



This is a repository copy of *Establishing minimally important differences for cardiac MRI end-points in pulmonary arterial hypertension*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/202812/>

Version: Published Version

Article:

Alabed, S. orcid.org/0000-0002-9960-7587, Garg, P., Alandejani, F. orcid.org/0000-0003-1227-401X et al. (24 more authors) (2023) Establishing minimally important differences for cardiac MRI end-points in pulmonary arterial hypertension. *European Respiratory Journal*, 62 (2). 2202225. ISSN 0903-1936

<https://doi.org/10.1183/13993003.02225-2022>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

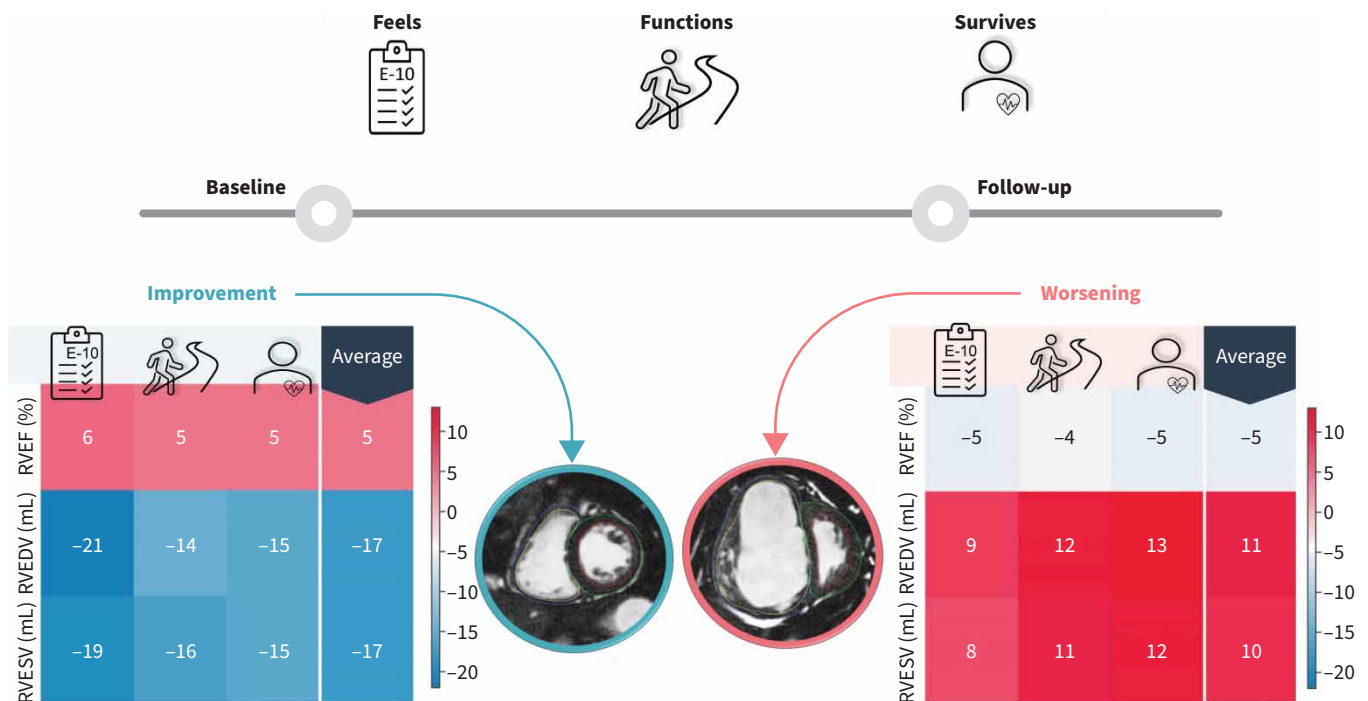


eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>



Establishing minimally important differences for cardiac MRI end-points in pulmonary arterial hypertension

Samer Alabed , Pankaj Garg, Faisal Alandejani , Krit Dwivedi, Ahmed Maiter, Kavita Karunasaagar, Smitha Rajaram, Catherine Hill, Steven Thomas, Rebecca Gosling, Michael J. Sharkey , Mahan Salehi, Jim M. Wild, Lisa Watson, Abdul Hameed , Athanasios Charalampopoulos, Haiping Lu, Alex M.K. Rothman, A.A. Roger Thompson , Charlie A. Elliot, Neil Hamilton, Christopher S. Johns, Iain Armstrong, Robin Condliffe , Rob J. van der Geest, Andrew J. Swift and David G. Kiely 



GRAPHICAL ABSTRACT Summary of cardiac magnetic resonance (CMR) absolute change minimally important differences (MIDs) for how a patient “feels” benchmarked to the emPHasis-10 (E-10) quality of life questionnaire, “functions” benchmarked to the incremental shuttle walk test or “survives” benchmarked to 1-year mortality post-follow-up. RVEF: right ventricular ejection fraction; RVEDV: right ventricular end-diastolic volume; RVESV: right ventricular end-systolic volume.



Establishing minimally important differences for cardiac MRI end-points in pulmonary arterial hypertension

Samer Alabed ^{1,2,3}, Pankaj Garg⁴, Faisal Alandejani ¹, Krit Dwivedi^{1,2}, Ahmed Maiter^{1,2}, Kavita Karunasaagarar^{1,2}, Smitha Rajaram^{1,2}, Catherine Hill², Steven Thomas^{1,2}, Rebecca Gossling¹, Michael J. Sharkey ¹, Mahan Salehi¹, Jim M. Wild^{1,3}, Lisa Watson⁵, Abdul Hameed ⁵, Athanasios Charalampopoulos⁵, Haiping Lu^{3,6}, Alex M.K. Rothman¹, A.A. Roger Thompson ^{1,5}, Charlie A. Elliot^{1,5}, Neil Hamilton^{1,5}, Christopher S. Johns^{1,2}, Iain Armstrong^{1,5}, Robin Condliffe ^{1,5}, Rob J. van der Geest⁷, Andrew J. Swift^{1,2,3,8,9} and David G. Kiely ^{1,3,5,8,9}

¹Department of Infection, Immunity and Cardiovascular Disease, University of Sheffield, Sheffield, UK. ²Department of Clinical Radiology, Sheffield Teaching Hospitals, Sheffield, UK. ³INSIGNEO, Institute for *in silico* Medicine, University of Sheffield, Sheffield, UK. ⁴Norwich Medical School, University of East Anglia, Norwich, UK. ⁵Sheffield Pulmonary Vascular Disease Unit, Royal Hallamshire Hospital, Sheffield, UK. ⁶Department of Computer Science, University of Sheffield, Sheffield, UK. ⁷Leiden University Medical Center, Leiden, The Netherlands. ⁸National Institute for Health and Care Research, Sheffield Biomedical Research Centre, Sheffield, UK. ⁹Joint senior authors.

Corresponding author: Samer Alabed (s.alabed@nhs.net)



Shareable abstract (@ERSpublications)

Changes in key cardiac MRI metrics reflect changes in how a patient “feels, functions or survives” in pulmonary arterial hypertension <https://bit.ly/3IE21Ow>

Cite this article as: Alabed S, Garg P, Alandejani F, *et al.* Establishing minimally important differences for cardiac MRI end-points in pulmonary arterial hypertension. *Eur Respir J* 2023; 62: 2202225 [DOI: 10.1183/13993003.02225-2022].

Copyright ©The authors 2023.

This version is distributed under the terms of the Creative Commons Attribution Licence 4.0.

Received: 20 Nov 2022
Accepted: 23 May 2023

Abstract

Background Cardiac magnetic resonance (CMR) is the gold standard technique to assess biventricular volumes and function, and is increasingly being considered as an end-point in clinical studies. Currently, with the exception of right ventricular (RV) stroke volume and RV end-diastolic volume, there is only limited data on minimally important differences (MIDs) reported for CMR metrics. Our study aimed to identify MIDs for CMR metrics based on US Food and Drug Administration recommendations for a clinical outcome measure that should reflect how a patient “feels, functions or survives”.

Methods Consecutive treatment-naïve patients with pulmonary arterial hypertension (PAH) between 2010 and 2022 who had two CMR scans (at baseline prior to treatment and 12 months following treatment) were identified from the ASPIRE registry. All patients were followed up for 1 additional year after the second scan. For both scans, cardiac measurements were obtained from a validated fully automated segmentation tool. The MID in CMR metrics was determined using two distribution-based (0.5SD and minimal detectable change) and two anchor-based (change difference and generalised linear model regression) methods benchmarked to how a patient “feels” (emPHasis-10 quality of life questionnaire), “functions” (incremental shuttle walk test) or “survives” for 1-year mortality to changes in CMR measurements.

Results 254 patients with PAH were included (mean±SD age 53±16 years, 79% female and 66% categorised as intermediate risk based on the 2022 European Society of Cardiology/European Respiratory Society risk score). We identified a 5% absolute increase in RV ejection fraction and a 17 mL decrease in RV end-diastolic or end-systolic volumes as the MIDs for improvement. Conversely, a 5% decrease in RV ejection fraction and a 10 mL increase in RV volumes were associated with worsening.

Conclusions This study establishes clinically relevant CMR MIDs for how a patient “feels, functions or survives” in response to PAH treatment. These findings provide further support for the use of CMR as a clinically relevant clinical outcome measure and will aid trial size calculations for studies using CMR.

Introduction

In patients with pulmonary arterial hypertension (PAH), symptoms and survival are determined primarily by right ventricular (RV) function. In PAH, a progressive pulmonary vasculopathy results in elevation of



mean pulmonary arterial pressure (mPAP) and an increase in RV afterload [1]. With disease progression and chronically elevated mPAP, the right ventricle undergoes remodelling, resulting in either adaptation and maintenance of output [2] or maladaptation, RV failure and consequently reduced survival [3]. Cardiac magnetic resonance (CMR) is the gold standard for assessing the right ventricle and shows potential in the assessment of PAH [4]. Impairments of RV function and associated increases in RV volumes can be detected and quantified by CMR, enabling prediction of clinical worsening and mortality [5] and aiding risk stratification [6]. In addition, CMR is sensitive to improvements in RV function following PAH therapy [7–10] and detects a larger treatment effect than the 6-min walk test (6MWT) [11]. In this context, CMR is an important tool for risk stratification and monitoring of disease and treatment response in PAH [4].

Phase 4 clinical studies of PAH therapies have recently utilised CMR as a primary end-point in addition to other composite outcomes [9, 10]. Assessing treatment response with CMR necessitates clinically relevant thresholds in order to determine improvement or worsening. However, only RV stroke volume (RVSV) measured on pulmonary artery phase contrast flow imaging has established thresholds [12], while those for volumetric CMR measurements on cine imaging with the exception of RV end-diastolic volume (RVEDV) remain unvalidated [13]. The introduction of automatic volumetric CMR measurements assessing RV changes over time has several advantages. It offers excellent repeatability in scan–rescan assessment and has higher accuracy than manual assessment [14]. In addition, results are generalisable across different centres and magnetic resonance imaging (MRI) systems, allowing standardised comparisons independent of the location of scan [14–16].

The US Food and Drug Administration has highlighted the need to identify clinical outcome measures for PAH therapy trials that reflect how a patient “feels, functions or survives” [17]. To reflect this, we aimed to identify clinically relevant thresholds for change in automatically derived CMR RV and left ventricular (LV) measurements, benchmarking against patient-reported outcome measures (“feels”), exercise testing (“functions”) and mortality (“survives”). Our results should aid the management of patients in the clinic by identifying clinically meaningful changes in key CMR metrics, and aid researchers by informing power calculations and the selection of end-points for clinical studies using CMR.

Methods

Study sample

Adult patients with PAH were identified from the ASPIRE (Assessing the Spectrum of Pulmonary hypertension Identified at a REferral centre) registry [18] between January 2010 and January 2022. Diagnosis of PAH was based on mPAP ≥ 25 mmHg and pulmonary arterial wedge pressure ≤ 15 mmHg and pulmonary vascular resistance (PVR) ≥ 3 WU, measured by right heart catheterisation. Patients were eligible for inclusion if they had: 1) baseline CMR prior to starting treatment and within 48 h of PAH diagnosis; 2) follow-up CMR at 12–24 months; and 3) at least 1 year follow-up after the follow-up scan. Patients were excluded if they did not have complete short-axis stack imaging for both baseline and repeat scans. The local ethics committee and institutional review board approved this study (ASPIRE; c06/Q2308/8).

Imaging procedures

MRI protocol

CMR was performed with 1.5 T MRI systems (Signa HDx; GE Healthcare, Chalfont St Giles, UK). Short-axis cine images were acquired using a cardiac-gated multislice balanced steady-state free precession sequence (20 frames per cardiac cycle, section thickness 10 mm, 0 mm inter-section gap, field of view 480 mm, acquisition matrix 256×200, flip angle 60°, bandwidth 125 kHz per pixel, repetition time/echo time (T_R/T_E) 3.7/1.6 ms). A stack of images in the short-axis plane was acquired, fully covering both ventricles from base to apex. End-systole was considered to be the smallest cavity area. End-diastole was defined as the first cine phase of the R-wave-triggered acquisition or largest volume. Patients were supine with a surface coil and with retrospective ECG gating.

Image analysis

An in-house deep-learning CMR segmentation tool was used to obtain fully automatic measurements [14]. The segmentation tool was trained in a multicentre, multivendor and multipathology dataset, and was previously validated by assessing: 1) accuracy against same-day invasive pulmonary haemodynamics and phase contrast flow imaging; 2) repeatability in a same-day scan–rescan cohort; 3) generalisability in an external testing cohort; and 4) mortality prediction in a large cohort with multiple cardiac and lung pathologies. The automatic contours included trabeculations in the blood pool and were obtained using MASS software (MASS research version 2020; Leiden University Medical Center, Leiden, The Netherlands).

Clinical parameters

The emPHasis-10 (E-10) questionnaire, a patient-reported outcome measure to assess health-related quality of life in patients with PAH, was completed at baseline and at the time of the follow-up scan from 2014 onwards. Each patient completed 10 questions ranked on a scale of 0 to 5, with a lower score indicating a better quality of life [19]. The incremental shuttle walk test (ISWT) was performed as part of routine patient evaluation according to the standard method [20]. Patients complete a 10 m length keeping in time to an external audible signal. Level 1 consists of three lengths (30 m) and each additional level adds one extra 10 m length to the preceding level. Each level takes 1 min to complete and the test finishes at the end of level 12, a distance of 1020 m. The patient continues until they are too breathless or unable to keep up the required pace. The REVEAL 2.0 and the 2022 European Society of Cardiology (ESC)/European Respiratory Society (ERS) risk scores were calculated from composite clinical parameters [21, 22] and modified to include the ISWT instead of the 6MWT [6, 23]. Mortality data were collected from the electronic records of the National Health Service (NHS) Personal Demographics Service. The NHS automatically updates the mortality records once a death is registered in the UK. All patients were followed up as part of the national service specification for patients with pulmonary hypertension for a minimum of 12 months.

Statistical analysis

Baseline characteristics are presented as proportion, mean \pm SD or median (interquartile range (IQR)). For each CMR parameter, both the absolute and relative differences were calculated. The absolute difference was determined by subtracting the baseline measurement from the follow-up measurement, while the relative difference was calculated as the ratio of the absolute difference to the baseline measurement.

We employed four methods to derive the minimally important difference (MID) estimates: two distribution-based and two anchor-based methods (table 1). Initially, Pearson's correlation analysis was used to determine at least a weak correlation ($r>0.20$) between the change in CMR measurements and the anchor [24, 25]. The difference in CMR parameters and difference in the E-10 or ISWT were regressed onto a scale of 6 units using z-score normalisation to assess the correlation of these changes. Any CMR measurement failing to meet the correlation threshold was subsequently excluded from further MID analysis [24].

Anchors for how a patient "feels" and "functions" were determined using a patient-reported outcome measure (E-10) and an assessment of exercise capacity (ISWT), respectively. Patients were considered to have improved, remained stable or worsened between baseline and follow-up, based on a change of 6 points in the E-10 [26, 27] or 47.5 m in the ISWT [28, 29]. For how a patient "survives", changes in CMR measurements in patients who survived 1 year post-follow-up scan were compared with patients who did not. For the anchor-based method, we derived MID estimates using change difference and regression analysis. The change difference was identified as the difference between the mean change in CMR measurements in patients who had improved or worsened (defined by the anchor) and the mean change in stable patients. The change difference method effectively adjusts the degree of change in the improved or worsened group according to the change observed in the stable group. For the regression analysis, we employed a generalised linear model regression to predict the difference in scores (DScore) between

TABLE 1 Methods employed to calculate the minimally important difference

Method	Type	Definition
0.5sd	Distribution-based	Estimated as 0.5 times the SD of the change in CMR measurements within each patient group (improved/worsened).
Minimal detectable change	Distribution-based	Calculated based on the SEM. The SEM was calculated by multiplying the SD of the change in CMR measurements by the square root of (1–reliability coefficient). The formula applied was: $MDC=1.96\times\sqrt{2\times SEM}$.
Change difference	Anchor-based	Difference between the mean change in CMR measurements in patients who had improved or worsened (according to the anchor) and the mean change in stable patients.
GLM regression	Anchor-based	Determined by the estimated coefficients for improvement and worsening, derived from a regression analysis using the anchor as a predictor for changes in CMR measurements.

SD: standard deviation; CMR: cardiac magnetic resonance; SEM: standard error of measurement; GLM: general linear model.

baseline and follow-up CMR measurements, as represented by:

$$DScore = k + \beta_b X_{better} + \beta_w X_{worse} + \beta_s X_{stable}$$

The DScore represents the change in CMR parameters and patient status (defined by the anchor as better (X_{better}) or worse (X_{worse})) was entered as a dummy variable in the model. The coefficients of better (β_b) and worse (β_w) in the regression model estimated the incremental difference in scores when patient status transitioned to better or worse compared with stable patients.

In the distribution-based approach, the MID was estimated based on the distribution of CMR measurements within each patient group (improved/worsened). In the first distribution method, the MID was estimated as 0.5 times the standard deviation ($0.5SD$) of the change in CMR measurements [30]. In the second distribution method, the minimal detectable change (MDC) was calculated using the formula: $MDC = 1.96 \times \sqrt{2 \times SEM}$, where SEM is the standard error of measurement calculated by multiplying the standard deviation of the difference in CMR measurements by the square root of 1 minus its reliability coefficient. Previously published consistency intraclass correlation coefficient values were used as the reliability coefficient of the CMR measurements [14]. Distribution-based methods were employed to facilitate the interpretation of the anchor-based method results. The standard error of measurement describes the variability between the observed and the true measurements. Changes in CMR measurements smaller than the corresponding standard error of measurement are more likely to represent an error of measurement rather than genuine changes [31]. In instances where the anchor-based MIDs were less than the standard error of measurement (*i.e.* indistinguishable from measurement error), the standard error of measurement was utilised as the MID [24].

Statistical analyses were carried out using the lifelines and pingouin Python libraries [32, 33] with a significance threshold of 0.05. Graphs were produced using the Matplotlib library [34] and Prism version 9 (GraphPad, La Jolla, CA, USA).

Results

A total of 254 treatment-naïve patients with PAH were included (figure 1). Patients had a mean \pm SD age of 53 \pm 16 years, with 79% female, 83% categorised as World Health Organization (WHO) Functional Class III, and an intermediate risk for mortality of 68% and 66% on REVEAL 2.0 and 2022 ESC/ERS risk models, respectively. 41% had PAH associated with connective tissue disease (PAH-CTD), 37% had idiopathic PAH (IPAH) and 22% had other types of PAH. The median (IQR) mPAP was 50 (41–59) mmHg and PVR was 812 (526–1154) $\text{dyn} \cdot \text{s} \cdot \text{cm}^{-5}$. The mean \pm SD E-10 score was 31 \pm 12 and ISWT walk distance was 207 \pm 175 m. Baseline patient characteristics are shown in table 2.

The median (IQR) duration between the baseline and repeat (follow-up) scan was 12 (7–16) months. Between the two scans patients were treated with phosphodiesterase 5 inhibitors (86%), endothelin receptor antagonists (72%), parenteral prostanoid (17%) and other medications (4%), with 32% receiving monotherapy, 48% dual combination and 20% triple combination therapy. During follow-up after the repeat scan, 25 out of 254 (10%) patients had died at 12 months and 123 out of 254 (48%) patients died during a median (IQR) period of 5 (3–7) years from baseline.

Table 3 shows the mean CMR measurements at baseline and follow-up for the different patient groups. The absolute differences in RV ejection fraction (RVEF) between baseline and follow-up have been reported in nine studies [8, 10, 11, 35–40] with a total of 321 treatment-naïve patients, and their pooled results compared with the current study are shown in supplementary figure S1 and presented in detail in supplementary table S1. The pooled mean difference in RVEF at 1 year was 6% (95% CI 3–8%) compared with 7% (95% CI 5–9%) in our study.

MIDs for how a patient “feels, functions or survives”

The E-10 and ISWT were performed mostly on the same day as the CMR and within 2 weeks, with a median (IQR) time between the CMR and E-10 of 0 (0–0) days and between the CMR and ISWT of 0 (0–8) days. Paired baseline and follow-up E-10 ($n=118$) and ISWT ($n=146$) categorised patients into improved ($n=42$ (35%) for E-10 and $n=57$ (39%) for ISWT), stable ($n=55$ (47%) for E-10 and $n=64$ (44%) for ISWT) and worsened ($n=21$ (18%) for E-10 and $n=25$ (17%) for ISWT) (table 2). The correlation between CMR parameters and the E-10 and ISWT was weak for RV parameters: RVEF ($r=-0.25$ and $r=0.20$, respectively), RVEDV ($r=0.28$ and $r=-0.28$, respectively) and RV end-systolic volume (RVESV) ($r=0.32$ and $r=0.34$, respectively). None of the LV parameters or RVSV ($r=0.10$ for both the E-10 and ISWT) showed a sufficient correlation with the anchors. The mean MID values for absolute and relative

		Treatment-naïve PAH			
		n=254			
Baseline	CMR	E-10	ISWT		
	n=254	n=118	n=146		
Follow-up	CMR	E-10	ISWT		
	n=254	n=118	n=146		
		Improved	42 (35%)	Improved	57 (39%)
		Stable	55 (47%)	Stable	64 (44%)
		Worse	21 (18%)	Worse	25 (17%)
1 year post-follow-up	Survived		229 (90%)		
	Died		25 (10%)		

FIGURE 1 Patient flowchart. First cardiac magnetic resonance (CMR) at baseline→repeat (follow-up) CMR at 12±6 months→1 year post-follow-up. The emPHasis-10 (E-10) health-related quality of life questionnaire and incremental shuttle walk test (ISWT) were performed mostly on the same day and within 2 weeks of CMR. PAH: pulmonary arterial hypertension.

improvement and worsening using the different MID methods are shown in figure 2 and supplementary figure S2. In summary, the overall MID means and range of means across methods for improvement were: 5% (3–9%) for RVEF, –17 mL (–6– –27 mL) for RVEDV and –17 mL (–11– –24 mL) for RVESV. For worsening, the values were –5% (–3– –9%) for RVEF, 11 mL (3–19 mL) for RVEDV and 10 mL (3–17 mL) for RVESV (graphical abstract). The highest relative change, indexed to the baseline value, was observed for RVEF (22% for improvement and –19% for worsening) (supplementary figure S2).

Discussion

Identifying clinically relevant thresholds for changes in CMR metrics that reflect how a patient “feels, functions or survives” has important implications for patient monitoring and the selection of therapy trial end-points. To the best of our knowledge, this is the first study to compare changes in CMR metrics with health-related quality of life in PAH in addition to measures of function and mortality, and the first to assess clinically relevant MIDs for automatic CMR measurements. MIDs for CMR metrics were identified using various distribution-based and anchor-based methods in an effort to generate reliable estimates [41]. The MIDs obtained using the change in E-10 score, ISWT walk distance and survival as anchors were remarkably consistent across methods and anchors, reinforcing the robustness of the MID estimates. Only RV metrics sufficiently correlated with the anchors to allow for MID calculations.

In this study of 254 patients, we observed a mean absolute difference of 7% in RVEF post-PAH treatment at an average of 1 year follow-up. Our result was comparable to the pooled estimate (6% mean difference in RVEF) identified from nine PAH studies, including 321 PAH patients with pre- and post-treatment RVEF measurements at 9–12 months. Two studies assessed change at 4 months following bosentan therapy: BENZA *et al.* [9] showed an absolute increase of 3% in RVEF in 84 patients, and WILKINS *et al.* [42] reported an increase in RVEDV of 6 mL and a decrease of RVESV of 2 mL in 12 patients. VAN DE VEERDONK *et al.* [37] assessed 52 patients at baseline and at 5 years follow-up, and found that a 3% absolute reduction in RVEF

TABLE 2 Baseline characteristics (n=254 treatment-naïve patients with pulmonary arterial hypertension (PAH))

Age (years)	55 (42–66)
Female	201 (79)
BSA (m ²)	1.80±0.35
PAH subcategory	
IPAH	94 (37)
CTD	105 (41)
CHD	21 (8)
Portal hypertension	16 (6)
Other PAH	18 (7)
WHO Functional Class	
II	10 (4)
III	210 (83)
IV	34 (13)
REVEAL 2.0 score	
≤6	11 (4)
7–8	173 (68)
≥9	70 (28)
2022 ESC/ERS risk	
Low	18 (7)
Intermediate	169 (66)
High	67 (26)
RHC parameters	
mPAP (mmHg)	50 (41–59)
PVR (dyn·s·cm ⁻⁵)	812 (526–1154)
PAWP (mmHg)	10 (8–12)
mRAP (mmHg)	9 (6–14)
CO (L·min ⁻¹)	4 (3–5)
S _{vO₂} (%)	64 (58–70)
Heart rate (beats·min ⁻¹)	77 (69–89)
PAH medication	
PDE5i	
Sildenafil	199 (78)
Tadalafil	20 (8)
ERA	
Ambrisentan	87 (34)
Bosentan	33 (13)
Macitentan	62 (24)
Parenteral prostanoid	43 (17)
Other	11 (4)
Therapeutic strategy	
Monotherapy	74 (32)
Dual combination	111 (48)
Triple combination	45 (20)

Data are presented as median (interquartile range), n (%) or mean±sd. BSA: body surface area; IPAH: idiopathic PAH; CHD: congenital heart disease; CTD: connective tissue disease; WHO: World Health Organization; ESC: European Society of Cardiology; ERS: European Respiratory Society; RHC: right heart catheterisation; mPAP: mean pulmonary arterial pressure; PVR: pulmonary vascular resistance; PAWP: pulmonary arterial wedge pressure; mRAP: mean right atrial pressure; CO: cardiac output; S_{vO₂}: mixed venous oxygen saturation; PDE5i: phosphodiesterase 5 inhibitor; ERA: endothelin receptor antagonist.

was associated with a lower survival rate in patients with decreased PVR. Our study identifies a 5% increase in RVEF and a 17 mL decrease in RVEDV and RVESV as MIDs for improvement, while a 5% decrease in RVEF and a 10 mL increase in RV volumes were associated with worsening.

CMR has been shown to be sensitive to change in response to treatment, with CMR detecting a larger treatment effect than the 6MWT [11]. BRADLOW *et al.* [7] previously estimated post-treatment thresholds for RV changes based on four studies with a total of 57 PAH patients. Although they suggested an absolute difference in RVEF of 3% and RV volumes of 10 mL as thresholds, these observations were minimally detectable changes based on the repeatability of the measurement and not benchmarked against clinically relevant outcomes such as changes in quality of life, exercise capacity or mortality. In contrast, in our study we have benchmarked CMR parameters against measures of how a patient “feels, functions or

TABLE 3 Changes in cardiac magnetic resonance measurements and clinical parameters: benchmarked to changes in how a patient “feels” (emPHasis-10 (E-10), n=118), “functions” (incremental shuttle walk test (ISWT), n=146) or “survives” (1-year mortality post-follow-up, n=254)

	Improved		Stable		Worsened	
	Baseline	Follow-up	Baseline	Follow-up	Baseline	Follow-up
E-10	n=42		n=55		n=21	
E-10 (points)	37±10	21±11	29±12	29±12	24±10	37±7
RVEF (%)	32±12	42±13	34±12	42±10	38±10	43±8
RVEDV (mL)	208±66	182±59	200±78	201±74	174±69	163±57
RVESV (mL)	146±60	110±53	138±68	122±60	112±56	94±40
ISWT	n=57		n=64		n=25	
Walk distance (m)	249±181	409±197	197±1501	196±148	259±16	158±142
RVEF (%)	31±10	42±10	34±11	40±10	33±12	38±10
RVEDV (mL)	235±70	216±86	208±79	196±74	187±69	192±70
RVESV (mL)	164±62	128±62	144±68	122±61	129±65	121±55
Mortality	Survivors n=229			Non-survivors n=25		
RVEF (%)	34±11	41±11			31±13	36±16
RVEDV (mL)	207±71	195±71			202±79	207±85
RVESV (mL)	141±62	118±56			144±72	139±75

Data are presented as mean±sd, unless otherwise stated. RVEF: right ventricular ejection fraction; RVEDV: right ventricular end-diastolic volume; RVESV: right ventricular end-systolic volume. Measurements were taken at baseline before treatment and at follow-up (12±6 months). Improvement was defined as an increase of at least 47.5 m in the ISWT, a decrease of at least 6 points in the E-10 score and 1 year survival post-follow-up, whereas worsening was defined as a reduction in the ISWT of at least 47.5 m, an increase of 6 points in the E-10 score or mortality. Otherwise, patients were described as stable.

survives”. Until now, the only CMR parameters with a clinically relevant and validated threshold in PAH have been RVSV derived from phase contrast flow imaging [12] and RVEDV measured from trans-axial cine images [13]; RVSV was anchored to a 6MWT, with 10 mL change identified as a threshold for important clinical effect, whereas RVEDV was anchored to a change in WHO Functional Class, with an 11% relative change identified as clinically relevant. In this study, we compared the change in CMR cine imaging to changes in patient-reported outcome measures (E-10 health-related quality of life) and exercise capacity (ISWT walk distance). The E-10 elicits how the patient “feels”, with domains reflecting the burden of breathlessness, fatigue and anxiety on patients with PAH [19], while the ISWT reflects how a patient “functions” by assessing exercise capacity [23, 29, 43]. Both the E-10 and ISWT have established MIDs, can assist in risk stratification of patients and have prognostic value, making them ideal benchmarks for assessing how a patient feels and functions [23, 26]. However, it must be noted that the E-10, ISWT and CMR all measure different components, confirmed by the weak correlation between differences in RV measurements and changes in the E-10 and ISWT, and therefore worsening in quality of life and exercise capacity can occur despite improvements in RV function. Despite being able to demonstrate that CMR metrics can detect MIDs for how a patient “feels, functions or survives”, a composite end-point that includes each of these domains will be superior to using a metric that focuses on a single measure such as cardiac function. Nonetheless, this study does provide further evidence for using CMR as a primary trial end-point in studies of PAH therapies by providing evidence that changes in key CMR metrics do reflect changes in how a patient feels and functions.

While an absolute change in CMR measurements gives an indication of direction, it does not take into account the baseline state of the patient to contextualise the magnitude of change. This may be relevant in patients with more severe disease where relative change is likely to be more sensitive to disease progression by accounting for a patient’s pre-treatment baseline and therefore should be considered alongside absolute changes in metrics [44]. In the current study, trabeculations were included in the blood pool. Further work to establish MIDs for CMR measurements excluding trabeculations would be of value and as technology evolves there is a need to develop standardised approaches to CMR measurements of the right ventricle.

Limitations

The major limitation of any longitudinal retrospective study is the inherent risk of selection bias. Inevitably, patients who survived until follow-up CMR imaging have had a less severe disease course

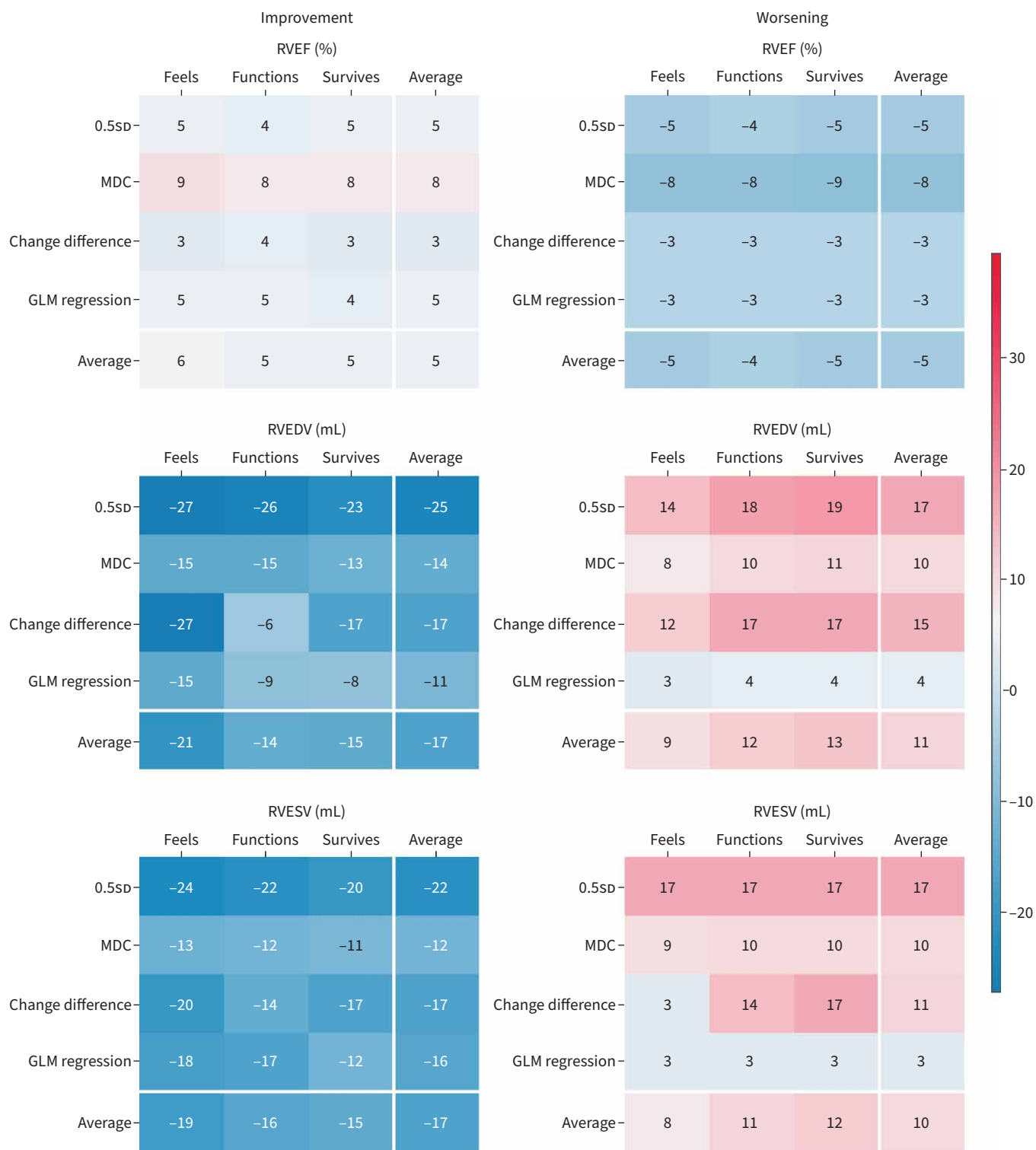


FIGURE 2 Heatmaps of mean minimally important differences (MIDs) for improvement (left panel) and worsening (right panel) for absolute changes in right ventricular (RV) parameters. The heatmaps display the MIDs calculated using four different assessment methods: 0.5SD, minimal detectable change (MDC), change difference and general linear model (GLM) regression (table 1). The values are colour coded, with blue representing lower MIDs and red representing higher MIDs. The white lines separate the average column and row for each parameter. RVEF: RV ejection fraction; RVEDV: RV end-diastolic volume; RVESV: RV end-systolic volume.

compared with those who died. However, patients having follow-up imaging might have been selected because the treating physician felt they were more at risk of deterioration and required additional monitoring; in our institution (Sheffield Pulmonary Vascular Disease Unit, Royal Hallamshire Hospital, Sheffield, UK), CMR imaging is regularly performed as part of routine follow-up in preference to echocardiography. Furthermore, our findings are based on a single-centre cohort and our mortality prediction thresholds are largely exploratory in nature and should be validated in external cohorts. Given the emergence of prospective studies using CMR imaging as a trial end-point and the increasing use of patient-reported outcome measures, consideration should be given to pooling data from such studies to refine our exploratory thresholds for how a patient “feels, functions or survives”. Larger datasets would also allow for the analysis of potential differences based on disease type (IPAH *versus* PAH-CTD) and sex differences [45, 46]. The E-10 was developed in 2014 and therefore patients included from 2010 to 2014 were only assessed with the ISWT. Finally, we have used the ISWT rather than the 6MWT as a measure of exercise capacity. The ISWT has the benefit over the 6MWT in that it is a maximal test and does not have a ceiling effect [23, 47], and thresholds exist for MIDs [28, 29]; however, data are more limited compared with the 6MWT in patients with PAH. Further study of CMR MIDs benchmarked to other measures of exercise capacity including 6MWT distance is required taking into account established MIDs for the 6MWT [46].

Conclusions

We have shown that CMR can identify MIDs for how a patient “feels, functions or survives”. In doing so this study provides further evidence that CMR has the characteristics of a clinical outcome measure. In addition, the findings of this study and the description of MIDs will aid trial size calculations for studies using CMR.

Data generated or analysed during the study are available from the corresponding author by request.

Conflict of interest: A.J. Swift has received research grant funding from Janssen Pharmaceuticals, and consultancy fees from Janssen Pharmaceuticals and General Electric. D.G. Kiely has received grant funding from Ferrer and Janssen Pharmaceuticals, and speaker and consultancy fees and funding for travel from Acceleron, Altavant, Ferrer, Janssen, MSD and United Therapeutics. The other authors have no relationships relevant to the contents of this paper to disclose.

Support statement: This research is partly funded by the National Institute for Health and Care Research (NIHR) Sheffield Biomedical Research Centre (NIHR203321). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. The study was supported by NIHR grant Al_AWARD01706 and by the Wellcome Trust (grant numbers 215799/Z/19/Z and 205188/Z/16/Z to A.J. Swift). The funders did not have any role in the design and conduct of the study; in the collection, analysis, and interpretation of the data; or in the preparation, review and approval of the paper. Funding information for this article has been deposited with the Crossref Funder Registry.

References

- 1 Kiely DG, Elliot CA, Sabroe I, *et al.* Pulmonary hypertension: diagnosis and management. *BMJ* 2013; 346: f2028.
- 2 Goh ZM, Alabed S, Shahin Y, *et al.* Right ventricular adaptation assessed using cardiac magnetic resonance predicts survival in pulmonary arterial hypertension. *JACC Cardiovasc Imaging* 2021; 14: 1271–1272.
- 3 Goh ZM, Balasubramanian N, Alabed S, *et al.* Right ventricular remodelling in pulmonary arterial hypertension predicts treatment response. *Heart* 2022; 108: 1392–1400.
- 4 Kiely DG, Levin D, Hassoun P, *et al.* Statement on imaging and pulmonary hypertension from the Pulmonary Vascular Research Institute (PVRI). *Pulm Circ* 2019; 9: 2045894019841990.
- 5 Alabed S, Shahin Y, Garg P, *et al.* Cardiac-MRI predicts clinical worsening and mortality in pulmonary arterial hypertension: a systematic review and meta-analysis. *JACC Cardiovasc Imaging* 2021; 14: 931–942.
- 6 Lewis RA, Johns CS, Cogliano M, *et al.* Identification of cardiac magnetic resonance imaging thresholds for risk stratification in pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2020; 201: 458–468.
- 7 Bradlow WM, Hughes ML, Keenan NG, *et al.* Measuring the heart in pulmonary arterial hypertension (PAH): implications for trial study size. *J Magn Reson Imaging* 2010; 31: 117–124.
- 8 Hassoun PM, Zamanian RT, Damico R, *et al.* Ambrisentan and tadalafil up-front combination therapy in scleroderma-associated pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2015; 192: 1102–1110.
- 9 Benza RL, Raina A, Gupta H, *et al.* Bosentan-based, treat-to-target therapy in patients with pulmonary arterial hypertension: results from the COMPASS-3 study. *Pulm Circ* 2018; 8: 2045893217741480.
- 10 Vonk Noordegraaf A, Channick R, Cottreel E, *et al.* The REPAIR study: effects of macitentan on RV structure and function in pulmonary arterial hypertension. *JACC Cardiovasc Imaging* 2022; 15: 240–253.

- 11 Swift AJ, Wilson F, Cogliano M, *et al.* Repeatability and sensitivity to change of non-invasive end points in PAH: the RESPIRE study. *Thorax* 2021; 76: 1032–1035.
- 12 van Wolferen SA, van de Veerdonk MC, Mauritz G-J, *et al.* Clinically significant change in stroke volume in pulmonary hypertension. *Chest* 2011; 139: 1003–1009.
- 13 Göransson C, Vejlstrop N, Carlsen J. Clinically important changes in right ventricular volume and function in pulmonary arterial hypertension assessed with cardiac magnetic resonance imaging. *Pulm Circ* 2022; 12: e12097.
- 14 Alabed S, Alandejani F, Dwivedi K, *et al.* Validation of artificial intelligence cardiac MRI measurements: relationship to heart catheterization and mortality prediction. *Radiology* 2022; 305: 68–79.
- 15 Alandejani F, Alabed S, Garg P, *et al.* Training and clinical testing of artificial intelligence derived right atrial cardiovascular magnetic resonance measurements. *J Cardiovasc Magn Reson* 2022; 24: 25.
- 16 Davies RH, Augusto JB, Bhuva A, *et al.* Precision measurement of cardiac structure and function in cardiovascular magnetic resonance using machine learning. *J Cardiovasc Magn Reson* 2022; 24: 16.
- 17 Sitbon O, Gombert-Maitland M, Granton J, *et al.* Clinical trial design and new therapies for pulmonary arterial hypertension. *Eur Respir J* 2019; 53: 1801908.
- 18 Hurdman J, Condliffe R, Elliot CA, *et al.* ASPIRE registry: Assessing the Spectrum of Pulmonary hypertension Identified at a Referral centre. *Eur Respir J* 2012; 39: 945–955.
- 19 Yorke J, Corris P, Gaine S, *et al.* emPHasis-10: development of a health-related quality of life measure in pulmonary hypertension. *Eur Respir J* 2014; 43: 1106–1113.
- 20 Singh SJ, Morgan MD, Scott S, *et al.* Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax* 1992; 47: 1019–1024.
- 21 Benza RL, Gombert-Maitland M, Miller DP, *et al.* The REVEAL Registry risk score calculator in patients newly diagnosed with pulmonary arterial hypertension. *Chest* 2012; 141: 354–362.
- 22 Humbert M, Kovacs G, Hoepfer MM, *et al.* 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. *Eur Respir J* 2023; 61: 2200879.
- 23 Lewis RA, Billings CG, Hurdman JA, *et al.* Maximal exercise testing using the incremental shuttle walking test can be used to risk-stratify patients with pulmonary arterial hypertension. *Ann Am Thorac Soc* 2021; 18: 34–43.
- 24 Revicki D, Hays RD, Cella D, *et al.* Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008; 61: 102–109.
- 25 Guyatt GH, Osoba D, Wu AW, *et al.* Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002; 77: 371–383.
- 26 Lewis R, Armstrong I, Bergbaum C, *et al.* EmPHasis-10 health-related quality of life score predicts outcomes in patients with idiopathic and connective tissue disease-associated pulmonary arterial hypertension: results from a UK multi-centre study. *Eur Respir J* 2020; 56: Suppl. 64, 3973.
- 27 Borgese M, Badesch D, Bull T, *et al.* EmPHasis-10 as a measure of health-related quality of life in pulmonary arterial hypertension: data from PHAR. *Eur Respir J* 2021; 57: 2000414.
- 28 Singh SJ, Puhan MA, Andrianopoulos V, *et al.* An official systematic review of the European Respiratory Society/American Thoracic Society: measurement properties of field walking tests in chronic respiratory disease. *Eur Respir J* 2014; 44: 1447–1478.
- 29 Singh SJ, Jones PW, Evans R, *et al.* Minimum clinically important improvement for the incremental shuttle walking test. *Thorax* 2008; 63: 775–777.
- 30 Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003; 41: 582–592.
- 31 Garratt A, Schmidt L, Mackintosh A, *et al.* Quality of life measurement: bibliographic study of patient assessed health outcome measures. *BMJ* 2002; 324: 1417.
- 32 Vallat R. Pingouin: statistics in Python. *J Open Source Softw* 2018; 3: 1026.
- 33 Davidson-Pilon C, Kalderstam J, Jacobson N, *et al.* CamDavidsonPilon/lifelines: v0.25.9. 2021. <https://zenodo.org/record/4505728> Date last accessed: 26 May 2023.
- 34 Hunter JD. Matplotlib: a 2D graphics environment. *Comput Sci Eng* 2007; 9: 90–95.
- 35 Michelakis ED, Tymchak W, Noga M, *et al.* Long-term treatment with oral sildenafil is safe and improves functional capacity and hemodynamics in patients with pulmonary arterial hypertension. *Circulation* 2003; 108: 2066–2069.
- 36 Roeleveld RJ, Vonk-Noordegraaf A, Marcus JT, *et al.* Effects of epoprostenol on right ventricular hypertrophy and dilatation in pulmonary hypertension. *Chest* 2004; 125: 572–579.
- 37 van de Veerdonk MC, Kind T, Marcus JT, *et al.* Progressive right ventricular dysfunction in patients with pulmonary arterial hypertension responding to therapy. *J Am Coll Cardiol* 2011; 58: 2511–2519.
- 38 van de Veerdonk MC, Huis in t Veld AE, Marcus JT, *et al.* Upfront combination therapy reduces right ventricular volumes in pulmonary arterial hypertension. *Eur Respir J* 2017; 49: 1700007.
- 39 van Wolferen SA, Boonstra A, Marcus JT, *et al.* Right ventricular reverse remodelling after sildenafil in pulmonary arterial hypertension. *Heart* 2006; 92: 1860–1861.

- 40 Peacock AJ, Crawley S, McLure L, *et al.* Changes in right ventricular function measured by cardiac magnetic resonance imaging in patients receiving pulmonary arterial hypertension-targeted therapy: the EURO-MR study. *Circ Cardiovasc Imaging* 2014; 7: 107–114.
- 41 Mouelhi Y, Jouve E, Castelli C, *et al.* How is the minimal clinically important difference established in health-related quality of life instruments? Review of anchors and methods. *Health Qual Life Outcomes* 2020; 18: 136.
- 42 Wilkins MR, Paul GA, Strange JW, *et al.* Sildenafil versus endothelin receptor antagonist for pulmonary hypertension (SERAPH) study. *Am J Respir Crit Care Med* 2005; 171: 1292–1297.
- 43 Billings CG, Lewis R, Hurdman JA, *et al.* The incremental shuttle walk test predicts mortality in non-group 1 pulmonary hypertension: results from the ASPIRE Registry. *Pulm Circ* 2019; 9: 2045894019848649.
- 44 Richeldi L, Ryerson CJ, Lee JS, *et al.* Relative versus absolute change in forced vital capacity in idiopathic pulmonary fibrosis. *Thorax* 2012; 67: 407–411.
- 45 Kawut SM, Lima JAC, Barr RG, *et al.* Sex and race differences in right ventricular structure and function: the Multi-Ethnic Study of Atherosclerosis-Right Ventricle Study. *Circulation* 2011; 123: 2542–2551.
- 46 Mathai SC, Puhan MA, Lam D, *et al.* The minimal important difference in the 6-minute walk test for patients with pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2012; 186: 428–433.
- 47 Billings CG, Hurdman JA, Condliffe R, *et al.* Incremental shuttle walk test distance and autonomic dysfunction predict survival in pulmonary arterial hypertension. *J Heart Lung Transplant* 2017; 36: 871–879.