach to categorize HFpEF patients into phenotypes that have already been validated in other large and well-described clinical trial populations.

Our study also has several limitations. First, the ejection fraction used to define HFpEF is now $>50\%$ according to recent guidelines,^{21,22} making the TOPCAT inclusion criterion of $EF > 45\%$ out of date. Second, we used BMI as a measure of body composition in defining the HFpEF phenotypes. Other measures of body composition are likely more informative in older adults with HFpEF but must be balanced against the ease of obtaining height and weight for measuring BMI in routine clinical care. Furthermore, BMI and waist circumference were correlated in this TOPCAT population ($r = 0.78$), suggesting that BMI still adds clinically important information to defining the HFpEF phenotypes. Third, we used only data from the Americas because of the previously described problems with the nature of the patients enrolled in Russia and Georgia, who had event rates far below what is expected of a typical HFpEF population.¹² This precluded comparison of phenotype-specific outcomes in the placebo group alone because of small sample size. Additionally, the phenotype-specific analyses, particular those stratified by treatment group, were underpowered to detect modest differences in some outcomes. Among those that were significant, we cannot exclude a type I statistical error because of multiple comparisons. We also did not detect a statistically significant difference in the rate of the primary outcome by treatment arm and HFpEF phenotype, which again may be due to small sample sizes and consequent lack of statistical power. Accordingly, the results of this study should be considered hypothesis generating.

Future directions

First, future research should validate the HFpEF phenotypes in the 'real-world' setting. This may be achieved by utilizing existing heart failure registry data (although challenges exist in finding patients with an accurate diagnosis of heart failure

and confirmed ejection fraction) and/or utilizing electronic medical record data.

Second, the findings from the current study should be confirmed in larger data sets powered for multiple comparisons, such as pooled data from multiple clinical trials and registries. Such a pooled data set may help (i) confirm our findings of phenotype-specific clinical and health status outcomes and (ii) confirm and further explore the discordant clinical and health status outcomes in Phenotypes B (younger, active women) and E (older male alcohol users).

Finally, testing novel interventions in the HFpEF population should take these phenotypes into consideration when targeting therapeutic interventions to patients most likely to benefit.

Conclusions

Rates of mortality and hospitalization associated with previously described data-driven HFpEF phenotypes were recapitulated in TOPCAT. In addition to differences in underlying pathophysiology, HFpEF phenotypes have important differences in mortality and hospitalization outcomes, biological response to spironolactone as well as health status outcomes. These insights should help guide future studies in patients with HFpEF by targeting outcomes and treatment modalities most relevant to specific HFpEF phenotypes.

Acknowledgements

The authors would like to thank JoAnn Lindenfeld, MD, for her feedback on early versions of this manuscript.

Funding

Dr. Kao was supported by National Institutes of Health, National Heart, Lung and Blood Institute Grant Number

K08HL125725 and by the American Heart Association grant number 71G3366030.

Conflict of interest

None declared.

Supporting information

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

Figure S1. Probability of HFpEF phenotype membership by HFpEF Phenotype.

Figure S2. Baseline KCCQ subscores by HFpEF phenotype.

Figure S3. All-cause mortality by phenotype ($P < 0.001$).

Figure S4. (A) Overall KCCQ score over time by phenotype.

Figure S4. (B) KCCQ subscores over time by phenotype.

Table S1. KCCQ scores by HFpEF phenotype.

Table S2. Association between baseline KCCQ and outcomes by phenotype, hazard ratio/10 points (95% confidence interval).

Table S3. Cox Proportional Hazards Model for the TOPCAT Primary Outcome using HFpEF Phenotypes Adjusted for Component Variables. Chi-square p-value for both models < 0.0001 .

Table S4. HR for TOPCAT Primary Outcome by Phenotype, adjusted for clinical variables used to construct the Phenotypes. Sex variable was not included for Phenotypes consisting of only one sex.

Table S5. Comparison of primary and secondary outcomes in the TOPCAT, I-PRESERVE and CHARM-Preserved trials and the association of the HFpEF phenotypes with the primary and secondary outcomes of each trial.

References

1. Nanayakkara S, Kaye DM. Management of heart failure with preserved ejection fraction: a review. *Clin Ther* 2015; **37**: 2186–2198.
2. Steinberg BA, Zhao X, Heidenreich PA, Peterson ED, Bhatt DL, Cannon CP, Hernandez AF, Fonarow GC. Get With the Guidelines Scientific Advisory C, Investigators. Trends in patients hospitalized with heart failure and preserved left ventricular ejection fraction: prevalence, therapies, and outcomes. *Circulation* 2012; **126**: 65–75.
3. Paulus WJ, Tschope C. A novel paradigm for heart failure with preserved ejection fraction: comorbidities drive myocardial dysfunction and remodeling through coronary microvascular endothelial inflammation. *J Am Coll Cardiol* 2013; **62**: 263–271.
4. Triposkiadis F, Giamouzis G, Parisis J, Starling RC, Boudoulas H, Skoularigis J, Butler J, Filippatos G. Reframing the association and significance of comorbidities in heart failure. *Eur J Heart Fail* 2016; **18**: 744–758.
5. Shah SJ, Kitzman DW, Borlaug BA, van Heerebeek L, Zile MR, Kass DA, Paulus WJ. Phenotype-specific treatment of heart failure with preserved ejection fraction: a multiorgan roadmap. *Circulation* 2016; **134**: 73–90.
6. Kao DP, Lewsey JD, Anand IS, Massie BM, Zile MR, Carson PE, McKelvie RS,

- Komajda M, McMurray JJ, Lindenfeld J. Characterization of subgroups of heart failure patients with preserved ejection fraction with possible implications for prognosis and treatment response. *Eur J Heart Fail* 2015; **17**: 925–935.
7. Yusuf S, Pfeffer MA, Swedberg K, Granger CB, Held P, McMurray JJV, Michelson EL, Olofsson B, Östergren J. Effects of candesartan in patients with chronic heart failure and preserved left-ventricular ejection fraction: the CHARM-Preserved Trial. *Lancet* 2003; **362**: 777–781.
 8. Massie BM, Carson PE, McMurray JJ, Komajda M, McKelvie R, Zile MR, Anderson S, Donovan M, Iverson E, Staiger C, Ptaszynska A, Investigators IP. Irbesartan in patients with heart failure and preserved ejection fraction. *N Engl J Med* 2008; **359**: 2456–2467.
 9. Pitt B, Pfeffer MA, Assmann SF, Boineau R, Anand IS, Claggett B, Clausell N, Desai AS, Diaz R, Fleg JL, Gordeev I, Harty B, Heitner JF, Kenwood CT, Lewis EF, O'Meara E, Probstfield JL, Shaburishvili T, Shah SJ, Solomon SD, Sweitzer NK, Yang S, McKinlay SM, Investigators T. Spironolactone for heart failure with preserved ejection fraction. *N Engl J Med* 2014; **370**: 1383–1392.
 10. Pfeffer MA, Claggett B, Assmann SF, Boineau R, Anand IS, Clausell N, Desai AS, Diaz R, Fleg JL, Gordeev I, Heitner JF, Lewis EF, O'Meara E, Rouleau JL, Probstfield JL, Shaburishvili T, Shah SJ, Solomon SD, Sweitzer NK, McKinlay SM, Pitt B. Regional variation in patients and outcomes in the Treatment of Preserved Cardiac Function Heart Failure With an Aldosterone Antagonist (TOPCAT) trial. *Circulation* 2015; **131**: 34–42.
 11. *Biologic Specimens and Data Repository Information Coordinating Center (BioLINCC)* <https://biolincc.nhlbi.nih.gov/home/>.
 12. Bristow MR, Enciso JS, Gersh BJ, Grady C, Rice MM, Singh S, Sopko G, Boineau R, Rosenberg Y, Greenberg BH. Detection and management of geographic disparities in the TOPCAT trial: lessons learned and derivative recommendations. *JACC Basic Transl Sci* 2016; **1**: 180–189.
 13. de Denus S, O'Meara E, Desai AS, Claggett B, Lewis EF, Leclair G, Jutras M, Lavoie J, Solomon SD, Pitt B, Pfeffer MA, Rouleau JL. Spironolactone metabolites in TOPCAT—new insights into regional variation. *N Engl J Med* 2017; **376**: 1690–1692.
 14. Desai AS, Claggett B, Lewis EF, Leclair G, Jutras M, Lavoie J, Solomon SD, Pitt B, Pfeffer MA, Rouleau JL. Spironolactone metabolites in TOPCAT—new insights into regional variation. *N Engl J Med* 2017; **367**: 1690–1692.
 15. Hegde SM, Claggett B, Shah AM, Lewis EF, Anand I, Shah SJ, Sweitzer NK, Fang JC, Pitt B, Pfeffer MA, Solomon SD. Physical activity and prognosis in the TOPCAT trial (Treatment of Preserved Cardiac Function Heart Failure With an Aldosterone Antagonist). *Circulation* 2017; **136**: 982–992.
 16. Lewis EF, Claggett B, Shah AM, Liu J, Shah SJ, Anand I, O'Meara E, Sweitzer NK, Rouleau JL, Fang JC, Desai AS, Retta TM, Solomon SD, Heitner JF, Stamos TD, Boineau R, Pitt B, Pfeffer MA. Racial differences in characteristics and outcomes of patients with heart failure and preserved ejection fraction in the Treatment of Preserved Cardiac Function Heart Failure trial. *Circ Heart Fail* 2018; **11**: e004457.
 17. Green CP, Porter CB, Bresnahan DR, Spertus JA. Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure. *J Am Coll Cardiol* 2000; **35**: 1245–1255.
 18. Joseph SM, Novak E, Arnold SV, Jones PG, Khattak H, Platts AE, Davila-Roman VG, Mann DL, Spertus JA. Comparable performance of the Kansas City Cardiomyopathy Questionnaire in patients with heart failure with preserved and reduced ejection fraction. *Circ Heart Fail* 2013; **6**: 1139–1146.
 19. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF 3rd, Feldman HI, Kusek JW, Eggers P, Van Lente F, Greene T, Coresh J, Ckd EPI. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009; **150**: 604–612.
 20. National Kidney F. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis* 2002; **39**: S1–S266.
 21. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr, Drazner MH, Fonarow GC, Geraci SA, Horwich T, Januzzi JL, Johnson MR, Kasper EK, Levy WC, Masoudi FA, McBride PE, McMurray JJ, Mitchell JE, Peterson PN, Riegel B, Sam F, Stevenson LW, Tang WH, Tsai EJ, Wilkoff BL, American College of Cardiology F, American Heart Association Task Force on Practice G. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol* 2013; **62**: e147–e239.
 22. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJ, Falk V, Gonzalez-Juanatey JR, Harjola VP, Jankowska EA, Jessup M, Linde C, Nihoyannopoulos P, Parissis JT, Pieske B, Riley JP, Rosano GM, Ruilope LM, Ruschitzka F, Rutten FH, van der Meer P, Authors/Task Force M. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J* 2016; **37**: 2129–2200.