Minimal Clinically Important Differences in 6-Minute Walk Test in Patients With HFrEF and Iron Deficiency

MUHAMMAD SHAHZEB KHAN, MD, MSC,[1] STEFAN D. ANKER, MD, PhD,[2] TIM FRIEDE, PhD,[3] EWA A. JANKOWSKA, MD,[4] MARCO METRA, MD,[5] ILEANA L PIÑA, MD, MPh,[6] ANDREW JS COATS, MD, PhD,[7] GIUSEPPE ROSANO, MD, PhD,[8] BERNARD ROUBERT, MSC,[9] UDO-MICHAEL GOEHRING, MD,[9] FABIO DORIGOTTI, MD,[9] JOSEP COMIN-COLET, MD, PhD,[10] DIRK J VAN VELDHUISEN, MD,[11] GERASIMOS S. FILIPPATOS, MD,[12,13] PIOTR PONIKOWSKI, MD,[4] AND JAVED BUTLER, MD, MPH, MBA[14,15]

Durham, North Carolina, Dallas, Texas, and Philadelphia, Pennsylvania, USA; Berlin and Göttingen, Germany; Wroclaw, Poland; Brescia and Rome, Italy; Sydney, Australia; Glattbrugg, Switzerland; Barcelona, Spain; Gronigen, The Netherlands; Nicosia, Cyprus; and Athens, Greece

ABSTRACT

Background: The 6-minute walk test (6MWT) is widely used to measure exercise capacity; however, the magnitude of change that is clinically meaningful for individuals is not well established in heart failure with reduced ejection fraction (HFrEF).

Objective: To calculate the minimal clinically important difference (MCID) for change in exercise capacity in the 6MWT in iron-deficient populations with HFrEF.

Methods: In this pooled secondary analysis of the FAIR-HF and CONFIRM-HF trials, mean changes in the 6MWT from baseline to weeks 12 and 24 were calculated and calibrated against the Patient Global Assessment (PGA) tool (clinical anchor) to derive MCIDs in improvement and deterioration.

Results: Of 760 patients included in the 2 trials, 6MWT and PGA data were available for 680 (89%) and 656 (86%) patients at weeks 12 and 24, respectively. The mean 6MWT distance at baseline was 281 $\pm$ 103 meters. There was a modest correlation between changes in 6MWT and PGA from baseline to week 12 ($r = 0.31$; $P < 0.0001$) and week 24 ($r = 0.43$; $P < 0.0001$). Respective estimates (95% confidence intervals) of MCID in 6MWT at weeks 12 and 24 were 14 meters (5;23) and 15 meters (3;27) for a "little improvement" (vs no change), 20 meters (10;30) and 24 meters (12;36) for moderate improvement vs a "little improvement,", -11 meters (-32;9.2) and -31 meters (-53;-8) for a "little deterioration" (vs no change), and -84 meters (-144;-24) and -69 meters (-118;-20) for "moderate deterioration" vs a "little deterioration".

Conclusions: The MCID for improvement in exercise capacity in the 6MWT was 14 meters−15 meters in patients with HFrEF and iron deficiency. These MCIDs can aid clinical interpretation of study data. (*J Cardiac Fail 2023;29:760−770*)

Key Words: Heart failure with reduced ejection fraction, minimal clinically important difference, 6-minute walk test.

From the [1]Division of Cardiology, Duke University Medical Center, Durham, North Carolina, USA; [2]Department of Cardiology and Berlin Institute of Health Center for Regenerative Therapies, German Center for Cardiovascular Research Partner Site Berlin, Charité Universitätsmedizin Berlin, Berlin, Germany; [3]Department of Medical Statistics, University Medical Center Göttingen, Göttingen, Germany; German Center for Cardiovascular Research, Göttingen, Germany; [4]Institute of Heart Diseases, Wroclaw Medical University Poland and Institute of Heart Diseases, University Hospital, Wroclaw, Poland; [5]Cardiology, ASST Spedali Civili and Department of Medical and Surgical Specialties, Radiological Sciences, and Public Health, Civil Hospitals and University of Brescia, Brescia, Italy; [6]Kimmel College of Medicine, Thomas Jefferson University, Philadelphia, Pennsylvania, USA; [7]Heart Research Institute, Sydney, Australia; [8]Centre for Clinical and Basic Research, Department of Medical Sciences, IRCCS San Raffaele Pisana, Rome, Italy; [9]Vifor Pharma, Glattbrugg, Switzerland; [10]Department of Cardiology, Bellvitge University Hospital and IDIBELL, University of Barcelona, Hospitalet de Llobregat, Barcelona, Spain; [11]Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; [12]University of Cyprus, Nicosia, Cyprus; [13]National and Kapodistrian University of Athens, Athens University Hospital Attikon, Athens, Greece; [14]Baylor Scott and White Research Institute, Dallas, Texas, USA and [15]Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA.

Manuscript received January 31, 2022; revised manuscript received September 15, 2022; revised manuscript accepted October 10, 2022.

Reprint requests: Javed Butler, MD, MPh, MBA, Baylor Scott and White Research Institute, 3434 Oak Street, Suite 501, Dallas, Texas 75204, USA. E-mail: Javed.Butler@bswhealth.org

1071-9164/$ - see front matter

https://doi.org/10.1016/j.cardfail.2022.10.423

Impaired functional capacity is common in patients with heart failure (HF).[1] With increasing focus on patient preference for better overall health status, in addition to a reduction in the use of traditional disease-specific endpoints, such as morbidity and mortality,[2,3] therapeutic targets in patients with HF have expanded to encompass improvements in functional capacity and health status (health-related quality of life).[4] The 6-minute walk test (6MWT) is an indicator of exercise capacity and prognosis in various cardiopulmonary conditions.[5–9] The 6MWT is 1 of the common instruments used to measure changes in exercise capacity in patients with HF and is increasingly included as a clinical trial endpoint in HF studies.[10,11] However, an understanding of the magnitude of change that is clinically meaningful to the patient (the minimal clinically important difference [MCID]) is fundamental to the interpretation of changes in 6MWT and decision making regarding the effectiveness of an intervention.

A wide range of 6MWT MCIDs has been previously reported across differing patient populations and studies: in patients with lung diseases, MCIDs ranging from 10–80 meters have been reported[12,13]; in patients with pulmonary arterial hypertension, the estimated MCID was approximated at 33 meters[14]; and in older adults with mobility impairments, the MCID has been estimated at 19–22 meters.[15] In patients with HF specifically, studies have suggested MCIDs ranging from 22–90 meters.[5,16] This large variation is likely to be due to the significant heterogeneity in 6MWT distances observed in patients with similar symptomatic presentations (within the same New York Heart Association [NYHA] class),[17] to differing study methods, to small sample sizes, and to proportions of patients lost to follow-up; nevertheless, the MCIDs for 6MWT changes in patients with HF have not been well established.

In the FAIR-HF (Ferinject Assessment in Patients with Iron Deficiency and Chronic Heart Failure[18]) and the CONFIRM-HF (Ferric CarboxymaltOse evaluatioN on perFormance in patients with IRon deficiency in coMbination with chronic Heart Failure[19]) studies, intravenous ferric carboxymaltose (FCM) (a nanoparticle iron-carbohydrate complex) improved mean 6MWT distance vs placebo in patients with HF with reduced ejection fraction (HFrEF) and iron deficiency. In this analysis, we used an anchor-based approach to establish 6MWT MCIDs in a large, pooled cohort of FAIR-HF and CONFIRM-HF patients, in association with FCM or placebo, using randomized clinical trial data.

Methods

The data for this analysis were drawn from 2 double-blind, placebo-controlled, parallel-group trials that evaluated the effects of intravenous FCM vs placebo in ambulatory patients with HFrEF and iron deficiency: FAIR-HF[18] and CONFIRM-HF.[19] The detailed designs and inclusion criteria for these studies have been published previously[20,21]; key study characteristics are shown in Table 1. The trials were approved by the appropriate regulatory authorities and ethics committees at each participating center, and all patients who participated in the individual randomized controlled trials provided written informed consent. The trials were conducted in strict compliance with the guidelines for Good Clinical Practice of the International Council for Harmonization and with the Declaration of Helsinki.

Assessment of Exercise Capacity Measurements

The 6MWT is a submaximal exercise test that involves measuring the distance walked over a span of 6 minutes. Participants were encouraged to walk on a straight, flat surface as fast as possible for 6 minutes on a marked course, with pauses as necessary. The maximum distance walked was recorded and used to estimate exercise capacity at baseline and weeks 4, 12 and 24 in FAIR-HF, and at baseline and weeks 6, 12, 24, 36 and 52 in CONFIRM-HF. For the purpose of this analysis, data from the placebo

Table 1. Key Characteristics of the Two Included Randomised Controlled Trials

	FAIR-HF[18,20]	CONFIRM-HF[19,21]
Randomization	2:1 (FCM:placebo)	1:1 (FCM:placebo)
Number of patients	459 (FCM: 304; placebo: 155)	301* (FCM: 150; placebo: 151)
Study duration	24 weeks	52 weeks
Patient population and HF details	Ambulatory patients with optimally treated CHF (NYHA class II/III) and iron deficiency	Ambulatory patients with optimally treated CHF (NYHA class II/III) and iron deficiency
Hemoglobin	≥9.5 and ≤13.5 g/dL	<15 g/dL
Primary endpoint	Change in PGA score and NYHA class from baseline to week 24	Change in 6MWT distance from baseline to week 24

*304 patients were randomized, but only 301 received study treatment and had any postbaseline assessment. CHF, chronic heart failure; CONFIRM-HF, Ferric CarboxymaltOse evaluatioN on perFormance in patients with IRon deficiency in coMbination with chronic Heart Failure; FAIR-HF, Ferinject Assessment in Patients with Iron Deficiency and Chronic Heart Failure; FCM, ferric carboxymaltose; HF, heart failure; NYHA, New York Heart Association; PGA, patient global assessment; 6MWT, 6-minute walk test.

and treatment arms were pooled from both studies at both the 12- and 24-week time points.

Assessment of Patients' Perceptions of Clinical Change

The Patient Global Assessment (PGA), a general health-related quality-of-life tool based on a Likert scale, was administered to evaluate the magnitude of change in patient-perceived HF status compared with the start of treatment by using the following response categories: +3: much improvement; +2: moderate improvement; +1: a little improvement; 0: no change; −1: a little deterioration; -2: moderate deterioration; and -3: much deterioration. The English version of the PGA has been translated for clinical use into 14 different languages by professional translators and double-checked by native-speaking Vifor Pharma employees. The PGA was administered at postrandomization visits before any other interview, assessment or procedure, and in this analysis, data from the 12- and 24-week time points were pooled from each study.

Assessment of Minimal Clinically Important Difference

The MCID for the 6MWT was evaluated using an anchor-based approach that calibrates exercise capacity against a clinically relevant external indicator. PGA, which has been used to assess patients' condition as directly perceived by the patients and has been used in several prior MCID calculations, was chosen as the clinical anchor against which changes in 6MWT were calibrated.[22-24] 6MWT mean change scores were calculated for each category of change in PGA to estimate the MCID. Patients with data on PGA and 6MWT at 2 or more time points were included. The assessment at week 12 was chosen as the key endpoint because it was considered the most appropriate time point to balance the minimum possible time duration required for an intervention to be impactful, considering recall ability and the risk of patient attrition and missing data. The assessment at week 24 was used to evaluate the stability of estimates over longer follow-up durations.

Statistical Analysis

Data were reported as number (%) for categorical variables and mean (standard deviation [SD]) for continuous variables. Patients were stratified by PGA category (much improvement, moderate improvement, a little improvement, no change, a little deterioration, moderate deterioration, and much deterioration), and mean (standard error [SE]) change from baseline in 6MWT was calculated for the subgroup of patients in each category. Anchor-based MCIDs with 95% confidence intervals (CIs) for the 6MWT were then calculated by subtracting the mean change from baseline in the 6MWT score in 1 PGA subgroup from that in the adjacent PGA subgroup. For example, the mean change in 6MWT for patients in the "no-change" PGA subgroup was subtracted from the mean change in 6MWT for the patients in the "little-improvement" PGA subgroup to give the MCID. MCIDs for "moderate" vs "mild" and "much" vs "moderate" improvements and corresponding deteriorations in 6MWT distance were similarly calculated. The influence of the treatment allocation on the MCID was assessed using ANCOVA (analysis of covariance) of the 6MWT change by PGA classes, treatment groups and the interaction PGA × treatment. A similar ANCOVA model was used to assess whether baseline 6MWT (divided into tertiles) had any influence on MCID.

Correlations between changes in 6MWT and PGA anchor values from baseline to various time points were calculated using a nonparametric Spearman rank correlation coefficient to assess the level of confidence in the interpretation of results. The linearity of the association between the analyzed 6MWT changes (continuous scale) and PGA at weeks 12 and 24 was assessed by analyzing residuals of the linear regression of each score by PGA.

Patients with missing baseline 6MWT values and/or who did not have ≥ 1 postbaseline 6MWT and/or PGA values prior to the time point of interest were excluded from all analyses. Patients who died or were hospitalized during the time period of interest were excluded from the main analysis, as were patients who did not have values for PGA or 6MWT at each of the specific 12- and 24-week time points. For the sensitivity analysis, patients with missing data who were known to be alive and not hospitalized were imputed by using the last observation carried forward, whereas patients who died or were hospitalized were categorized as having experienced "much deterioration" for the PGA analysis, and improvement in 6MWT for such patients was considered 0 meters for that specific time point. Two-tailed *P* values were used for all correlation and regression assessments, with $P < 0.05$ considered statistically significant. SAS version 9.4 (SAS Institute, Cary, NC, USA; 2000−2004) was used to conduct all analyses.

Results

Of the 760 patients in the pooled FAIR-HF and CONFIRM-HF population, 6MWT change from baseline data were available for 684 (90%) and 660 (87%) patients at weeks 12 and 24, respectively, while PGA data were available for 685 (90%) and 669 (88%) patients at weeks 12 and 24, respectively. Overall, 680 patients had baseline 6MWT values and

nonmissing values for both PGA and 6MWT at week 12, and 656 patients had baseline 6MWT values and nonmissing values for both PGA and 6MWT at week 24; these patients were included in the main MCID analysis. The sensitivity analysis with imputation for death, hospitalization and last observation carried forward included 738 and 739 patients at weeks 12 and 24, respectively. A flow diagram delineating patients included in main and sensitivity analyses is shown in Supplementary Fig. 1.

The mean (SD) age of the 760 patients was 68 (10) years, 51% were female, and 67% were in NYHA class III. Almost all patients (> 99%) were from a white/European ethnic background. The mean distance recorded for the 6MWT at baseline was 281 ± 103 meters. Table 2 summarizes the pooled baseline and clinical characteristics of participants enrolled in the 2 studies.

Correlation Between Change in 6MWT Distance and PGA Score

There were statistically significant but modest correlations between change in 6MWT and PGA score from baseline to week 12 ($r = 0.31$; $P < 0.0001$) and week 24 ($r = 0.43$; $P < 0.0001$) (Supplementary Fig. 2 and 3). For changes in 6MWT at weeks 12 and 24, no substantial deviations from a normal distribution were detected, and there were low correlations with PGA (expressed as deterioration and improvement).

Mean Change in 6MWT Distance Across PGA Categories

The distributions of patients across the 7 PGA categories (from much improvement to much deterioration) at weeks 12 and 24 are shown in Fig. 1, and

Table 2. Pooled Baseline and Clinical Characteristics of the Included Patients

	FCM pool (n = 454)	Placebo pool (n = 306)	Total (N = 760)
Age, years	67.8 (10.1)	68.2 (10.4)	68.0 (10.2)
Female sex	226 (49.8)	159 (52.0)	385 (50.7)
White European ethnicity	452 (99.6)	305 (99.7)	757 (99.6)
NYHA class III	321 (70.7)	186 (60.8)	507 (66.7)
LVEF, %	33.6 (6.7)	34.7 (6.9)	34.1 (6.8)
BMI, kg/m^2	28.1 (4.7)	28.6 (5.4)	28.3 (5.0)
6MWT distance, m	278.6 (102.8)	285.1 (104.2)	281.2 (103.3)
Comorbidities			
Hypertension	373 (82.2)	259 (84.6)	632 (83.2)
Diabetes mellitus	131 (28.9)	82 (26.8)	213 (28.0)
Smoking	133 (29.3)	82 (26.8)	215 (28.3)
Atrial fibrillation	493 (53.9)	431 (57.7)	924 (55.6)
Prior MI	500 (54.7)	395 (52.9)	895 (53.9)
Prior stroke	99 (10.8)	103 (13.8)	202 (12.2)
Prior coronary revascularisation	312 (34.1)	278 (37.2)	590 (35.5)
Hb, g/dL	12.1 (1.3)	12.2 (1.4)	12.1 (1.3)
Hb category			
<10 g/dL	26 (5.7)	12 (3.9)	38 (5.0)
≥10 and <12 g/dL	181 (39.9)	120 (39.2)	301 (39.6)
≥12 g/dL	247 (54.4)	174 (56.9)	421 (55.4)
Serum ferritin, ng/mL	54.0 (52.6)	58.6 (55.6)	55.9 (53.8)
Serum ferritin category			
<50 ng/mL	266 (58.6)	172 (56.2)	438 (57.6)
≥50 and <100 ng/mL	138 (30.4)	95 (31.1)	233 (30.7)
≥100 ng/mL	50 (11.0)	39 (12.8)	89 (11.7)
TSAT, %	18.5 (14.5)	17.4 (8.3)	18.1 (12.4)
TSAT category			
≤10%	94 (20.7)	61 (19.9)	155 (20.4)
>10 and ≤20%	213 (46.9)	140 (45.8)	353 (46.5)
>20%	147 (32.4)	105 (34.3)	252 (33.2)
eGFR, mL/min/1.73 m^2	64.4 (20.8)	64.2 (22.5)	64.3 (21.5)
eGFR <60 mL/min/1.73 m^2	179 (39.4)	137 (44.8)	316 (41.6)
Ischemic HF etiology	370 (81.5)	249 (81.4)	619 (81.4)
Concomitant medication			
ARNI or SGLT2 inhibitor	0 (0.0)	0 (0.0)	0 (0.0)
ACEI or ARB or ARNI	423 (93.2)	283 (92.5)	706 (92.9)
Beta-blocker	393 (86.6)	267 (87.3)	660 (86.8)
Aldosterone antagonists	237 (52.2)	147 (48.0)	384 (50.5)
Triple therapy	194 (42.7)	122 (39.9)	316 (41.6)

Values are mean (SD) for continuous variables and n (%) for categorical variables.

ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; ARNI, angiotensin receptor-neprilysin inhibitor; BMI, body mass index; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration; eGFR, estimated glomerular filtration rate; FCM, ferric carboxymaltose; Hb, hemoglobin; LVEF, left ventricular ejection fraction; MI, myocardial infarction; NYHA, New York Heart Association; SD, standard deviation; TSAT, transferrin saturation; 6MWT, 6-minute walk test.

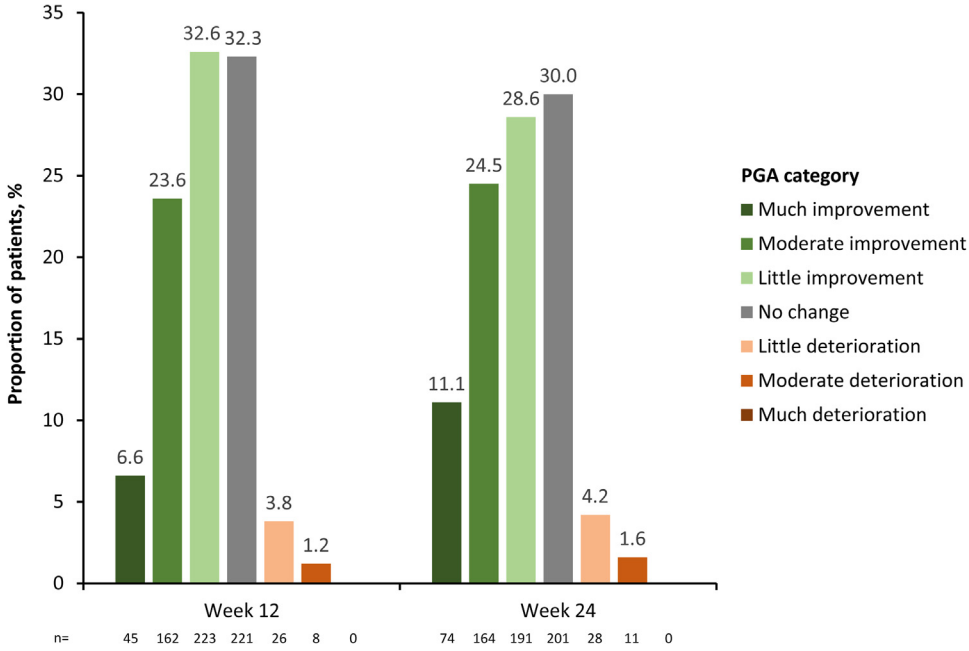

Fig. 1. Patient distribution across PGA categories at weeks 12 and 24. The figure shows the percentage of patients (N = 760) in each of the PGA categories at week 12 and week 24. The PGA categories (much improvement, moderate improvement, a little improvement, no change, a little deterioration, moderate deterioration, and much deterioration) are coded according to the color legend shown in the figure. No imputation was performed in this analysis. HfrEF, heart failure with reduced ejection fraction; PGA, patient global assessment.

the mean changes in 6MWT distance within each of these PGA categories at weeks 12 and 24 are shown in Fig. 2. At week 12, 45 patients (7%) had experienced "much improvement", 162 (24%) had experienced "moderate improvement", 223 (33%) had experienced "little improvement", and 221 (32%) had experienced "no change" in PGA score vs baseline. The mean (SE) change in 6MWT distance was 57 meters (10 meters) for those with "much improvement" in PGA score; 45 meters (4 meters)

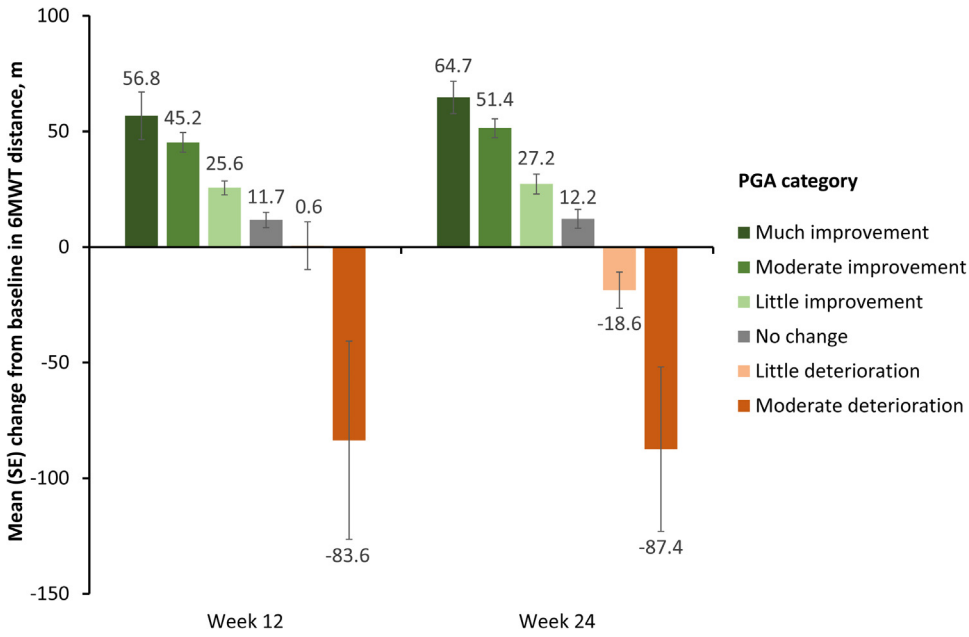

Fig. 2. Mean change in 6MWT across PGA categories at weeks 12 and 24. The figure shows the mean (SE) change in 6MWT distance from baseline to week 12 and week 24 in each PGA subgroup (much improvement, moderate improvement, a little improvement, no change, a little deterioration, moderate deterioration, and much deterioration) coded according to the color legend shown in the figure. No imputation was performed in this analysis. 6MWT, 6-minute walk test; PGA, patient global assessment; SE, standard error.

for those with "moderate improvement" in PGA score; 26 meters (3 meters) for those with a "little improvement" in PGA score; and 12 meters (3 meters) for those with "no change" in PGA score; results were similar at week 24 (Fig. 2).

At week 12, 26 patients (4%) had experienced "little deterioration" in PGA score vs baseline, and 8 patients (1%) had experienced "moderate deterioration"; no patients experienced "much deterioration". These results were similar at week 24 (Fig. 1). The mean (SE) change in 6MWT distance was 1 meter (10 meters) for those with little deterioration and -84 meter (43 meters) for those with moderate deterioration in PGA score at week 12; corresponding values at week 24 were -19 meters (8 meter) and -87 meters (36 meter), respectively (Fig. 2).

Estimates of Minimal Clinically Important Difference

The results of the 6MWT MCID main, treatment subset, and sensitivity analyses at weeks 12 and 24 are shown in Table 3. Using PGA as the clinical anchor, the MCID (95% CI) for "little improvement" (vs "no change") in 6MWT in the main analysis was 14 meters (5;23) at week 12 and 15 meters (3;27) at week 24, and the MCID for "little deterioration" (vs "no change") was -11 meters (-32;9) at week 12 and -31 meters (-53;-8) at week 24. The estimated MCID (95% CI) for "moderate improvement" vs "little improvement" was 20 meters (10;30) at week 12 and 24 meters (12;36) at week 24, whereas the MCID for 'moderate deterioration' vs "little deterioration" was -84 meters (-144;-24) at week 12 and -69 meters (-118;-20) at week 24. The estimated MCID for "much improvement" vs "moderate improvement" was 12 meters (-8;31) at week 12 and 13 meters (-2;28) at week 24. Estimates of the MCID for "much deterioration" vs "moderate deterioration" could not be calculated due to a lack of patients with "much deterioration" in PGA score at either time point. Results of the sensitivity analysis, including patients who died or were hospitalized, and using imputations are also shown in Table 3.

On assessing the significance of treatment allocation on the MCID, none of the interactions were found to be statistically significant (*P* values for little improvement vs no change, moderate improvement vs little improvement, much improvement vs moderate improvement, little deterioration vs no change, and moderate deterioration vs little deterioration were 0.19, 0.13, 0.69, 0.13 and 0.93, respectively, at week 12. For week 24, *P* values were 0.21, 0.61, 0.75, 0.22, and 0.40, respectively), but a clear scale effect was observed; there was a larger difference between the PGA classes in the FCM group compared with placebo. Similarly, no statistically significant interaction was seen on analyzing the influence of baseline tertiles ($<$ 240 meters [n = 244], 240$-$ $<$ 321 meters [n = 262] and $>$ 321 meters [n = 253]) of 6MWT on MCID (*P* values for little improvement vs no change, moderate improvement vs little improvement, much improvement vs moderate improvement, little deterioration vs no change, and moderate deterioration vs little deterioration were 0.15, 0.22, 0.16, 0.32, and 0.47, respectively, at week 12. For week 24, *P* values were 0.71, 0.77, 0.67, 0.25, and 0.30, respectively). However, a scale effect was

Table 3. MCID for Improvement and Deterioration in 6MWT at Weeks 12 and 24

	MCID (95% CI), m					
	Improvement			Deterioration		
PGA subgroups	Little (vs no change)	Moderate (vs little)	Much (vs moderate)	Little (vs no change)	Moderate (vs little)	Much (vs moderate)
Main Analysis						
Week 12	13.8 (5.1; 22.5)	19.7 (9.7; 29.6)	11.6 (-7.6; 30.7)	-11.1 (-31.5; 9.2)	-84.2 (-144.0; -24.3)	N/A
Week 24	15.0 (3.4; 26.7)	24.2 (12.4; 36.0)	13.4 (-1.7; 28.4)	-30.8 (-53.2; -8.4)	-68.8 (-117.9; -19.8)	N/A
FCM Subset Analysis						
Week 12	17.7 (6.0; 29.4)	20.7 (8.2; 33.2)	14.9 (-7.2; 37.2)	-28.8 (-56.8; -0.6)	-81.6 (-167.1; 3.8)	N/A
Week 24	20.1 (4.4; 35.8)	22.2 (9.0; 35.7)	15.4 (-3.0; 33.8)	-11.9 (-53.4; 29.6)	-94.5 (-191.2; 2.1)	N/A
Placebo Subset Analysis						
Week 12	5.9 (-7.4; 19.1)	3.8 (-11.6; 19.3)	6.3 (-26.8; 39.3)	2.9 (-26.2; 32.0)	-86.6 (-177.9; 4.8)	N/A
Week 24	5.3 (-12.2; 22.8)	15.7 (-7.7; 39.1)	10.1 (-10.8; 31.1)	-40.7 (-65.5;-15.9)	-53.0 (-110.1; 3.9)	N/A
Sensitivity Analysis*						
Week 12	14.1 (5.6; 22.5)	18.2 (8.5; 27.9)	18.2 (-1.6; 38.0)	-11.6 (-31.3; 8.2)	-73.1 (-129.0; - 17.3)	-154.0 (-274.5; -33.6)
Week 24	16.9 (5.8; 27.9)	23.3 (11.9; 34.8)	16.7 (0.9; 32.4)	-25.4 (-46.9; -3.9)	-60.2 (-106.9; -13.6)	-169.1 (-247.3; -90.9)

*Including imputation for patients who died or were hospitalized (categorized as having experienced much deterioration for the PGA analysis and 0 meters for improvement in the 6MWT; Note: the method for handling of deaths and hospitalizations when computing MCID is not well established), and the last observation carried forward for patients known to be alive and not hospitalized with baseline and $\geq$ 1 postbaseline value. CI, confidence interval; MCID, minimal clinically important difference; N/A, not available; PGA, patient global assessment; 6MWT, 6-minute walk test.

observed; patients who had lower baseline distances showed greater improvement.

Discussion

In this pooled analysis of patients with HFrEF and iron deficiency, we report 2 key findings. First, MCID estimates of "little improvement" vs "no change" in 6MWT among patients with HFrEF and iron deficiency were 14 meters and 15 meters at weeks 12 and 24, respectively. This suggests that even small changes in the 6MWT are clinically meaningful, with the MCID for improvement remaining largely similar between the 2 time points studied. Second, MCID estimates of "little deterioration" in the 6MWT among patients with HFrEF and iron deficiency were -11 meters and -31 meters at weeks 12 and 24, respectively. These results suggest that, in contrast to improvements, the MCID for deterioration may become greater and more difficult to achieve over longer follow-up durations. Together, these results have clinical implications because data on MCID in distance walked during the 6MWT can aid clinicians and researchers in establishing therapeutic thresholds that are objective, measurable and patient-centered. These, in turn, can be used to assess the efficacy of interventions for improving the exercise capacity of patients with HFrEF and iron deficiency.

Comparison With MCIDs Observed in Other Analyses

Compared with previous studies that used the Global Rating Scale (GRS) as a clinical anchor to investigate MCIDs for the 6MWT in patients with HF, the MCID estimates observed in our analysis of the study's patients who had HFrEF and iron deficiency were lower[25,26]; using a 15-point GRS collapsed to 7 levels as a clinical-based anchor, Spertus et al. reported an estimate of 55 meters for moderate improvement over a 6-week time period,[23] compared with 34 meters at 12 weeks in this study; using a 5-point GRS, O'Keeffe et al. found clinically meaningful changes of 43 meters to correspond with "little improvement" at 4-week follow-up,[25] in contrast to the 14 meters at 12 weeks observed here. However, these prior studies were limited by small sample sizes, and a large proportion of patients were lost to follow-up.

It should also be noted that MCIDs are not universal and cannot be used to compare different patient populations[27]; rather, MCIDs are more contextual, depending on various factors, such as baseline characteristics and demographics, disease severity, the clinical anchor used, and calculation methods.[27–29] This may have accounted for the lower MCIDs observed in the current analysis of patients with HFrEF and iron deficiency. In this analysis, the MCIDs for deterioration were -11 meters and -31 meters at weeks 12 and 24, respectively, indicating that even small deteriorations in 6MWT distance can be clinically meaningful. However, it should be noted that in patients with HFrEF, the 6MWT distance is an independent predictor of poorer survival and greater likelihood of hospitalization.[30] Therefore, patients who have already experienced even small deteriorations in their condition and have achieved a correspondingly shorter 6MWT distance may be at risk of poorer outcomes; this, rather than just the degree of change in 6MWT distance, needs to be taken into account. To the best of our knowledge, this is the first study to report estimates of MCID for improvement and deterioration in the 6MWT in patients with HFrEF and iron deficiency (with or without anemia) using clinical trial data. It is also important to emphasize that although "little improvement" has generally been considered MCID, other authors have used varying thresholds, such as "moderate improvement," to define MCID. These differences exist because of the different perceptions of what is "clinically meaningful." The clinically meaningful thresholds may vary based on several factors, such as invasiveness and the cost of the intervention.

Magnitudes of MCID Across the Spectrum of Improvements and Deteriorations

We found that the MCIDs for "little improvement" (vs "no change"), for "moderate improvement" (vs "little improvement"), and for "much improvement" (vs "moderate improvement") in the 6MWT were consistent and ranged from 12–20 meters between categories. Moreover, it is important to note that the MCID for "little improvement" (vs no change) remained relatively stable between weeks 12 and 24 (14 meters and 15 meters, respectively), indicating that improvements may remain stable over longer follow-ups. In contrast, the MCID for "little deterioration" (vs no change) showed a decrease between weeks 12 and 24 (-11 meters and -31 meters, respectively). We also observed that the MCID for "moderate deterioration" (vs "little deterioration") was much larger than that for "moderate improvement" (vs "little improvement"), indicating that it may be relatively easier to achieve a moderate clinically meaningful improvement in functional capacity than to achieve a moderate clinically meaningful deterioration. Future studies should evaluate MCIDs for improvements and deteriorations for longer follow-up periods to assess stability and adjust for measured interventions.

Use of MCIDs When Interpreting Clinical Trial Data

Interpretation of future clinical trials such as HEART-FID (Randomized Placebo Controlled Trial of

Ferric Carboxymaltose as Treatment of Heart Failure with Iron Deficiency) can be based on the MCIDs presented in this study.[31] For instance, the proportion of patients who experience clinically meaningful changes in health in response to treatments can be reported. In addition, the reported thresholds can be used to categorize patients as responders or nonresponders in terms of improvement or deterioration in clinical status, which can then be used to compare the effectiveness of treatments at an individual patient level so as to establish the minimum number of patients needed to treat in order to detect clinically relevant changes and to measure the stability of the response. Provided that an adequate safety margin is considered, clinically relevant improvements in exercise capacity may also be used as a measure in the approval process for new therapeutic agents.[32] Future studies should include other interventions to further establish MCIDs for improvement and deterioration in patients with HFrEF.

Study Limitations

There are several limitations in this study. First, generalizability is limited because the patient population consisted mainly of white, adult patients with HFrEF (predominantly NYHA class III) and with iron deficiency (with or without anemia). For patients with different ethnic backgrounds, preserved ejection fraction, more or less severe NYHA class, and HF without iron deficiency, different MCIDs may apply. Second, this analysis assessed only changes over a relatively short period, 12−24 weeks; some of the interventions employed, such as interdisciplinary care, may benefit from longer follow-ups. Third, approximately 60% of the patients in this analysis received intravenous FCM; iron repletion in patients with initial iron deficiency may specifically increase functional capacity and perceived quality of life by separate mechanisms. Nevertheless, pooling data from FCM and placebo arms to calculate MCID allowed inclusion of subjects across the whole range of PGA categories, with exploratory, nonsignificant differences between FCM and placebo arms reflecting high variability across subjects in the subjective PGA endpoint. Fourth, although the 6MWT is a simple, low-cost test with good reliability, peak oxygen consumption is considered the gold-standard for assessing aerobic capacity.[33] However, peak oxygen consumption is expensive and not easy to measure and was not recorded in the FAIR-HF and CONFIRM-HF trials. It was measured in the EFFECT-HF study of FCM in HfrEF,[34] but that study was not included in the present pooled analysis because of its open-label design. Last, analysis of MCID for deterioration was based on a small number of patients, and it can be expected that post hoc analyses of clinical trials may favor the responders and survivors.

Conclusions

In conclusion, the MCID for improvement in the 6MWT in patients with HFrEF, iron deficiency and a 6MWT distance of 281 meters at baseline was 14 meters at 12 weeks, suggesting that even small changes in the 6MWT can be clinically meaningful. Our findings could be used by clinicians to evaluate the efficacy of specific interventions in improving exercise capacity. Furthermore, these thresholds could be used in sample-size calculations in future trials involving patients with HFrEF, aiding in assessing efficacies and, thus, drug approvals for this population. Further exploration using various clinical anchors, interventions and calculation methods should be undertaken in future studies to validate and establish the MCID for 6MWT in patients with HFrEF and iron deficiency, with or without anemia.

Funding

The FAIR-HF and CONFIRM-HF trials were funded by Vifor Pharma. GMCR was supported by funding of the Italian Ministry of Health (Ricerca Corrente: 20/1819).

Lay summary

- The 6-minute walk test (6MWT, which measures the maximum distance a person can walk in 6 minutes) is an indicator of exercise capacity; it is important to know what the size of change in the 6MWT distance over time represents in terms of a meaningful change for an individual.
- In this study of patients with heart failure and iron deficiency, a 6MWT distance increase of 14 meters or more in 12 weeks was shown to represent a meaningful improvement, whereas a decrease of 31 meters or more represented a meaningful worsening.
- These findings could be used by doctors to evaluate whether certain treatments may provide a patient with meaningful improvements in exercise capacity.

This analysis combined data from 2 studies in patients with heart failure and iron deficiency. A patient survey was used to measure quality of life, and the 6-minute walk test (6MWT) was used to measure exercise capacity. The authors investigated the magnitude of change in 6MWT distance that corresponded with a meaningful change in quality of life. Results showed that a 6MWT distance increase of 14 meters or more in 12 weeks represented a meaningful improvement, whereas a decrease of 31 meters or more represented a meaningful worsening. These

findings could be used to help doctors interpret the results of clinical studies.

Conflicts of Interest

SDA has received research grants and personal fees from Vifor Int and Abbott Vascular (IIT/Trial steering committee work), personal fees from Bayer, Boehringer Ingelheim and Impulse Dynamics (Trial steering committee work), Novartis, Cardiac Dimensions and Occlutech (Adivsory committee work), and Servier (Registry Steering Committee). TF reports support for statistical consultancies and personal fees from Vifor for the present manuscript, consulting fees for statistical consultancies and personal fees from Bayer, CSL Behring, Galapagos, Minoryx, Vifor, Novartis, and LivaNova outside of the current work; payment for educational events from Fresenius Kabi outside of the current work; personal fees from Novartis, Eli Lilly, Bayer, BiosenseWebster, Janssen, Roche, and Enanta for participation on a Data Safety Monitoring Board. EAJ has received research grants and personal fees from Vifor Pharma (co-PI of the AFFIRM trial); personal fees from Bayer, Novartis, Abbott, Boehringer Ingelheim, Pfizer, Servier, AstraZeneca, Berlin Chemie, Cardiac Dimensions, Fresenius, Respicardia, Takeda, Swixx Biopharma, and Gedeon Richter; treasurer of the Executive Committee for the Heart Failure Association. MM has received personal fees from Vifor Pharma (Executive Committee member), Amgen (Executive Committee member and National PI), AstraZeneca, Abbott vascular, Bayer (participation in Advisory Boards), Boehringer Ingelheim (advisory board member), Servier (participation in Advisory Boards and speeches at sponsored symposia), Edwards Therapeutics (speeches at sponsored symposia), Actelion (DMC Member), LivaNova (Executive Committee member), and Windtree Therapeutics (Executive Committee member and Advisory Board). ILP reports personal fees from Boehringer Ingelheim outside of the submitted work. AJSC reports personal fees from AstraZeneca, Bayer, Boehringer Ingelheim, Menarini, Novartis, Nutricia, Servier, Vifor, Abbott, Actimed, Arena, Cardiac Dimensions, Corvia, CVRx, Enopace, ESN Cleer, Faraday, Gore, Impulse Dynamics, and Respicardia outside of the submitted work. BR, UMG and FD are full-time employees of Vifor Pharma. JCC reports unrestricted grants from Vifor Pharma and Novartis; consulting fees from Vifor Pharma, AstraZenica and Boehringer Ingelheim; and honoraria for conference activities from Vifor Pharma, AstraZenica and Boehringer Ingelheim. GSF reports grants from the European Commission; payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Bayer and Boehringer Ingelheim; participation on a data safety monitoring board or advisory board from Bayer and Boehringer Ingelheim; leadership or fiduciary role in the Heart Failure Association; and other financial or nonfinancial interests as a committee member for Medtronic, Vifor Pharma, Amgen, Servier, and Novartis. PP reports participation in clinical trials for and grants and personal fees from Vifor Pharma during the conduct of the study; participation in clinical trials for and personal fees from Amgen, Bayer, Novartis, AbbottVascular, Boehringer Ingelheim, Pfizer, Servier, Astra Zeneca, Cibiem, BMS, and Impulse Dynamics outside of the submitted work; participation in clinical trials for Cardiac Dimensions outside of the submitted work; and personal fees from Berlin Chemie outside of the submitted work. JB reports personal consulting fees from Abbott, Adrenomed, Amgen, Applied Therapeutics, Array, AstraZeneca, Bayer, Boehringer Ingelheim, CVRx, G3 Pharma, Impulse Dynamics, Innolife, Janssen, LivaNova, Luitpold, Medtronic, Merck, Novartis, NovoNordisk, Relypsa, Sequana Medical, and Vifor Pharma, and payment for lectures, presentations, speakers' bureaus, manuscript writing, and educational events from AstraZeneca, Eli-Lilly, Janssen, and Novartis. DJVV, GR and MSK have no conflicts of interest relating to the performance of this research or in the evaluation and publication process of the manuscript.

Acknowledgments

The authors acknowledge assistance with language editing, formatting, preparation of figures, and submission provided by Helen Sims (AXON Communications), funded by Vifor Pharma.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.cardfail.2022.10.423.

References

1. Arena R, Cahalin LP, Borghi-Silva A, Phillips SA. Improving functional capacity in heart failure: the need for a multifaceted approach. Curr Opin Cardiol 2014;29:467–74.
2. Stevenson LW, Hellkamp AS, Leier CV, Sopko G, Koelling T, Warnica JW, et al. Changing preferences for survival after hospitalization with advanced heart failure. J Am Coll Cardiol 2008;52:1702–8.
3. Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. BMJ 2010;341:c3920.
4. Forman DE, Arena R, Boxer R, Dolansky MA, Eng JJ, Fleg JL, et al. Prioritizing functional capacity as a principal end point for therapies oriented to older adults with cardiovascular disease: a scientific statement for

healthcare professionals from the American Heart Association. Circulation 2017;135:e894–918.
5. Shoemaker MJ, Curtis AB, Vangsnes E, Dickinson MG. Triangulating clinically meaningful change in the six-minute walk test in individuals with chronic heart failure: a systematic review. Cardiopulm Phys Ther J 2012;23:5–15.
6. Opasich C, Pinna GD, Mazza A, Febo O, Riccardi R, Riccardi PG, et al. Six-minute walking performance in patients with moderate-to-severe heart failure; is it a useful indicator in clinical practice? Eur Heart J 2001;22:488–96.
7. Cahalin L, Pappagianopoulos P, Prevost S, Wain J, Ginns L. The relationship of the 6-min walk test to maximal oxygen consumption in transplant candidates with end-stage lung disease. Chest 1995;108:452–9.
8. Rostagno C, Galanti G, Comeglio M, Boddi V, Olivo G, Serneri GG. Comparison of different methods of functional evaluation in patients with chronic heart failure. Eur J Heart Fail 2000;2:273–80.
9. Cahalin LP, Mathier MA, Semigran MJ, Dec GW, DiSalvo TG. The six-minute walk test predicts peak oxygen uptake and survival in patients with advanced heart failure. Chest 1996;110:325–32.
10. Hamo CE, Gheorghiade M, Butler J. Novel endpoints for heart failure clinical trials. Curr Heart Fail Rep 2017;14:210–6.
11. Anker SD, Schroeder S, Atar D, Bax JJ, Ceconi C, Cowie MR, et al. Traditional and new composite endpoints in heart failure clinical trials: facilitating comprehensive efficacy assessments and improving trial efficiency. Eur J Heart Fail 2016;18:482–9.
12. Polkey MI, Spruit MA, Edwards LD, Watkins ML, Pinto-Plata V, Vestbo J, et al. Six-minute-walk test in chronic obstructive pulmonary disease: minimal clinically important difference for death or hospitalization. Am J Respir Crit Care Med 2013;187:382–6.
13. Holland AE, Nici L. The return of the minimum clinically important difference for 6-minute-walk distance in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2013;187:335–6.
14. Mathai SC, Puhan MA, Lam D, Wise RA. The minimal important difference in the 6-minute walk test for patients with pulmonary arterial hypertension. Am J Respir Crit Care Med 2012;186:428–33.
15. Bohannon RW, Crouch R. Minimal clinically important difference for change in 6-minute walk test distance of adults with pathology: a systematic review. J Eval Clin Pract 2017;23:377–81.
16. Shoemaker MJ, Curtis AB, Vangsnes E, Dickinson MG. Clinically meaningful change estimates for the six-minute walk test and daily activity in individuals with chronic heart failure. Cardiopulm Phys Ther J 2013;24:21–9.
17. Yap J, Lim FY, Gao F, Teo LL, Lam CS, Yeo KK. Correlation of the New York Heart Association Classification and the 6-Minute Walk Distance: a systematic review. Clin Cardiol 2015;38:621–8.
18. Anker SD, Comin Colet J, Filippatos G, Willenheimer R, Dickstein K, Drexler H, et al. Ferric carboxymaltose in patients with heart failure and iron deficiency. N Engl J Med 2009;361:2436–48.
19. Ponikowski P, van Veldhuisen DJ, Comin-Colet J, Ertl G, Komajda M, Mareev V, et al. Beneficial effects of long-term intravenous iron therapy with ferric carboxymaltose in patients with symptomatic heart failure and iron deficiency. Eur Heart J 2015;36:657–68.
20. Anker SD, Colet JC, Filippatos G, Willenheimer R, Dickstein K, Drexler H, et al. Rationale and design of Ferinject assessment in patients with IRon deficiency and chronic Heart Failure (FAIR-HF) study: a randomized, placebo-controlled study of intravenous iron supplementation in patients with and without anaemia. Eur J Heart Fail 2009;11:1084–91.
21. Ponikowski P, van Veldhuisen DJ, Comin-Colet J, Ertl G, Komajda M, Mareev V, et al. Rationale and design of the CONFIRM-HF study: a double-blind, randomized, placebo-controlled study to assess the effects of intravenous ferric carboxymaltose on functional capacity in patients with chronic heart failure and iron deficiency. ESC Heart Fail 2014;1:52–8.
22. Cheung YT, Foo YL, Shwe M, Tan YP, Fan G, Yong WS, et al. Minimal clinically important difference for the functional assessment of cancer therapy: cognitive function in breast cancer patients. J Clin Epidemiol 2014;67:811–20.
23. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61:102–9.
24. Butler J, Khan MS, Mori C, Filippatos GS, Ponikowski P, Comin-Colet J, et al. Minimal clinically important difference in quality of life scores for patients with heart failure and reduced ejection fraction. Eur J Heart Fail 2020;22:999–1005. https://doi.org/10.1002/ejhf.1810.
25. O'Keeffe ST, Lye M, Donnellan C, Carmichael DN. Reproducibility and responsiveness of quality of life assessment and six-minute walk test in elderly heart failure patients. Heart 1998;80:377–82.
26. Spertus J, Peterson E, Conard MW, Heidenreich PA, Krumholz HM, Jones P, et al. Monitoring clinical changes in patients with heart failure: a comparison of methods. Am Heart J 2005;150:707–15.
27. Wright A, Hannon J, Hegedus EJ, Kavchak AE. Clinimetrics corner: a closer look at the minimal clinically important difference (MCID). J Man Manip Ther 2012;20:160–6.
28. Lauridsen HH, Hartvigsen J, Manniche C, Korsholm L, Grunnet-Nilsson N. Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients. BMC Musculoskelet Disord 2006;7:82.
29. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J 2007;7:541–6.
30. Matsumoto K, Xiao Y, Homma S, Thompson JLP, Buchsbaum R, Ito K, et al. Prognostic impact of 6 min walk test distance in patients with systolic heart failure: insights from the WARCEF trial. ESC Heart Fail 2021;8:819–28.
31. Mentz RJ, Ambrosy AP, Ezekowitz JA, Lewis GD, Butler J, Wong YW, et al. HEART-FID Trial Investigators. Randomized placebo-controlled trial of ferric carboxymaltose in heart failure with iron deficiency: rationale and design. Circ Heart Fail 2021;14:e008100. https://doi.org/10.1161/CIRCHEARTFAILURE.120.008100.
32. Zannad F, Garcia AA, Anker SD, Armstrong PW, Calvo G, Cleland JGF, et al. Clinical outcome endpoints in heart failure trials: a European Society of Cardiology

Heart Failure Association consensus document. Eur J Heart Fail 2013;15:1082–94.
33. Ross RM, Murthy JN, Wollak ID, Jackson AS. The six-minute walk test accurately estimates mean peak oxygen uptake. BMC Pulm Med 2010;10:31.
34. van Veldhuisen DJ, Ponikowski P, van der Meer P, Metra M, Böhm M, Doletsky A, et al. Effect of ferric carboxymaltose on exercise capacity in patients with chronic heart failure and iron deficiency. Circulation 2017;136:1374–83.