Anesthesiology and
Perioperative Science

ORIGINAL RESEARCH

Reproducibility of heart rate recovery measures across differing modalities and intensities of submaximal exercise in healthy volunteers: A healthy volunteer study

Haroon Minhas[1*] , Christopher Morton[1], Martin Shaw[1] and Ben Shelley[1,2]

Abstract

Purpose Submaximal exercise tests (SET) are an easier to conduct and less physically demanding alternative to cardiopulmonary exercise testing for assessment of pre-operative fitness. Assessment of heart rate recovery (HRR) following SETs offers an easily obtained marker of vagal reactivation and sympathetic withdrawal and therefore provides an indication of an individual's fitness. This study sought to assess the reproducibility of HRR across different workloads and modalities of SET.

Methods Thirty-four healthy volunteers were recruited to undergo two SETs at 40% and 60% of their predicted maximal workload. A second group of thirty-one healthy volunteers were recruited to undergo three SETs of differing modalities; Cycle ergometry, Step and Shuttle Walk tests. HRR was quantified using the conventional indices HRR1 and HRR2 (calculated as heart rate (HR) on exercise cessation minus HR at 1 and 2 min of rest) and a novel area under the HRR vs. time curve (HRR_{AUC}) method. Reproducibility of results was assessed using intra-class correlation coefficient (ICC) and limits of agreements.

Results The results showed that HRR1 and HRR2 were poorly reproducible across differing workloads and exercise modalities (ICC < 0.45 for all comparisons) whereas HRR_{AUC} proved to be at least moderately reproducible (ICC > 0.52 for all comparisons).

Conclusions These results suggest that HRR_{AUC} may be a superior way of quantifying HRR following SETs, adding objectivity to SET results. Quantifying HRR_{AUC} could prove to have useful clinical applications for pre-operative risk assessment, assessing fitness to undergo treatment and monitoring disease progression.

Keywords Heart rate recovery, Reproducibility, Submaximal exercise testing, Cycle ergometry, Shuttle walk test, Step test

*Correspondence:
Haroon Minhas
haroon.minhas@nhs.scot
[1] Academic Unit of Anesthesia, Perioperative Medicine and Critical Care Research Group, University of Glasgow, Glasgow G4 OSF, UK
[2] Department of Cardiothoracic Anaesthesia, Golden Jubilee National Hospital, Glasgow G81 4DY, UK

1 Introduction

Cardiopulmonary exercise testing (CPET), where patients exercise to peak exertion provides an objective indicator of an individual's aerobic fitness and is commonly advocated for preoperative risk stratification [1, 2]. Whilst CPET is considered the gold standard tool for assessing an individual's fitness; it is not widely available, costly to conduct and the maximal nature may place a potentially unnecessary physiological

Minhas *et al. Anesthesiology and Perioperative Science*   (2024) 2:11

Page 2 of 12

burden on individuals with significant co-morbidities [3]. Submaximal exercise testing (SET) offers a cheaper, less physically demanding and easier to conduct alternative to CPET; during SETs individuals exercise to the point of subjective limitation [3]. Unlike CPET however, there is no widely accepted protocol for conducting and assessing the results of SETs [3]. Heart rate recovery (HRR) is thought to be governed by vagal reactivation and sympathetic withdrawal [4–7]. Imai et al. showed that there are significant differences in HRR between athletes and heart failure patients suggesting a relationship between HRR and exercise capacity [5]. Further studies have confirmed this finding, suggesting that HRR is a surrogate marker of an individual's fitness [4, 7, 8]. Assessment of HRR following SETs could therefore provide an easily obtained indicator of an individual's fitness [4–7] therefore allowing SETs to become widely adopted tool for pre-operative risk assessment, monitoring responses to pre-habilitation and treatment, assessing fitness to undergo treatment and monitoring disease progression.

HRR1, calculated as the difference between heart rate (HR) at cessation of exercise and HR at 1 min of rest (HRR1), is a commonly used method for quantification of HRR. This was first validated by Cole et al. [9] in a population of individuals at risk of coronary artery disease referred for exercise testing. They found that a HRR1 ≤ 12 beats per minute (bpm) after maximal exercise was a strong predictor of mortality in the 6-year follow-up period. HRR1 (following maximal exertion) has since been shown to be a useful predictive and prognostic tool in a variety of clinical populations including as a predictor of poorer postoperative outcomes in patients undergoing non-cardiac surgery [10]. In a multivariable analysis impaired HRR1 response following a 6-min walk test was independently associated with a significantly increased risk of postoperative cardiopulmonary complications in patients undergoing lung resection surgery for lung cancer [11]. A secondary analysis of the measurement of exercise tolerance before surgery study demonstrated an impaired HRR1, in multivariable analysis, was independently associated with a significantly increased risk of myocardial injury in the perioperative period [10]. A recent study comparing HRR following an orthostatic challenge (a markedly 'submaximal' haemodynamic challenge) found that patients who suffered from perioperative myocardial injury had significantly slower HRR as compared with patients who remained free of myocardial injury [12]. Furthermore, in a recent secondary analysis of prospectively collected pre-operative HRR data following six-minute walk test (a submaximal exercise test) in patients undergoing lung cancer resection our group has demonstrated that area under the HRR vs. time curve (HRR$_{AUC}$), a novel method

for quantifying HRR devised by our research team, was significantly elevated in patients who suffered from cardiopulmonary complications (55.2 bpm*min) compared with those without (38.9 bpm*min) (95% confidence interval (CI) 0.13–32.5, $p = 0.048$) [13].

However, there has been significant debate surrounding the reproducibility of HRR following SET [14–16]. Studies have shown that at fixed workloads HRR1 and HRR2 are reproducible measures [14, 16], however when the workloads between tests are varied the reproducibility of HRR decreases [15]. Furthermore studies have also found HRR1 to be significantly different across different modalities of maximal exercise [17, 18].

Since by their nature, the workload for SETs cannot be predetermined and given the submaximal nature of SETs the workload may vary depending on exercise modality. Before testing of the validity of HRR following SETs as a marker of fitness in clinical populations we sought to assess the influence of both exercise intensity and modality on the reproducibility of HRR. We sought to determine the reproducibility of HRR1, HRR2 and a novel area under the HRR vs. time curve (HRR$_{AUC}$) method for quantifying HRR, across different workloads and different modalities of submaximal exercise testing in healthy volunteers. We hypothesized that by including all heart rate data for a full six minutes of recovery, and as a result being influenced less by peak heart rate during exercise, our novel HRR$_{AUC}$ method would be more reproducible than HRR1/2.

2 Methods

We used the STROBE cohort checklist when writing our report [19]. SEARCH I (exercise intensity study) was externally peer reviewed and funded by Medical Research Scotland. SEARCH II (exercise modality study) was externally peer reviewed and funded by the British Journal of Anesthetics and Royal College of Anesthetists. Ethical approval for both studies was granted by the Golden Jubilee Research Institute (GJRI). GJRI reference numbers: 18/ANAES/03 (SEARCH-I) and 19/ANAES/03 (SEARCH-II). Since both studies recruited healthy volunteers, they were exempt from NHS Research Ethics procedures. Written informed consent was obtained from all volunteers in both studies. The studies were conducted in accordance with standards set by the Declaration of Helsinki and in accordance with local statutory requirements [20]. Sample size calculation was not performed in both studies, however the sample size is consistent with recommendations for pilot studies [21].

2.1 Exercise intensity study

Thirty-four healthy volunteers were recruited by convenience sampling method to undergo two SETs by cycle ergometry at two different workloads at the Golden Jubilee

Research Institute, from June 2018 to August 2018. Inclusion criteria were > 18 years old and World Health Organization performance status ≤ 1. Exclusion criteria were; taking β blockers, pregnant or where exercise testing was contraindicated per the American Thoracic Society's contraindications for six-minute walk test [22]. Further details regarding the inclusion and exclusion criteria are shown in Supplementary Table 1.

2.2 Cycle ergometry

The cycle ergometry test followed a modified version of Astrand-Rhyming cycle ergometer test [23]. Volunteers took part in 2 SETs of differing workloads in a randomized order with a minimum of one hour of rest between each test. The workload was varied between each test; 40% and 60% of the volunteer's predicted maximal workload in watts (Wmax). The Wmax was calculated using the following formula [24]:

$$Wmax = \left(\frac{20.4 \ (\text{Height in cm}) - 8.74 \ (\text{Age}) - 288 \ (\text{Sex}) - 1,909 \text{kpm}/\min}{6.116} \right) \times 1.08$$

Volunteers were requested not to consume any caffeine three hours prior to SETs [25]. Prior to the SET the seat position and handlebar position on the cycle ergometer were adjusted such that they were comfortable for the volunteers. Before the SET, the volunteers read a standard written instruction, stating that the protocol and all volunteers underwent a practice test at 20% volunteers Wmax to familiarize them with the controls of the cycle ergometer and protocol. After sitting on the bike at rest to record their baseline HR, volunteers were then instructed to cycle until they achieved and maintained a cadence of 50–60 revolutions per minute. During the first minute of the test volunteers cycled unloaded to warm up. The resistance was then increased to the predetermined target workload over the next two minutes. Volunteers continued to cycle at this effort level for the next three minutes. Volunteers were instructed to cease pedaling at 6-min and to stay stationary for 6 min of rest. Volunteers were asked not to speak during the exercise test or recovery as previous studies have shown this significantly impacts HR variability [4]. If during the test volunteers reached a HR > 85% age predicted maximal HR (HR_{max}) (calculated using the formula from Tanaka et al. [26]) exercise test was terminated.

Volunteers were permitted to terminate the test at any stage. HR data was continuously measured and recorded at 30 s intervals during and after exercise was measured using an electrocardiography (ECG) GE DASH 4000 device and GE CRIKTON DINAMAP device and stored on paper-based case report form.

2.3 Exercise modality study

A second cohort of thirty-one healthy volunteers were recruited by convenience sampling method to undergo three SETs of differing exercise modalities, using the same inclusion and exclusion criteria (Supplementary Table 1) at the Golden Jubilee Research Institute, from January 2020 to March 2020. All volunteers underwent three SETs of differing modalities in a randomized order, with a minimum of one hour of rest between each modality; cycle ergometry, step test and shuttle walk test. The exercise protocol for the Shuttle Walk test was developed by modifying and combining elements from the 6-min Walk Test and Shuttle Walk test [22, 27]. The Step Test followed a modified version of Step Oximetry test [28]. Both protocols were modified to match the modified version of Astrand-Rhyming cycle ergometer test used in the exercise intensity study [22, 23, 28]. Pilot testing and data from exercise intensity study demonstrated that

three minutes of exercise was sufficient to achieve target HR of > 50% HR_{max} across all three modalities of exercise. During the first two minutes of exercise the volunteers were encouraged to exercise until they achieved a target HR of > 50% of their age predicted maximal HR (HR_{max}) calculated using the formula from Tanaka et al. [26]. The target HR of > 50% HR_{max} was selected as a mid-way point between the workloads in exercise intensity study (40% and 60%). Volunteers were then maintained at this HR for one further minute. At the end of the minute the volunteers were asked to stop exercising immediately and stay seated at rest for six minutes. If the volunteers failed to achieve HR of > 50% HR_{max} within the first three minutes, they were encouraged to exercise further until they achieved and maintained a HR of > 50% HR_{max} for one minute. The exercise test was terminated at six minutes, regardless of whether volunteers achieved target HR. If during the test volunteers reached a HR > 85% age predicted maximal HR (HR_{max}) exercise test was terminated. All other factors were kept the same as exercise intensity study. HR was recorded at 30 s intervals throughout the exercise tests and recovery using ECG GE DASH 4000 device and wrist-worn HR monitor (Nonin WristOx$_2$® Model 3150 OEM).

2.4 Cycle ergometry

The Cycle ergometry test was carried out as per the protocol described above at 50% of the volunteer's predicted maximal workload in watts (Wmax) and volunteers were

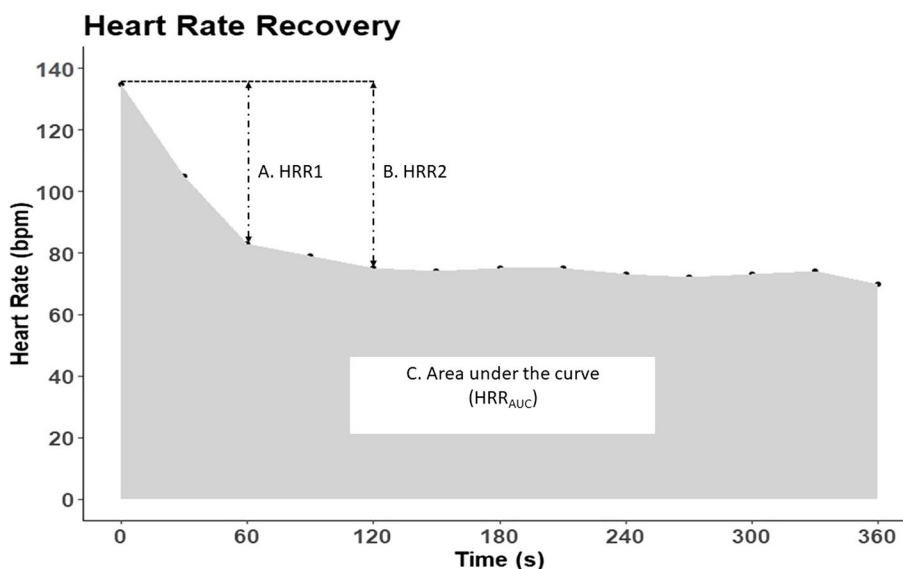**Fig. 1** Methods for quantifying Heart rate recovery (HRR). Example HRR vs time curve; time 0 reflects cessation of exercise. **A** HRR1—Absolute difference between HR at exercise cessation and HR at 1 min of recovery. **B** HRR2—Absolute difference between heart rate (HR) at exercise cessation and HR at 2 min of recovery. **C** Area under the curve (HRR$_{AUC}$)—Area under the HRR vs. Time curve over 360 s of recovery computed using trapezoid method

exercised until they achieved and maintain target HR of > 50% HR$_{max}$ for one minute.

2.5 Step test
The Step test followed a modified version of the Step Oximetry test using a 15 cm step [29]. Volunteers were asked to climb up and down on the step at a self-selected pace but were encouraged to progressively increase their pace until target HR of > 50% HR$_{max}$ was achieved. Volunteers were then asked to maintain this pace to maintain HR of > 50% HR$_{max}$ for one minute.

2.6 Shuttle walk test
The Shuttle Walk test followed a protocol modified version of the the 6-min Walk Test and Shuttle Walk Test [22, 27]. Volunteers were asked to walk back and forth in a 10-m corridor until volunteers achieved and maintained target HR of > 50% HR$_{max}$ for minute. At the end of the test participants were asked to walk back to the start and sit down for six minutes of rest. During exercise HR was monitoring using a wrist worn HR monitor and with ECG during the recovery phase.

2.7 Statistics
HRR was quantified using the following methods, shown in Fig. 1:

A- HRR1—Absolute difference between HR at exercise cessation and HR at 1 min of recovery

B- HRR2—Absolute difference between HR at exercise cessation and HR at 2 min of recovery

C- Area under the curve (HRR$_{AUC}$)—Area under the HRR vs. Time curve over 360 s of recovery computed using trapezoid method

Normality of distribution of data was assessed by Shapiro-Wilk test. Categorical data is presented as counts and percentages and continuous data as mean (± standard deviation) or median [interquartile range] as appropriate to distribution. To ascertain the reproducibility of HRR indices Pearson's correlation coefficient (*r*) and Intraclass correlation coefficient (ICC) (single rater, absolute agreement, 2-way mixed-effects model) were calculated [30]. To assess agreement between HRR indices following exercise at differing intensity and modality, Bland-Altman analysis was conducted. Data analysis was carried out using Rstudio (© 2009 – 2019 Rstudio, Inc.).

3 Results
Thirty-four volunteers were recruited to the exercise intensity study, the median age of the sample was 34 years (22–54) and 15 volunteers (44%) were female. Thirty-two volunteers were recruited to exercise modality study, the median age of the sample was 26 years (21–44) and

Table 1 Demographic information on volunteers in the SEARCH I (exercise intensity) and SEARCH II (exercise modality) studies

Demographics	Search I	Search II
No. of volunteers	34	31
Age	34 [22–54]	26 [21–44]
Sex		
Females	15 (44%)	16 (52%)
Height (cm)	174 (±9.5)	172 (±9.2)
Weight (kg)	75.9 (±14.3)	72.9 (±14.3)
BMI (kg/m^2)	24.8 (±3.1)	24.6 (±3.9)
Smoking history	1 (3%)	5 (16%)
Currently smoking	1 (3%)	2 (7%)
Currently drinking alcohol	34 (100%)	19 (61%)
Units per week	13 [6–15]	2 [0–8]
Hours of exercise per week[a]	4 [3–5]	4 [2–5]
Self-assessment of walking pace		
Slow pace (i.e. < 3mph)	2 (6%)	1 (3%)
Steady average pace	8 (24%)	11 (36%)
Brisk pace	20 (59%)	17 (55%)
Fast pace (i.e. > 4mph)	4 (12%)	2 (%)7
Chronic co-morbidity	5 (15%)	5 (16%)
History of cancer	2 (6%)	1 (3%)
COPD	1 (3%)	1 (3%)
Arterial Hypertension	4 (12%)	3 (10%)
Diabetes Mellitus	1 (3%)	2 (7%)
Medications	8 (24%)	11 (39%)
ACE inhibitor/ARB	3 (9%)	1 (3%)
Calcium channel blocker	4 (12%)	3 (10%)
Alpha blocker	0 (0%)	1 (3%)
Inhaled steroid	1 (3%)	3 (10%)
Inhaled bronchodilator	1 (3%)	2 (7%)
Statin	5 (15%)	1 (3%)
Antiplatelet	2 (6%)	0 (0%)
Other	0 (0%)	14 (45%)
WHO clinical performance status		
0	34 (100%)	30 (97%)
1	0 (0%)	1 (3%)

Categorical variables are given as counts and percentages and continuous variables are given as mean ± standard deviation or median [Interquartile range] as appropriate to distribution.

[a] Volunteers that undertook > 5 h of exercise per week were excluded. *BMI* Body mass index, *ACE* Angiotensin Converting Enzyme, *ARB* Angiotensin II receptor blocker, *COPD* Chronic Obstructive Pulmonary Disease, *WHO* World Health Organization

16 volunteers (52%) were female. Baseline characteristics from both samples are shown in are shown in Table 1. The HR value at peak exercise on the Bike test could not be recorded due to a HR monitor fault in one volunteer and two volunteers did not achieve target HR of > 50% HR_{max} during the Walk test despite encouragement; these subjects were retained in the main analysis, but a sensitivity analysis was conducted by excluding these participants.

Table 2 Intra-class Correlation Coefficient and Pearson's Correlation Coefficient values for HRR models across different workloads (SEARCH-I study (exercise intensity))

HRR Model	40% Wmax vs. 60% Wmax	
	ICC (95% CI)	*r* (*p* values)
HRR1	0.15 (−0.12–0.43) *n* = 33	0.22 (*p* = 0.22) *n* = 33
HRR2	0.35 (−0.1–0.68) *n* = 33	**0.62 (*p* = 0.001)** **n = 33**
HRR$_{AUC}$	**0.58 (0.14–0.8)** **n = 33**	**0.68 (*p* < 0.001)** **n = 33**

Results highlighted in bold *p* < 0.05

ICC Intra-class Correlation Coefficient (single rater, absolute agreement, 2-way mixed-effects model), *r* Pearson's Correlation Coefficient, *HRR1* Heart rate recovery at 1 min, *HRR2* Heart rate recovery at 2 min, *HRR$_{AUC}$* Area under the HRR vs. Time Curve, *CI* Confidence interval

HRR1 and HRR2 were poorly reproducible when workloads were varied between tests (ICC < 0.40 for all comparisons). HRR1 had very weak positive correlation across different workloads ($r = 0.22$, $p = 0.22$), whilst for HRR2, the association was stronger ($r = 0.62$, $p = 0.001$). The HRR$_{AUC}$ method proved to be moderately reproducible across different workloads (ICC 0.58, 95% CI 0.14–0.80) and demonstrated a moderate positive correlation across different workloads ($r = 0.68$, $p < 0.001$), Table 2.

Bland-Altman plots and scatterplots for HRR1, HRR2 and HRR$_{AUC}$ from the exercise intensity study are shown in Fig. 2. Visual inspection of the Bland-Altman plots did not reveal any evidence of heteroscedastic scatter. Bland-Altman analysis of the exercise intensity study revealed a systematic bias such that exercise conducted at 60% Wmax on average estimated a higher value for HRR$_{AUC}$, HRR1 and HRR2 than at 40% Wmax, though this bias was less for the HRR$_{AUC}$ method than HRR1 and HRR2 (mean difference 6.87% versus 18.33% and 29.21% respectively), Fig. 2.

HRR1 and HRR2 were poorly reproducible between different modalities of exercise (ICC < 0.50). HRR1 for Cycle Ergometry vs. Step Test showed a moderate positive association ($r = 0.53$, $p = 0.003$) whilst for the comparisons of Cycle Ergometry vs. Shuttle Walk Test and Step Test vs. Shuttle Walk Test, HRR1 showed a low positive correlation ($r < 0.4$). Sensitivity analysis conducted by excluding individual 2 and 8 showed an improvement in these associations ($r = 0.42$, $p = 0.03$) and ($r = 0.53$, $p = 0.003$) respectively). The HRR$_{AUC}$ method proved to be moderately reproducible between different modalities of exercise, Cycle ergometry vs. Step Test (ICC = 0.63 (0.14–0.84)), Cycle ergometry vs. Shuttle Walk Test (ICC = 0.54 (−0.05–0.81)) and Step Test vs. Shuttle Walk
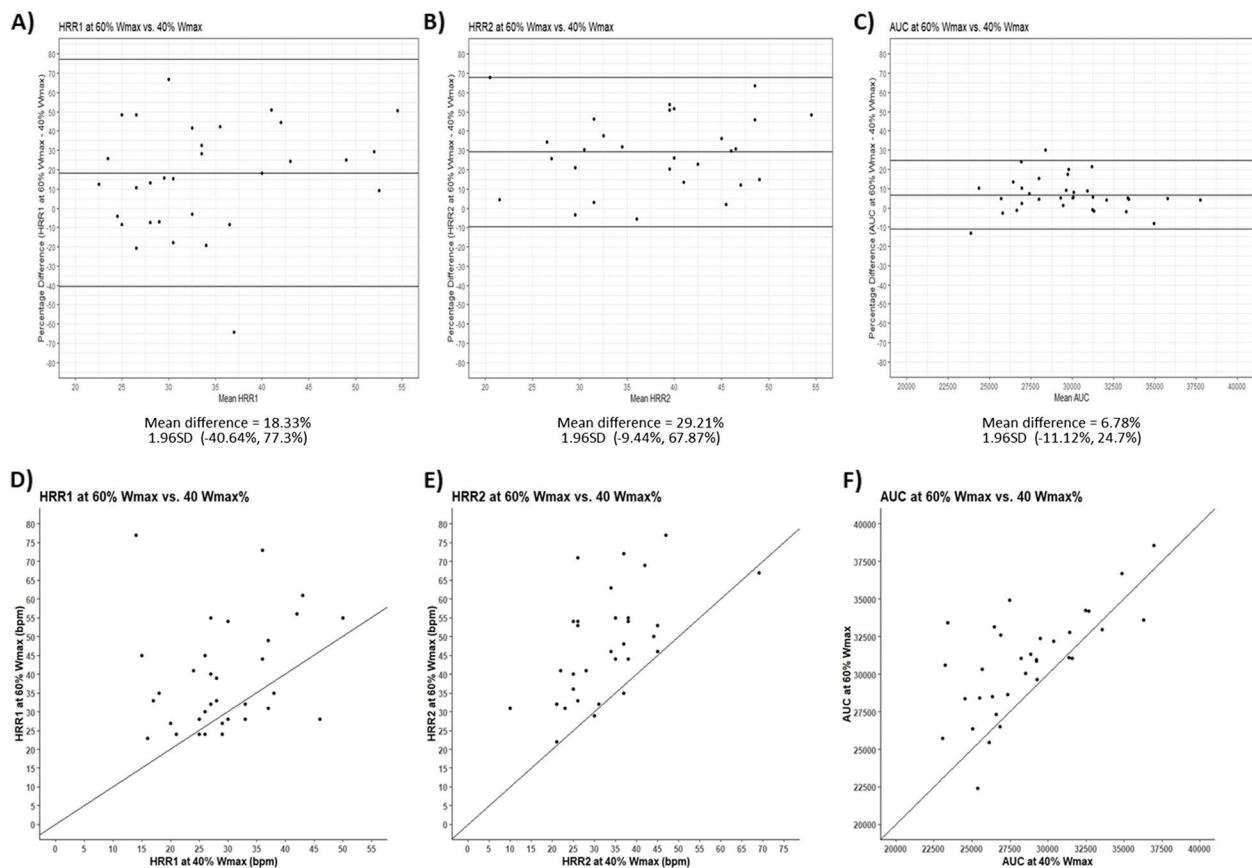
Fig. 2 Bland-Altman plots and Scatterplots of SEARCH I (exercise intensity). Bland-Altman plots and Scatterplots reflecting agreement between HRR values across two different workloads (60% and 40% of calculated maximal workload (Wmax)) for HRR1, HRR2 and HRR$_{AUC}$ values. Bland-Altman plots; **A** HRR1 60% Wmax vs. 40% Wmax, **B** HRR2 60% Wmax vs. 40% Wmax and **C** HRR$_{AUC}$ 60% Wmax vs. 40% Wmax. Scatterplots; **D** HRR1 60% Wmax vs. 40% Wmax, **E** HRR2 60% Wmax vs. 40% Wmax and **F** HRR$_{AUC}$ 60% Wmax vs. 40% Wmax. The reference line represents the line of equality: y = x, bpm = beats per minute

Table 3 Intra-class Correlation Coefficient and Pearson's Correlation Coefficient values for HRR models across exercise modalities (SEARCH- II study (exercise modality))

HRR Model	Cycle Ergometry vs. Step Test		Cycle Ergometry vs. Shuttle Walk		Step Test vs. Shuttle Walk	
	ICC (95% CI)	*r* (*p* values)	ICC (95% CI)	*r* (*p* values)	ICC (95% CI)	*r* (*p* values)
HRR1	**0.45 (0.10–0.70)** *n* = 30	**0.53 (*p* = 0.003)** *n* = 30	0.23 (−0.11–0.53) *n* = 30	0.25 (*p* = 0.18) *n* = 30	0.24 (−0.09–0.54) *n* = 31	**0.39 (*p* = 0.03)** *n* = 31
HRR1[a]			**0.40 (0.05—0.67)** *n* = 28	**0.42 (*p* = 0.03)** *n* = 28	0.35 (−0.08–0.66) *n* = 29	**0.53 (*p* = 0.003)** *n* = 29
HRR2	0.19 (−0.14–0.50) *n* = 30	0.21 (*p* = 0.26) *n* = 30	0.20 (−0.09–0.49) *n* = 30	0.27 (*p* = 0.15) *n* = 30	0.19 (−0.10–0.49) *n* = 31	**0.39 (*p* = 0.03)** *n* = 31
HRR2[a]			0.29 (−0.04–0.58) *n* = 28	0.36 (*p* = 0.06) *n* = 28	0.25 (−0.10–0.58) *n* = 29	**0.47 (*p* = 0.01)** *n* = 29
HRR$_{AUC}$	**0.63 (0.14–0.84)** *n* = 30	**0.75 (*p* < 0.001)** *n* = 30	0.54 (−0.05–0.81) *n* = 30	**0.76 (*p* < 0.001)** *n* = 30	**0.81 (0.63–0.91)** *n* = 31	**0.84 (*p* < 0.001)** *n* = 31
HRR$_{AUC}$[a]			0.52 (−0.05–0.80) *n* = 28	**0.75 (*p* < 0.001)** *n* = 28	**0.80 (0.61–0.90)** *n* = 29	**0.83 (*p* < 0.001)** *n* = 29

Results highlighted in bold *p* < 0.05

[a] Results from sensitivity analysis: conducted by excluding 2 volunteers who did not achieve target HR on the Shuttle Walk test
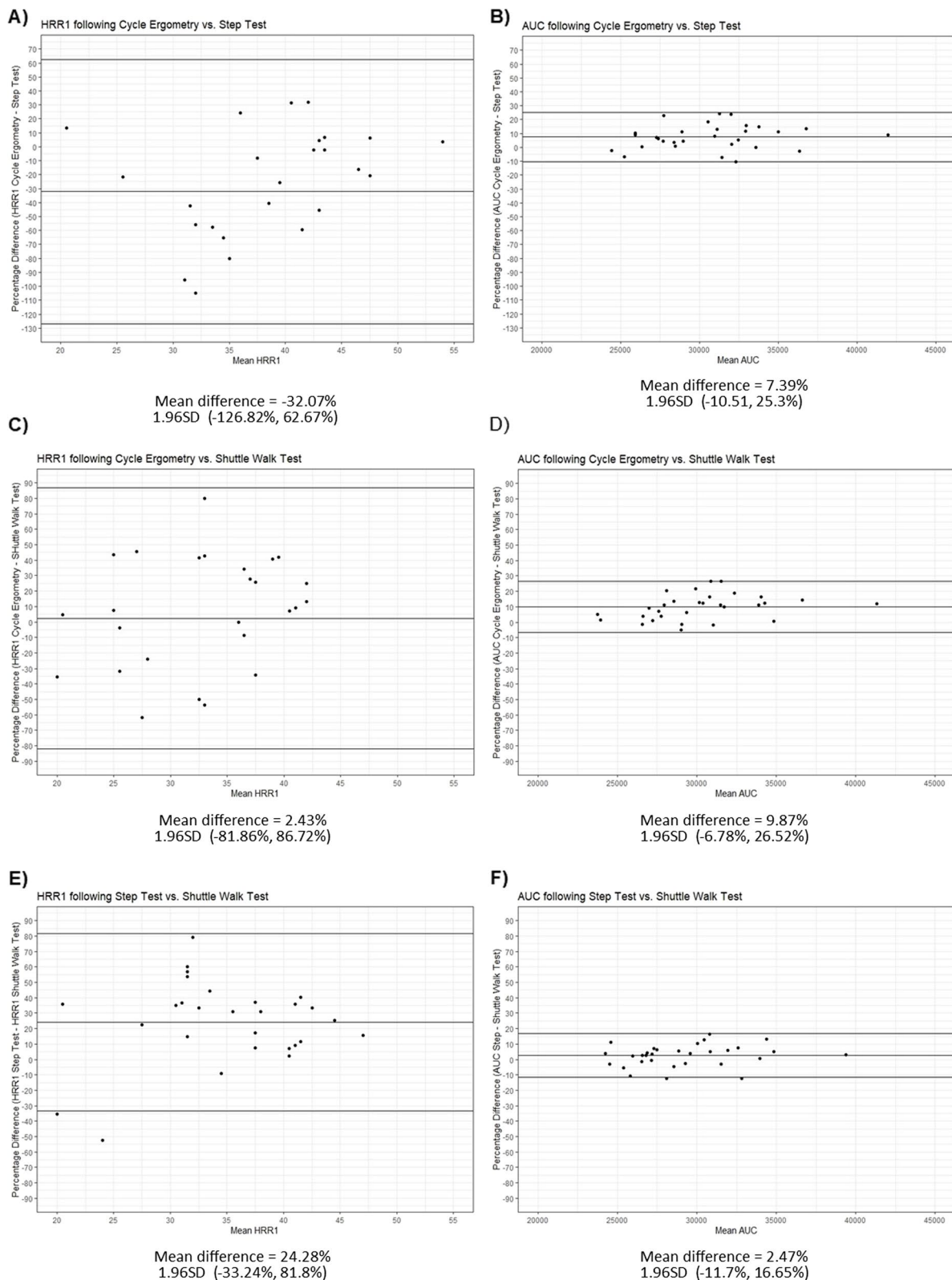
Fig. 3 Bland-Altman plots SEARCH-II (exercise modality) for HRR1and HRR$_{AUC}$. Bland-Altman plots reflecting percentage difference between HRR values across three different exercise modalities (Cycle ergometry, Step test and Shuttle walk test) HRR1 and HRR$_{AUC}$. Bland-Altman plots; **A** HRR1 Cycle ergometry vs. Step Test, **B** HRR$_{AUC}$ Cycle ergometry vs. Step Test, **C** HRR1 Cycle ergometry vs. Shuttle walk test, **D** HRR$_{AUC}$ Cycle ergometry vs. Shuttle walk test, E) HRR1 Step Test vs. Shuttle walk test and F) HRR$_{AUC}$ Step Test vs. Shuttle walk test
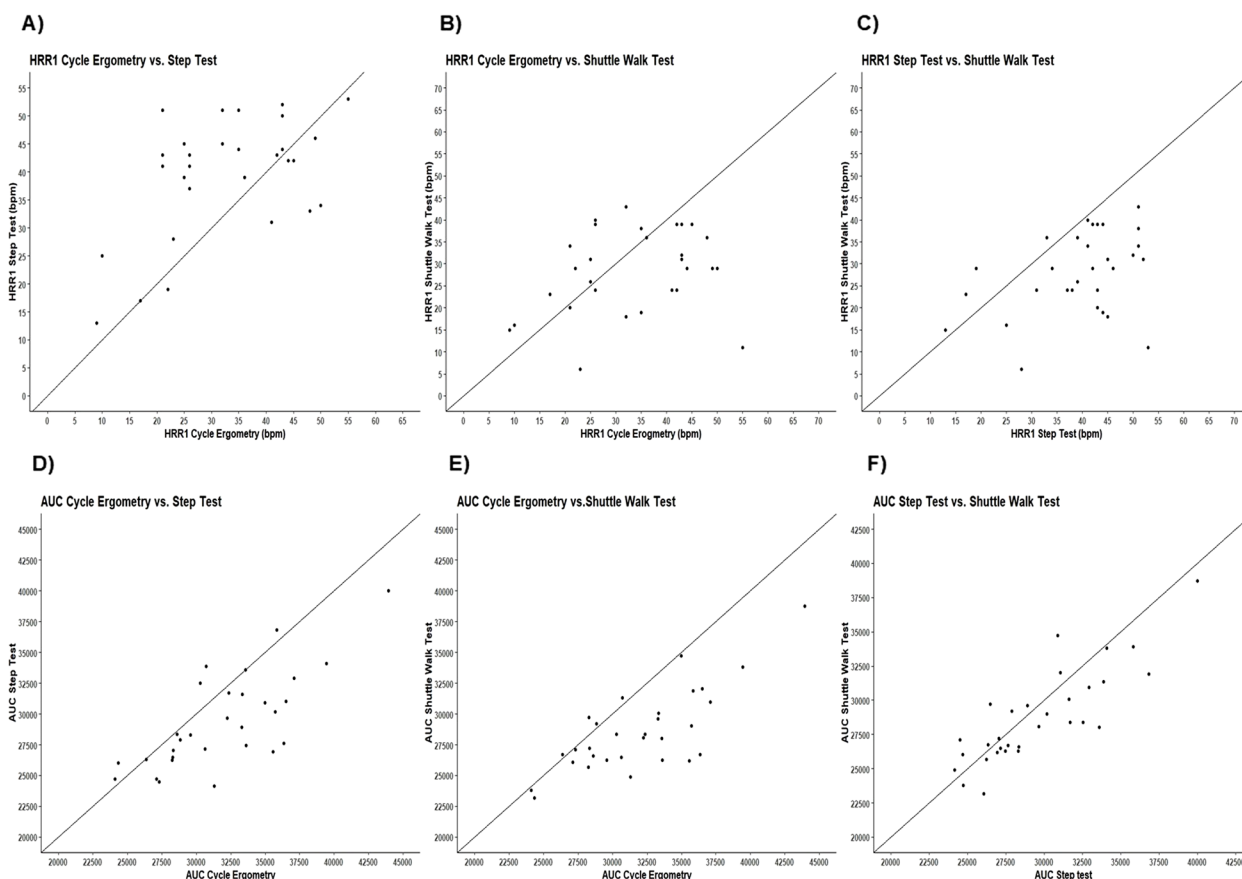
Fig. 4 Scatterplots SEARCH-II (exercise modality) for HRR1 and HRR$_{AUC}$. Scatter plots reflecting agreement between HRR values across three different exercise modalities (Cycle ergometry, Step test and Shuttle walk test) HRR1 and HRR$_{AUC}$ Scatter plots: **A** HRR1 Cycle ergometry vs. Step Test, **B** HRR1 Cycle ergometry vs. Shuttle walk test, **C** HRR1 Step Test vs. Shuttle walk test, **D** HRR$_{AUC}$ Cycle ergometry vs. Step Test, **E** HRR$_{AUC}$ Cycle ergometry vs. Shuttle walk test, and **F**) HRR$_{AUC}$ Step Test vs. Shuttle walk test. *The reference line represents the line of equality: y = x, bpm = beats per minute*

Test (ICC = 0.80 (0.61–0.90)). The HRR$_{AUC}$ method demonstrated very strong positive correlation across all comparisons ($r \geq 0.75$), in Table 3.

Bland-Altman plots and scatter plots for HRR1, HRR2 and HRR$_{AUC}$ from the exercise modality study are shown in, Figs. 3, 4 and 5 respectively. Visual inspection of the Bland-Altman plots did not reveal any evidence of heteroscedastic scatter. Bland-Altman analysis of HRR1 from the exercise modality study revealed a systemic bias such that the Step Test, on average estimated a higher value for HRR1 across both comparisons involving the Step Test (mean difference of 32.07% for Step Test vs. Cycle Ergometry and 24.28% for Step Test vs. Shuttle Walk Test), Fig. 3. Bland–Altman analysis of HRR$_{AUC}$ also revealed a systemic bias such that the Cycle ergometry test on average estimated a higher value for HRR$_{AUC}$ for the same person across both comparisons involving the Cycle Ergometry though this bias was less for the HRR$_{AUC}$ method than HRR1 (mean difference of 7.29% for Cycle Ergometry vs. Step Test and 2.47% for Cycle Ergometry vs. Shuttle Walk Test), Fig. 3. The Bland-Altman and scatterplots for HRR2 are shown in Fig. 5 and demonstrate a systemic bias such that the Step Test, on average estimated a higher value for HRR2 across both comparisons involving the Step Test (mean difference of 25.05% for Step Test vs. Cycle Ergometry and 30.32% for Step Test vs. Shuttle Walk Test), Fig. 5.

4 Discussion
The main findings of these studies are that the novel HRR$_{AUC}$ method for quantifying HRR following SETs demonstrates moderate reproducibility across different workloads and different modalities of exercise where HRR1 and HRR2 are poorly reproducible. This suggests that the novel HRR$_{AUC}$ may be a more objective method for quantifying HRR following SET.
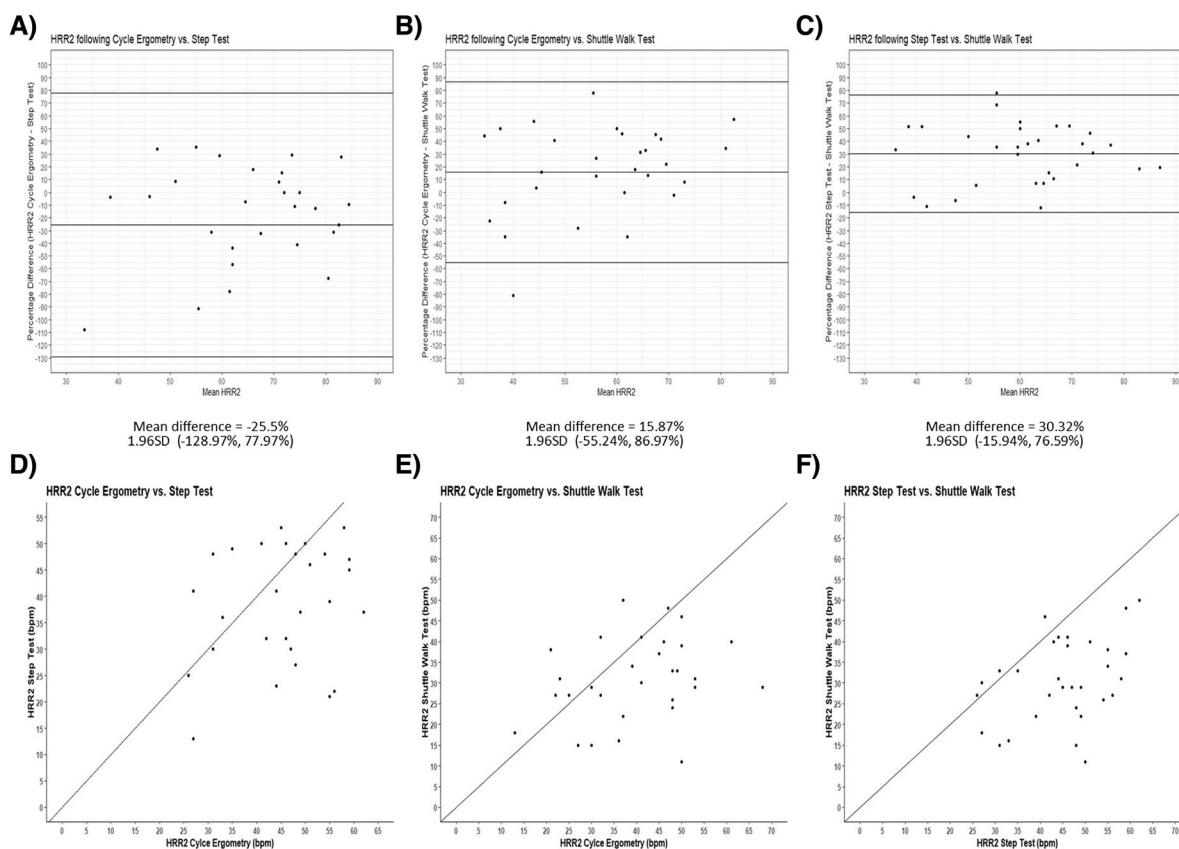
Fig. 5 Bland-Altman plots and Scatterplots of SEARCH II (exercise modality) for HRR2. Bland-Altman plots and Scatterplots reflecting agreement between HRR2 values across three different exercise modalities (Cycle ergometry, Step test and Shuttle walk test). Bland-Altman Plots: **A** HRR2 Cycle ergometry vs. Step Test, **B** HRR2 Cycle ergometry vs. Shuttle walk test, **C** HRR2 Step Test vs. Shuttle walk test. Scatterplots: **D** HRR2 Cycle ergometry vs. Step Test, **E** HRR2 Cycle ergometry vs. Shuttle walk test, and **F** HRR2 Step Test vs. Shuttle walk test. *The reference line represents the line of equality: y = x, bpm = beats per minute*

Previous studies have shown that at fixed workloads HRR1 and HRR2 are reproducible measures. Arduini et al. [14] recruited 21 healthy individuals to undergo a set of cycle ergometer submaximal exercise tests at 65% and 80% age predicted maximum heart rate followed by passive recovery; this was then repeated 2 weeks later. These authors demonstrated that at fixed workloads HRR1 and HRR2 were reproducible measures between individuals at each workload (mean ICC between trial 1 and trial 2: 0.827 and 0.747 respectively), but no comparison was made of HRR between the different workloads. Similarly, Orini et al. [15] repeated exercise stress tests at 50% of the absolute predicted maximum workload in 1187 individuals 34.2 ± 2.8 months apart and found that at fixed workloads (when the difference in workload between tests was ≤ 10 watts) HRR indices were reproducible. However, they found that in participants where the difference in workload between tests was $> 10W$, this negatively impacted the reproducibility of HRR1 and HR response during exercise, suggesting that both HR

response to exercise and HRR may be effort dependent. The results of the current study are in keeping with these findings and demonstrate that when workloads are varied (by 20W between cycle ergometry tests), HRR1 and HRR2 are poorly reproducible. The novel HRR_{AUC} method however appears to cope better with this difference in workload and demonstrates moderate reproducibility across different workloads (ICC 0.58, 95% CI (0.14–0.8)). This is further emphasized by the Bland-Altman analysis which shows that there is poor agreement with very wide limits of agreement between HRR1 and HRR2 values when the workload is varied. HRR_{AUC} however demonstrates better agreement and narrower limits of agreement with changing workload.

Whilst it has been widely accepted that in healthy individuals, the time constant of heart rate decay in the first 30 s of recovery (early phase of recovery) is almost exclusively a function of vagal reactivation [5], there is evidence to challenge this philosophy which suggests that early phase of recovery is instead due to a combination

of vagal reactivation and sympathetic withdrawal [6, 7]. Pierpont et al. [6], modelled the relative contributions of the sympathetic and parasympathetic nervous system on HRR following maximal and submaximal exercise. This demonstrated that the relative contributions of each arm of the autonomic nervous system were dependent on the intensity of exercise. Higher intensities of exercise (which lead to increased sympathetic activation during exercise) were associated with increased contribution of sympathetic withdrawal to HRR in the early phase of recovery, whilst at lesser exercise intensities (which led to lesser sympathetic activation during exercise) were associated with lesser contribution of sympathetic withdrawal to HRR in the early phase of recovery. Gourine et al. [4], suggested a similar physiology for HR during exercise; increasing workload during exercise is associated with increased sympathetic contribution to HR control. They also suggested that there is a threshold of HR beyond which rises in HR during exercise are governed exclusively by sympathetic activation. This suggests, as is perhaps intuitive, the level of sympathetic activation during exercise is dependent on the intensity of exercise. Subsequently the degree of sympathetic withdrawal during recovery depends on the degree of sympathetic activation during exercise. From these findings, it is reasonable to hypothesize that since different individuals would exercise to different intensities during SETs, they would have different levels of sympathetic activation during exercise and in turn will have different levels of sympathetic withdrawal during recovery. The HRR response following SETs would therefore naturally be dependent on exercise intensity (i.e., effort dependent). We hypothesized that since HRR1/2 ignore the HRR data beyond the first 2 min of recovery and may miss key information contained within the HRR curve regarding an individual's HRR dynamics and autonomic innervation beyond two minutes. Visual inspection of the HRR vs. time plots revealed that in many volunteers' HR had not plateaued until 4 or more minutes into recovery. Further, by quantifying the dynamics of HRR for the entire 6 min of recovery, the HRR_{AUC} is likely to be less influenced by peak HR which in turn is influenced by peak workload during exercise.

Previous studies have documented that HRR1 differs significantly between different modes of maximal exercise. Maeder et al. [17] found that HRR1 values obtained following maximal cycling were significantly higher than following maximal treadmill exercise, both in healthy subjects and in heart failure patients. In their study, peak oxygen consumption (VO_2) obtained during the treadmill test was significantly higher than that obtained during cycle ergometry despite similar peak HR, suggesting increased work associated with the treadmill

test. The authors hypothesized that exercises involving greater muscle mass (reflected by a higher VO_2), require a higher heart rate during the early phase of recovery to maintain cardiac output and allow clearance of metabolites that have accumulated in these muscles; thus, resulting in a slower rate of HRR. Randive et al. [18] found similar results; HRR1 was significantly lower during leg ergometry as compared with arm ergometry (where peak VO_2 was lower) despite similar peak HR between both tests. The current study also found HRR1 and HRR2 to be poorly reproducible across different exercise modalities despite similar peak HR; this may reflect the differences in muscle use between the differing exercise tests, in keeping with the hypothesis of Maeder et al. Moreover, our study also found HRR_{AUC} to be highly reproducible between Shuttle walk vs. Step test comparison, given that these two exercises involve greater muscle mass than cycle ergometry this adds further evidence to the hypothesis that muscle mass involvement may also influence the early phase of heart rate recovery. Despite this however the novel HRR_{AUC} method demonstrates moderate reproducibility across all three comparisons. We hypothesize, by including HRR data beyond the early phase the HRR_{AUC} method proves to be less influenced by factors impacting early phase of recovery such as greater muscle mass and therefore demonstrates moderate reproducibility despite differing exercise modality and therefore muscle mass involvement. Lastly, the Bland-Altman analysis also revealed poor agreement and very wide limits of agreement for HRR1 and HRR2 across all comparisons. In comparison the HRR_{AUC} method demonstrates better agreement between different exercise modalities with mean percentage difference < 10% across all comparisons. Furthermore, the limits of agreement are much narrower for the HRR_{AUC} across all three comparisons which is emphasized by visual inspection of the Bland–Altman plots in Fig. 3.

Two volunteers (volunteer 2 and 8) did not achieve target HR > 50% of HR_{max} on the Shuttle Walk test; both of whom were young and comparatively fit suggesting that the Shuttle Walk test may not be sufficient work for younger fitter individuals and hence an inappropriate test in this demographic. Sensitivity analysis conducted by excluding these individuals showed an increase in reproducibility for both HRR1 and HRR2, but interestingly did not affect the reproducibility of HRR_{AUC}. This suggests that the HRR_{AUC} may require a much lower minimum peak work to yield reproducible results. Further work is however required to establish the minimum peak effort required to yield reproducible HRR responses and consolidate this finding.

Whilst these studies identify important novel findings for HRR following SETs, they are not without their

limitations. These studies included mostly young healthy individuals which limits the extrapolation of results to an older age group; this simply reflects the ease of recruiting students as healthy volunteers to take part in an exercise study. The study excluded volunteers taking beta blockers as this could potentially affect heart rate during test however a large proportion of participants were consuming alcohol, and a small proportion of the participants were on calcium channel blockers which may have affected the heart rate, though this likely reflects the clinical population in which HRR may be used. Standardization of workloads between exercise modalities may have increased reproducibility, however this would be technically challenging without expensive exercise equipment. We did however aim to maintain volunteers at HR_{max} at peak exercise between 50%–85% of the volunteers age predicted maximum HR. Furthermore, to be widely adopted as an easily obtained indicator of fitness, SETs need to be able to be conducted without expensive exercise equipment which is not as widely available. The duration of exercise between modalities of exercise also varied which may have been another potential confounder but all volunteers achieved and maintained target HR (HR > 50% of age predicted HR_{max}) before cessation of exercise. Lastly, whilst the results of this study show that HRR_{AUC} method has better reproducibility across differing workload and exercise modalities as compared with HRR1 and HRR2, there is currently limited evidence supporting the utility of HRR_{AUC} method as for assessing pre-operative fitness in individuals highlighting the need for further studies. Further research is also required to establish the effect of exercise time on the reproducibility of HRR and the minimum peak effort required to produce a HRR response.

5 Conclusion

The current study demonstrates that HRR_{AUC} is a moderately reproducible measure across different workloads and modalities of exercise and therefore, may be a superior method to traditional methods for quantifying HRR following SETs. There is currently limited evidence supporting the utility of HRR_{AUC} method as for assessing pre-operative fitness in individuals highlighting the need for further studies. Further research is also required to establish the effect of exercise time on the reproducibility of HRR and the minimum peak effort required to produce a HRR response. This may allow HRR to be more widely utilized across various clinical populations for pre-operative risk assessment, monitoring responses to pre-habilitation and to treatment, assessing fitness to undergo treatment and monitoring disease progression.

Abbreviations

CI	Confidence interval
CPET	Cardiopulmonary exercise testing
ECG	Electrocardiography
HR	Heart rate
HR_{max}	Age predicted maximum heart rate
HRR	Heart rate recovery
HRR1	Heart rate recovery 1 min (Absolute difference between heart rate at exercise cessation and heart rate at 1 min of recovery)
HRR2	Heart rate recovery 2 min (Absolute difference between heart rate at exercise cessation and HR at 2 min of recovery)
HRR_{AUC}	Area under the HRR vs. Time curve over 360 s of recovery computed using trapezoid method
ICC	Intraclass correlation coefficient
SET	Submaximal exercise test(s)

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s44254-023-00042-2.

> **Additional file 1: Supplementary Table 1.** Exclusion criteria as per the American Thoracic Society's contraindications for exercise testing.

Authors' contributions
All experiments were performed at the Golden Jubilee Research Institute. Conception and design of the work: H.M., C.M., M.S. and B.S. Data acquisition: H.M. and C. M. Analysis and interpretation: H.M., C.M., M.S. and B.S. Authors H.M., C.M., M.S. and B.S. contributed to the drafting of work and the revision for critically important content. H.M., C.M., M.S. and B.S. approved the final version of the manuscript and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Availability of data and materials
Data that supports the findings of this manuscript is available from corresponding author upon reasonable request subject to application to Golden Jubilee National Hospital Research and Development and in line with the Golden Jubilee Research Institute data sharing policies.

Declarations

Ethics approval and consent to participate
Ethical approval for both studies in this article was granted by the Golden Jubilee Research Institute (GJRI). GJRI reference numbers: 18/ANAES/03 (SEARCH-I) and 19/ANAES/03 (SEARCH-II). Since both studies recruited healthy volunteers, they were exempt from NHS Research Ethics procedures. Written informed consent was obtained from all volunteers in both studies. The studies were conducted in accordance with standards set by the Declaration of Helsinki and in accordance with local statutory requirements.

Consent for publication
All authors gave their content for publication.

Competing interests
The authors of this publication have no competing interests.

References

1. Pritchard A, Burns P, Correia J, Jamieson P, Moxon P, Purvis J, et al. ARTP statement on cardiopulmonary exercise testing 2021. BMJ Open Respir Res. 2021;8(1):e001121.
2. Kristensen SD, Knuuti J, Saraste A, Anker S, Bøtker HE, De Hert S, et al. 2014 ESC/ESA guidelines on non-cardiac surgery: cardiovascular assessment and management: the joint task force on non-cardiac surgery: cardiovascular assessment and management of the European Society of Cardiology (ESC) and the european Society of Anaesth. Eur Heart J. 2014;35(35):2383–431.
3. Noonan V, Dean E. Submaximal exercise testing: clinical application and interpretation. Phys Ther. 2000;80(8):782–807.
4. Gourine AV, Ackland GL. Cardiac vagus and exercise. Physiology. 2019;34(1):71–80.
5. Imai K, Sato H, Hori M, Kusuoka H, Ozaki H, Yokoyama H, et al. Vagally mediated heart-rate recovery after exercise is accelerated in athletes but blunted in patients with chronic heart-failure. J Am Coll Cardiol. 1994;24(6):1529–35.
6. Pierpont GL, Adabag S, Yannopoulos D. Pathophysiology of exercise heart rate recovery: a comprehensive analysis. Ann Noninvasive Electrocardiol. 2013;18(2):107–17.
7. Michael S, Graham KS, Davis GM. Cardiac autonomic responses during exercise and post-exercise recovery using heart rate variability and systolic time intervals—a review. Front Physiol. 2017;8:301.
8. Machhada A, Trapp S, Marina N, Stephens RCM, Whittle J, Lythgoe MF, et al. Vagal determinants of exercise capacity. Nat Commun. 2017;8:1–7.
9. Cole CR, Blackstone EH, Pashkow FJ, Snader CE, Lauer MS. Heart-rate recovery immediately after exercise as a predictor of mortality. N Engl J Med. 1999;341(18):1351–7.
10. Ackland GL, Abbott TEF, Minto G, Clark M, Owen T, Prabhu P, et al. Heart rate recovery and morbidity after noncardiac surgery: planned secondary analysis of two prospective, multi-centre, blinded observational studies. PLoS ONE. 2019;14(8):e0221277.
11. Ha D, Choi H, Zell K, Raymond DP, Stephans K, Wang XF, et al. Association of impaired heart rate recovery with cardiopulmonary complications after lung cancer resection surgery. J Thorac Cardiovasc Surg. 2014;149(4):1168–73.
12. May SM, Reyes A, Martir G, Reynolds J, Paredes LG, Karmali S, et al. Acquired loss of cardiac vagal activity is associated with myocardial injury in patients undergoing noncardiac surgery: prospective observational mechanistic cohort study. Br J Anaesth. 2019;123(6):758–67.
13. Ismahel H, Hughes C, Lafferty B, Shelley B. Prediction of postoperative cardiopulmonary complications via assessment of heart rate recovery after submaximal exercise testing. Anaesthesia, 2023;78(10):1295–7.
14. Arduini A, Gomez-Cabrera MC, Romagnoli M. Reliability of different models to assess heart rate recovery after submaximal bicycle exercise. J Sci Med Sport. 2011;14(4):352–7.
15. Orini M, Tinker A, Munroe PB, Lambiase PD. Long-term intra-individual reproducibility of heart rate dynamics during exercise and recovery in the UK Biobank cohort. PLoS One. 2017;12(9): e0183732.
16. Fecchio RY, Brito L, Leicht AS, Forjaz CLM, Peçanha T. Reproducibility of post-exercise heart rate recovery indices: a systematic review. Auton Neurosci. 2019;221:102582.
17. Maeder MT, Ammann P, Rickli H, Brunner-La Rocca HP. Impact of the exercise mode on heart rate recovery after maximal exercise. Eur J Appl Physiol. 2009;105(2):247–55.
18. Ranadive SM, Fahs CA, Yan H, Rossow LM, Agliovlastis S, Fernhall B. Heart rate recovery following maximal arm and leg-ergometry. Clin Auton Res. 2011;21(2):117–20.
19. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007;335(7624):806-8.
20. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013;310(20):2191-4.
21. In J. Introduction of a pilot study. Korean J Anesthesiol. 2017;70(6):601-5.
22. American Thoracic Society. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med. 2003. 2003;166(1):111–7.
23. Test A ryhming A r CE, Test T, Max EVO. Cycle ergometer testing. 2014.
24. Jones NL, Makrides L, Hitchcock C, Chypchar T, McCartney N. Normal standards for an incremental progressive cycle ergometer test. Am Rev Respir Dis. 1985;131(5):700–8.
25. Arnaud MJ. The pharmacology of caffeine. Prog Drug Res. 1987;31:273–313.
26. Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. J Am Coll Cardiol. 2001;37(1):153–6.
27. Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE. Development of a shuttle walking test of disability in patients with chronic airways obstruction. Thorax. 1992;47(12):1019–24
28. Fox BD, Sheffy N, Vainshelboim B, Fuks L, Kramer MR. Step oximetry test: a validation study. 2018;5(1):e000320.
29. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63.
30. Buchheit M, Laursen PB, Ahmaidi S. Parasympathetic reactivation after repeated sprint exercise. Am J Physiol-Heart Circul Physiol. 2007;293(1):H133–41.

Publisher's Note