Canadian Science Publishing | Applied Physiology, Nutrition, and Metal

Test-retest reliability of cardiopulmonary exercise testing in patients with atrial fibrillation and determination of exercise responders to high-intensity interval training and moderate-to-vigorous intensity continuous training

Journal:	*Applied Physiology, Nutrition, and Metabolism*
Manuscript ID	Draft
Manuscript Type:	Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Terada, Tasuku; University of Nottingham, School of Life Sciences Keir, Daniel; University of Western Ontario, Kinesiology Murias, Juan; Hamad Bin Khalifa University, Faculty of Kinesiology Vidal-Almela, Sol; University of Ottawa Heart Institute, Exercise Physiology and Cardiovascular Health Lab; Institut du Savoir Montfort, Buckley, John; Keele University Reed, Jennifer; University of Ottawa Heart Institute; University of Ottawa
Is the manuscript for consideration in a Special Issue or Collection?:	Not applicable (regular submission)
Keyword:	aerobic evaluation < exercise, exercise intensity < exercise, exercise performance < exercise, exercise physiology < exercise, aerobic exercise < exercise, exercise testing < exercise, exercise training < exercise

Title: Test-retest reliability of cardiopulmonary exercise testing in patients with atrial fibrillation and determination of exercise responders to high-intensity interval training and moderate-to-vigorous intensity continuous training

Running Title: Reliability of CPET in Afib

Authors: Tasuku Terada,[1,2] Daniel A. Keir,[3,4,5] Juan M. Murias,[6] Sol Vidal-Almela,[2] John Buckley,[7] Jennifer L. Reed,[2,8,9]

Affiliations:

[1]School of Life Sciences, Division of Physiology, Pharmacology, and Neuroscience; University of Nottingham, Nottingham, UK

[2]Exercise Physiology and Cardiovascular Health Lab, Division of Cardiac Prevention and Rehabilitation, University of Ottawa Heart Institute, Ontario, Canada

[3]School of Kinesiology, The University of Western Ontario, Ontario, Canada

[4]Lawson Health Research Institute, London, Ontario, Canada

[5]Cardiac Rehabilitation and Secondary Prevention Program, St. Joseph's Health Care, Ontario, Canada

[6]College of Health and Life Sciences, Hamad Bin Khalifa University, Doha, Qatar

[7]School of Allied Health Professions, Keele University, Stafford, UK

[8]School of Human Kinetics, Faculty of Health Sciences, University of Ottawa, Ontario, Canada

[9]Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada

Corresponding Author:

Jennifer L. Reed, RKin, PhD

Exercise Physiology and Cardiovascular Health Lab, Division of Cardiac Prevention and Rehabilitation, University of Ottawa Heart Institute, Ontario, Canada

E-mail: JReed@ottawaheart.ca

Telephone: 1-613-696-7392

Twitter: @DrJenniferReed

Abstract: 250 words

All element word counts: 6041 words

Tables/Figures: 3 figures and 1 table

References: 40

ABSTRACT

Introduction: Disabling atrial fibrillation (AF)-related symptoms (e.g., palpitations and dyspnea) may influence day-to-day cardiopulmonary exercise testing (CPET) performance, which can affect exercise prescription for high-intensity interval training (HIIT) and moderate-to-vigorous intensity continuous training (M-VICT) and their outcomes. This study examined the test-retest reliability of CPET in patients with AF and assessed the proportion of participants achieving minimal detectable changes (MDC) in peak oxygen consumption ($\dot{V}O_{2peak}$) following HIIT and M-VICT. **Methods**: Participants were randomized into HIIT or M-VICT after completing two baseline CPETs: one with cardiac stress technologists ($CPET_{diag}$) and the other with a research team of exercise specialists ($CPET_{research}$). Additional CPET was completed following 12 weeks of twice-weekly training. Reliability of $CPET_{diag}$ and $CPET_{research}$ was assessed by intraclass correlation coefficient (ICC) and dependent t-tests. The MDC score was calculated for $\dot{V}O_{2peak}$ using a reliable change index. The proportion of participants achieving MDC was compared between HIIT and M-VICT using chi-square analysis. **Results**: Eighteen participants (69±7 years, 33% females) completed two baseline CPETs. ICC was significant for all measured variables. However, peak power output (PO_{peak}: 124±40 vs. 148±40 watts, $p<0.001$) and HR (HR_{peak}: 136±22 vs. 148±30 bpm, $p=0.023$) were significantly greater in $CPET_{research}$ than $CPET_{diag}$. Few participants achieved MDC in $\dot{V}O_{2peak}$ (5.6 mL/kg/min) with no difference between HIIT (0%) and M-VICT (10.0%, $p=0.244$). **Conclusions**: PO_{peak} and HR_{peak} differed significantly in patients with AF when CPETs were repeated under different settings. Caution must be practiced when prescribing exercise intensity based on these measures as under-prescription may increase the number of exercise non-responders. **Key words**: cardiopulmonary fitness, arrhythmia, peak oxygen consumption, cardiovascular rehabilitation

Introduction

Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia(Dai et al., 2021) characterized by a rapid and irregular heartbeat that is associated with disabling symptoms, such as palpitations, dyspnea, and dizziness (Lip et al., 2016). AF increases the risk of stroke by five-fold and the risk of all-cause mortality by nearly two-fold (Ruddox et al., 2017, Wolf et al., 1991). As AF progresses from paroxysmal to persistent and permanent forms, the risks of thromboembolism (Chiang et al., 2012) and mortality (Mentel et al., 2020, Shukla and Curtis, 2014) increase. The age-adjusted prevalence of AF approximately quadrupled from 20 to 96 per thousand in males and 14 to 49 per thousand in females, respectively, between 1998 and 2007 (Schnabel et al., 2015). In 2017, AF afflicted more than 37 million people globally (Dai et al., 2021). With the rise in AF incidence and prevalence, it meets the criteria for a 21st century cardiovascular disease epidemic (Kornej et al., 2020).

Low aerobic fitness in patients with AF (Andrade et al., 2020, Mertens and Kavanagh, 1996) may contribute to the high risk of morbidity and mortality (Kokkinos et al., 2023). Indeed, in patients with AF, an increase in one metabolic equivalent (MET, a measure of aerobic fitness) is associated with a 12% lower risk of morbidity and 15% lower risk of all-cause mortality (Garnvik et al., 2020), and increases of 2.0-3.0 METs are associated with approximately 31% lower mortality risks (Kokkinos et al., 2023). Although exercise training in the form of moderate-to-vigorous intensity continuous training (M-VICT) increases aerobic fitness (American College of Sports Medicine, 2021), growing evidence indicates that high-intensity interval training (HIIT) may elicit greater improvements in aerobic fitness in people with cardiovascular disease (Hannan et al., 2018).

Exercise intensity for HIIT and M-VICT in patients with AF has been prescribed as a fraction of peak oxygen uptake ($\dot{V}O_{2peak}$) (Mertens and Kavanagh, 1996), peak heart rate (HR_{peak}) (Malmo et al., 2016), peak power output (PO_{peak}) (Reed et al., 2022), or based on ratings of perceived exertion (RPE) (Mertens and Kavanagh, 1996, Risom et al., 2016) achieved during a symptom-limited incremental cardiopulmonary exercise test (CPET). CPET protocols usually involve a gradual increase in the exercise work rate until patients reach volitional exhaustion or exhibit symptoms for which continued exercise is contraindicated. Despite some limitations of using fraction of maximal/peak responses for aerobic exercise prescription (Iannetta et al., 2020b), this approach is often used to prescribe individualized exercise intensity to maximize the effects of exercise training on aerobic fitness (MacInnis and Gibala, 2017).

To effectively prescribe exercise intensity or assess the effectiveness of interventions on aerobic fitness, variables determined from CPET must be reliable. In patients with stable angina (Sullivan et al., 1984) and coronary artery disease (Coeckelberghs et al., 2016, Faricier et al., 2023), CPET provides highly reliable measures of $\dot{V}O_{2peak}$ and HR_{peak} (intraclass correlation [ICC] ≥0.88). However, in patients with heart failure, $\dot{V}O_{2peak}$ (Elborn et al., 1990), HR_{peak}, and RPE (Bensimhon et al., 2008) have been shown to be greater in the second CPET when two CPETs were performed over two weeks, suggesting that the test-retest reliability may depend on clinical conditions or test-retest familiarization. Notably, 17 of these patients (57%) had AF in one of these studies (Elborn et al., 1990). In patients with AF, the complex presentation of disabling AF-related symptoms (e.g., palpitations, dyspnea, weakness, and dizziness) combined with the effects of medications to control HR may influence day-to-day exercise performance during CPET beyond the expected biological variation seen in other non-arrhythmia populations.

Establishing the test-retest reliability of CPET allows for the determination of (1) consistency and variability in exercise intensity prescription for patients with AF, and (2) minimal detectable change (Chelune, 1993) that distinguishes the training effects from error and test exposure at the individual level (O'Connell et al., 2021, Chelune, 1993). The primary purpose of this study was to examine the test-retest reliability of variables available from CPET (i.e., $\dot{V}O_{2peak}$, PO_{peak}, HR_{peak}, and RPE) in adult participants with persistent and permanent AF prior to embarking on a 12 week HIIT and M-VICT. The secondary purpose was to compare the proportion of participants achieving minimal detectable change after participating in 12 weeks of HIIT and M-VICT.

Materials and methods

Study design: We analyzed data collected in a previous RCT that compared the effects of HIIT and M-VICT on functional capacity (6-minute walk test distance) and quality of life in patients with persistent and permanent AF (Reed et al., 2022). The test-retest reliability was assessed by analyzing the data from a subgroup of participants who completed CPET on two separate occasions within two weeks at baseline (i.e., before the exercise training intervention). Data from participants who completed CPETs at baseline ($CPET_{pre}$) and following HIIT or M-VICT ($CPET_{post}$) were analyzed to compare the proportion of participants achieving the minimal detectable change in $\dot{V}O_{2peak}$ between the groups. The study was approved by the Ottawa Health Science Network Research Ethics Board (protocol number: 20150427-01H) and all participants provided written consent prior to participating in the study.

Participants: Detailed description of participants has been reported (Reed et al., 2022). Briefly, this study included patients with diagnosed persistent and permanent AF who were ≥40 years old and had a resting HR ≤110 bpm. Patients who were participating in routine exercise training (≥twice a week) or diagnosed with unstable angina, uncontrolled diabetes, severe mitral or aortic stenosis, or hypertrophic obstructive cardiomyopathy with significant obstruction were excluded. Demographic and medical information of participants were retrieved from the electronic medical records system (EPIC). Study staff measured stature, body mass, percent body fat (bioelectrical impedance analysis), resting HR and blood pressure. Participants' time spent in AF was assessed by a 24-hr Holter monitor. Physical activity levels were measured over seven days by an ActiGraph GT3X accelerometer (ActiGraph, Pensacola FL).

Baseline cardiopulmonary exercise testing (CPET): At baseline, participants completed two separate CPETs within two weeks. One CPET was completed on an electronically-braked cycle ergometer (ergoselect 100, ergoline, Germany), with gas exchange and ventilatory responses measured using a Vmax metabolic cart (Sensormedics, Yorba Linda, CA, United States). This CPET was supervised by cardiac stress technologists in the Department of Cardiac Imaging ($CPET_{diag}$). On a separate day, another CPET was completed by a certified study staff member ($CPET_{research}$) using an electronically-braked cycle ergometer (ergoselect 200, ergoline, Germany) and Oxycon Mobile portable indirect calorimetry (CareFusion, Canada). The open circuit calorimetry used in both CPETs provided breath-by-breath measures of oxygen uptake ($\dot{V}O_2$) and carbon dioxide exhalation ($\dot{V}CO_2$), based on which the respiratory exchange ratio (RER) was automatically calculated by the software. A 12-lead ECG was used to monitor HR continuously for $CPET_{diag}$ and a 3-lead ECG for $CPET_{research}$. The test termination criteria for

both tests were volitional exhaustion (e.g., participants requesting to stop or unable to sustain the cadence) or onset of severe symptoms (American College of Sports Medicine, 2021). Staff performing CPETs were blinded to the results of the other CPET.

The initial power output for $CPET_{diag}$ was set at 20 watts and increased by 20 watts per minute thereafter. The initial power output of $CPET_{research}$ was set at 40 watts, which did not change for the first two minutes, followed by an increase of 20 watts per minute thereafter. Consequently, the power output only differed during the first minute (Figure 1). RPE was assessed at the end of each minute and at volitional exhaustion using the Borg's 6-20 RPE Scale (Borg, 1970). For $\dot{V}O_2$ and $\dot{V}CO_2$, average measures captured over a 20-sec window were used for the analyses. The highest attained 20-second average $\dot{V}O_2$ during the exercise test was defined as $\dot{V}O_{2peak}$. The first ventilatory threshold (VT_1, gas exchange threshold) and second ventilatory threshold (VT_2, respiratory compensation point) were determined by an investigator blinded to the participant ID and conditions (i.e., $CPET_{diag}$ or $CPET_{research}$). As previously described, VT_1 and VT_2 were determined by plotting $\dot{V}CO_2$, ventilation ($\dot{V}_E$), $\dot{V}_E/\dot{V}O_2$, and $\dot{V}_E/\dot{V}CO_2$ against $\dot{V}O_2$ (Keir et al., 2022). Participants were instructed to take their prescribed medications as usual prior to CPETs.

Exercise intervention: Following the baseline measures, participants were randomized 1:1 to HIIT or M-VICT. Each HIIT session was performed on a cycle ergometer and was 23 minutes in duration, including: a 2-min warm-up at 50% of PO_{peak}; two blocks of 8 x 30-sec high-intensity bouts interspersed with 30-sec active recovery, with 4 min active recovery between the blocks; and, a 1-min cool-down at 25% PO_{peak}. Exercise intensity of the high-intensity bouts was based on the $CPET_{diag}$ results and set at 50% of PO_{peak} for the first week and increased by 10% after

every other session. When participants were able to tolerate the prescribed 100% PO_{peak}, they were permitted to increase the intensity up to 150% of their PO_{peak}. All HIIT sessions were supervised by study personnel. The M-VICT sessions were supervised by cardiovascular rehabilitation staff. Each M-VICT session was 60 minutes in duration, including: 10 to 15-minute warm-up, 30 minutes of continuous aerobic conditioning (i.e. walking or jogging, cycling, elliptical, rowing) at HR within 67-95% HR_{peak} or RPE between 12 and 16, and a 15-minute cool-down with stretching and strengthening exercise. Both HIIT and M-VICT sessions were performed 2 days/week for 12 weeks.

Follow-up CPET: Willing participants completed an additional CPET following the allocated exercise interventions (i.e., HIIT or M-VICT). The $CPET_{post}$ was performed by the study group using the Oxycon Mobile. The protocol used for the $CPET_{post}$ was the same as the $CPET_{research}$.

Minimal detectable change: Based on data obtained from two baseline CPETs (i.e., $CPET_{diag}$ and $CPET_{research}$) and using a reliable change index score of 1.96, the minimal detectable change in $\dot{V}O_{2peak}$ was calculated (O'Connell et al., 2021, Faricier et al., 2023). Participants with $\dot{V}O_{2peak}$ changes exceeding the minimal detectable change following HIIT and M-VICT were considered as positive responders.

Data analysis: All analyses were performed using SPSS statistical software version 29 (IBM Corp, Armonk, NY, USA). Descriptive statistics were used to summarize the characteristics of participants. Both relative and absolute reliabilities were assessed to determine the test-retest reliability of CPETs. Relative reliability was assessed using the intraclass correlation coefficient

(ICC) based on a two-way random-effects model and type consistency using average measures ($ICC_{3,1}$). The ICC was classified as follows: <0.5 as poor, ≥0.5 and <0.75 as moderate, ≥0.75 and <0.9 as good, and ≥0.9 as excellent (Koo and Li, 2016). Paired t-tests were used to compare $CPET_{diag}$ and $CPET_{research}$. Bland-Altman plots were used to assess the limits of agreement. The presence of proportional bias was assessed by the linear regression of differences on averages before plotting the Bland-Altman plots (Weir, 2005). When proportional bias was present, modified Bland-Altman plots were plotted.

For changes in aerobic fitness, the normality of data was assessed by the Shapiro-Wilk test. Repeated-measures analysis of variance (ANOVA) was used to examine the main effect of time and time by exercise intervention interaction effects. When two baseline CPETs were completed (i.e., $CPET_{diag}$ and $CPET_{research}$), the latter CPET conducted prior to attending prescribed exercise interventions were used for the analyses. The proportions of participants achieving minimal detectable change in $\dot{V}O_{2peak}$ were compared between HIIT and M-VICT using chi-square analyses. A p value <0.05 was considered statistically significant. All continuous data are presented as mean ± SD.

Results

Participants: Of 94 participants with persistent or permanent AF consented, 18 (33% persistent AF, mean age: 69 ± 7 years, 33% females) completed two CPETs within two weeks at baseline and 23 completed $CPET_{pre}$ and $CPET_{post}$ (n=13 in HIIT and n=10 in M-VICT). Participants' characteristics are summarized in **Table 1**. Overall, participants were not physically active (13 ± 11 minutes of moderate-intensity physical activity per day and 1 ± 1 minutes of vigorous-intensity physical activity per day) and spent most of the day in AF (≥ 95%). There was a large

variability in $\dot{V}O_{2peak}$ among participants at baseline, ranging from 11.3 to 31.9 mL/kg/min (i.e., very poor to excellent (American College of Sports Medicine, 2021)).

Test-retest reliability of baseline CPETs: The two baseline CPETs were separated by 7 ± 2 days. The majority of participants completed $CPET_{diag}$ first (n=12, 67%). All CPETs were terminated when participants reached volitional exhaustion; no CPETs were terminated by a supervisor due to the onset of exercise contraindications. Seventeen participants (94%) achieved a respiratory exchange ratio (RER) ≥1.10 during $CPET_{diag}$ and $CPET_{research}$. The variables measured during the two baseline CPETs are presented in **Figure 2**. At volitional exhaustion, the PO_{peak} (124 ± 40 vs. 148 ± 40 watts, p<0.001, *d*=-1.006) and HR_{peak} (136 ± 22 vs. 148 ± 30 bpm, p=0.023, *d*=-0.720) were significantly greater for $CPET_{research}$ compared to $CPET_{diag}$. However, the mean RPE at volitional exhaustion (18 ± 3 vs. 19 ± 1 points, p=0.129, *d*=-0.564) and the $\dot{V}O_{2peak}$ (18.6 ± 5.9 vs. 19.3 ± 5.1 mL/kg/min, p=0.393, *d*=-0.207) did not differ significantly between $CPET_{diag}$ and $CPET_{research}$.

Between $CPET_{diag}$ and $CPET_{research}$, the ICC was significant for $\dot{V}O_{2peak}$, PO_{peak}, HR_{peak}, and RPE at exhaustion. The ICC of $\dot{V}O_{2peak}$ was excellent, whereas the ICC of the other measures were in a good range. At VT_1, there were no differences in $\dot{V}O_2$ (10.3 ± 2.0 vs. 10.9 ± 1.8 mL/kg/min, p=0.191, *d*=-0.355) or $\%\dot{V}O_{2peak}$ (58.6 ± 11.5 vs. 58.8 ± 13.7 %, p=0.929, *d*=-0.023) between $CPET_{diag}$ and $CPET_{research}$. Their ICCs were significant (p=0.022 for $\dot{V}O_2$ and p=0.002 for $\%\dot{V}O_{2peak}$, respectively). VT_2 could only be identified in three participants for both $CPET_{diag}$ and $CPET_{research}$, and thus, was excluded from the analysis.

The Bland-Altman plots are presented in **Figure 3**. The limits of agreement were wide for PO_{peak} and HR_{peak} with significant systematic bias between the repeated tests. Although the

mean difference was small for $\dot{V}O_{2peak}$ between $CPET_{diag}$ and $CPET_{research}$, the limits of agreement were similarly wide (±6.2 mL/kg/min). For RPE, there were larger discrepancies between the two tests when average RPE was lower.

Exercise training responders: The characteristics of participants who completed $CPET_{pre}$ and $CPET_{post}$ are summarized in **Table 1**. Exercise adherence (i.e., % of exercise sessions attended) was high for both HIIT (91 ± 11%) and M-VICT (85 ± 10%) and did not differ between the groups (p=0.830). At week 10, 85% of participants in HIIT and 80% in M-VICT achieved the prescribed exercise intensity targets, with no difference in the proportion of participants achieving exercise target between the groups (p=0.772).

At week 12, no significant change in $\dot{V}O_{2peak}$ (18.8 ± 5.2 vs. 19.3 ± 4.6 mL/kg/min, main effect of time, p=0.386) or time by exercise intervention interaction effects (p=0.984) were found. The minimal detectable change for $\dot{V}O_{2peak}$ was 5.6 mL/kg/min. Following HIIT and M-VICT, there were no significant differences in the proportion of participants who achieved a reliable increase in $\dot{V}O_{2peak}$ (0 [0%] in HIIT and 1 [10%] in M-VICT, p=0.244).

Discussion

CPET plays an important role in patients with AF to: (1) diagnose morbidity and mortality risks; (2) prescribe individualized exercise intensity; (3) determine the effectiveness of exercise training interventions; and, (4) assess the prognosis of clinical conditions (Mezzani et al., 2009). However, for these methods to be valid in patients with AF, the reliability of CPET needs to be established for this condition. Because two repeated tests were completed by different groups (diagnostics and research personnel) using different metabolic systems, we present conservative reliability that includes intra-individual, inter-tester, and inter-device variabilities.

Between $CPET_{diag}$ and $CPET_{research}$, the ICC of PO_{peak} and HR_{peak} were in the good range. However, both PO_{peak} and HR_{peak} were significantly higher in $CPET_{research}$ compared to $CPET_{diag}$. In those unaccustomed to maximal and exhaustive efforts, increased confidence and expectations may permit heightened exercise tolerance and achievement of higher exercise intensity. Because a larger proportion of participants completed $CPET_{research}$ after $CPET_{diag}$ (12 of 18 patients, 67%), it was possible that greater PO_{peak} and HR_{peak} observed in $CPET_{research}$ were due to a learning or familiarization effect. To test the effects of the testing order, we performed dependent t-tests to compare the variables obtained in the first and second CPETs (results not shown). These comparisons showed no differences between the first and second CPETs (all $p>0.05$), highlighting that significantly higher PO_{peak} and HR_{peak} during $CPET_{research}$ were not due to learning effects. Consequently, the observed differences in PO_{peak} and HR_{peak} may reflect different testing settings, such as different degrees of encouragement provided in research versus clinical settings (Andreacci et al., 2002). Professionals involved in diagnostic evaluations often perform CPET to identify the underlying risks, whereas the research teams may often be more interested in peak/maximal performance. To partially support this, although the RPE did not differ significantly between $CPET_{research}$ and $CPET_{diag}$, the average RPE score was higher and variability smaller during $CPET_{research}$ compared to $CPET_{diag}$ (18 ± 3 vs. 19 ± 1 points, $p=0.129$, d=-0.564), suggesting a consistent and greater level of effort during $CPET_{research}$.

Despite significantly higher PO_{peak} and HR_{peak} in $CPET_{research}$ compared to $CPET_{diag}$, there was no difference in $\dot{V}O_{2peak}$. This was unexpected because, in healthy individuals, the expected gain in $\dot{V}O_2$ within the severe-intensity domain for a ramp of 20 watts/min ranges between 7.6 and 9.2 mL/min/watt (Iannetta et al., 2019). This is equivalent to approximately 2.0 to 2.4 mL/kg/min for a ~24 watts difference using the group mean body mass of our patients. The lack

of significant difference in $\dot{V}O_{2peak}$ despite an increase in PO_{peak} could be explained by physiological or technological factors. From a physiological perspective, this could reflect the achievement of maximal oxygen consumption ($\dot{V}O_{2max}$) in $CPET_{research}$ with additional work facilitated by anaerobic energy contributions. However, this is unlikely because a plateau in $\dot{V}O_2$ was not observed. Alternatively, the high reliability for $\dot{V}O_{2peak}$ in our study may be confounded by a systematic offset between $\dot{V}O_2$ measuring devices as accuracy and error are likely to differ between systems (Van Hooren et al., 2023). The excellent test-retest ICC for the $\dot{V}O_{2peak}$ aligns with previous findings in patients with coronary artery disease (ICC 0.95; 95% CI, 0.92-0.97) but the study reported no differences in the PO and HR at $\dot{V}O_{2peak}$ (Coeckelberghs et al., 2016). Thus, the use of two different metabolic systems complicates the interpretation of within-participant changes in gas exchange variables and may have contributed to the dissociation of $\dot{V}O_2$ from the HR and PO data. The wide limits of agreement for $\dot{V}O_{2peak}$ depicted in the Bland-Altman plots (±6.2 mL/kg/min) also represents the combination of inter-individual and inter-device variabilities. Similar to $\dot{V}O_{2peak}$, significant and good test-retest ICCs for the $\dot{V}O_2$ and $\%\dot{V}O_{2peak}$ at VT_1 need to be interpreted with caution given the different metabolic systems used. Future studies in AF should reevaluate the reliability of $\dot{V}O_{2peak}$ (and other gas exchange and ventilatory indices) using the same metabolic system. That said, it is important to note that many cardiac centres are equipped with different metabolic systems from different manufacturers and often lack a standardized exercise testing protocol.

After 12 weeks of exercise training, there were no significant changes in $\dot{V}O_{2peak}$. We also found that only one participant in M-VICT achieved the minimal detectable change in $\dot{V}O_{2peak}$. In the HIIT group, the exercise intensity prescription was based on PO_{peak} achieved during $CPET_{diag}$. As shown in the original study (Reed et al., 2022), $\%HR_{peak}$ and RPE of the

HIIT sessions were performed at our vigorous intensity classification (American College of Sports Medicine, 2021). However, given the significantly lower PO_{peak} in $CPET_{diag}$ when compared to $CPET_{research}$, the exercise prescription for HIIT may still have been insufficient to induce significant changes in $\dot{V}O_{2peak}$. It is possible that more rigorous and effective exercise prescriptions can be implemented when CPET is led by exercise specialists. In addition, PO_{peak} at volitional exhaustion depends on the ramp slope (i.e., the rate and magnitude of PO increase) (Iannetta et al., 2020a, Keir et al., 2018). This concept is likely transferable to patients with clinical conditions (Keir et al., 2018, Bowen et al., 2012) and makes it challenging to prescribe appropriate exercise intensity based on PO_{peak}. Considering that exercise intensity prescription based on HR may be confounded by arrhythmia and rate-controlling medications in AF, further research is warranted to elucidate the best way of prescribing an effective individualized exercise intensity for patients living with AF.

The M-VICT group was supervised by the cardiovascular rehabilitation team. The average RPE of the M-VICT group over the 12 weeks of intervention remained between 12 and 13 (i.e., moderate intensity (American College of Sports Medicine, 2021)) and HR during the exercise was also in a moderate-intensity range (Reed et al., 2022). Thus, the M-VICT group may also have required a greater exercise intensity to improve $\dot{V}O_{2peak}$. Additionally, a higher exercise training volume predicts a greater proportion of individuals achieving an increase in aerobic fitness >0.5 METs, a threshold that has been considered clinically meaningful (Bonafiglia et al., 2021). As recommended by the Canadian Cardiovascular Society AF guidelines (Andrade et al., 2020), a greater exercise volume (i.e., ≥200 minutes weekly) might be required to increase $\dot{V}O_{2peak}$ in patients with AF when exercise intensity is moderate. It is important to note that there was a significant improvement in six-minute walk test distance in

HIIT (21.3±34.1 m) and M-VICT (13.2±55.2 m) (Reed et al., 2022). This inconsistency (i.e., an increase in six-minute walk test distance but not in $\dot{V}O_{2peak}$) may be due to different statistical power or an increase in submaximal thresholds without a rise in $\dot{V}O_{2peak}$. Alternatively, there may be a dissociation between changes in six-minute walk test distance and $\dot{V}O_{2peak}$. The six-minute walk test distance significantly correlates cross sectionally with $\dot{V}O_{2peak}$ in patients with coronary artery disease (Mandic et al., 2013). However, the correlation between longitudinal changes in six-minute walk test distance and $\dot{V}O_{2peak}$ is unclear. Considering that most participants in HIIT and M-VICT did not change $\dot{V}O_{2peak}$, finding the optimal exercise prescription to increase responders is warranted for patients with AF.

This study has limitations. First, we examined the test-retest reliability of patients with AF using different testers and metabolic systems. Thus, the reliability we observed can be attributed to the combination of intra-individual and inter-testing environment variabilities. As discussed above, this limitation also reflects a reality of many cardiovascular rehabilitation environments where different metabolic systems are used. Second, the order of CPET was not randomized. While our subsequent analyses showed no differences between the first and second CPETs, randomized crossover study is needed to establish the reliability of CPET in AF. Third, the timing of two CPETs were not standardized. While participants were instructed to take prescribed medications prior to CPETs, their impact may have differed depending on the time of CPETs. In addition, although we confirmed the administration of prescribed medications prior to $CPET_{research}$, this was not documented for $CPET_{diag}$. This may have affected HR responses to $CPET_{diag}$. Fourth, because $CPET_{post}$ was only performed by those willing to complete this additional CPET, the participants included in the HIIT versus M-VICT analyses may have been more motivated. However, we observed no improvement in aerobic fitness. Last, the study may

have lacked the power to detect changes in aerobic fitness over time or differences between HIIT and M-VICT.

Conclusion

We found that PO_{peak} and HR_{peak} can differ significantly when CPETs are repeated under different settings in patients with persistent and permanent AF. Caution must be practiced when prescribing exercise intensity based on these measures as under-prescription of exercise intensity can result in increased number of non-responders. Exercise interventions to increase responders are warranted in patients with AF.

Arthur's contribution: TT, DAK, JMM, and JLR contributed to the conception or design of the work. TT, SVA, and JLR contributed to the acquisition of data. TT, DAK, JMM, and JLR contributed to data analysis and interpretation. TT drafted the manuscript. DAK, JMM, SVA, JB, and JLR critically revised the manuscript. All gave final approval and agreed to be accountable for all aspects of work, ensuring integrity and accuracy.

Data availability statement: A data sharing agreement can be requested to our legal team.

Conflict of interest: None

Funding: This study was supported by the Innovations Fund of the Alternate Funding Plan for the Academic Health Sciences Centre of the Ministry of Ontario, New Investigators Award in Clinical Rehabilitation from CIHR, Heart and Stroke Foundation of Emerging Research Leaders Initiative.

References:

American College of Sports Medicine 2021. ACSM's Guidelines for Exercise Testing and Prescription. Eleventh Editin. *ACSM's Guidelines for Exercise Testing and Prescription. Eleventh Edition.*

Andrade, J. G., Aguilar, M., Atzema, C., Bell, A., Cairns, J. A., Cheung, C. C., Cox, J. L., Dorian, P., Gladstone, D. J., Healey, J. S., Khairy, P., Leblanc, K., Mcmurtry, M. S., Mitchell, L. B., Nair, G. M., Nattel, S., Parkash, R., Pilote, L., Sandhu, R. K., Sarrazin, J. F., Sharma, M., Skanes, A. C., Talajic, M., Tsnag, T. S. M., Verma, A., Verma, S., Whitlock, R., Wyse, D. G., Macle, L. The 2020 Canadian Cardiovascular Society/Canadian Heart Rhythm Society Comprehensive Guidelines for the Management of Atrial Fibrillation. *Can J Cardiol,* 36**,** 1847-1948.

Andreacci, J. L., Lemura, L. M., Cohen, S. L., Urbansky, E. A., Chelland, S. A. & Von Duvillard, S. P. 2002. The effects of frequency of encouragement on performance during maximal exercise testing. *J Sports Sci,* 20**,** 345-52.

Bensimhon, D. R., Leifer, E. S., Ellis, S. J., Fleg, J. L., Keteyian, S. J., Pina, I. L., Kitzman, D. W., Mckelvie, R. S., Kraus, W. E., Forman, D. E., Kao, A. J., Whellan, D. J., O'Conor, C. M., Russell, S. D. 2008. Reproducibility of peak oxygen uptake and other cardiopulmonary exercise testing parameters in patients with heart failure (from the Heart Failure and A Controlled Trial Investigating Outcomes of exercise traiNing). *Am J Cardiol,* 102**,** 712-7.

Bonafiglia, J. T., Preobrazenski, N., Islam, H., Walsh, J. J., Ross, R., Johannsen, N. M., Martin, C. K., Church, T. S., Slentz, C. A., Ross, L. M., Kraus, W. E., Kenny, G. P., Goldfield, G. S., Prud'homme, D., Sigal, R. J., Earnest, C. P. & Gurd, B. J. 2021. Exploring differences in cardiorespiratory fitness response rates across varying doses of exercise

409 training: A retrospective analysis of eight randomized controlled trials. *Sports Med,* 51**,**
410 1785-1797.
411 Borg, G. 1970. Perceived exertion as an indicator of somatic stress. *Scand J Rehabil Med,* 2**,** 92-
412 8.
413 Bowen, T. S., Cannon, D. T., Begg, G., Baliga, V., Witte, K. K. & Rossiter, H. B. 2012. A novel
414 cardiopulmonary exercise test protocol and criterion to determine maximal oxygen
415 uptake in chronic heart failure. *J Appl Physiol (1985),* 113**,** 451-8.
416 Chelune, H. G. N., I Richard; Luders, Hans; Seldlak, Jeffery; Awad, A Issam1993. Individual
417 change after epilepsy surgery: Practice effects and base-rate information.
418 *Neuropsychology,* 7**,** 41-52.
419 Chiang, C. E., Naditch-brule, L., Murin, J., Goethals, M., Inoue, H., O'Neill, J., Silva-cardoso, J.,
420 Zharinov, O., Gamra, H., Alam, S., Ponikowski, P., Lewalter, T., Rosenqvist, M. & Steg,
421 P. G. 2012. Distribution and risk profile of paroxysmal, persistent, and permanent atrial
422 fibrillation in routine clinical practice: insight from the real-life global survey evaluating
423 patients with atrial fibrillation international registry. *Circ Arrhythm Electrophysiol,* 5**,**
424 632-9.
425 Coeckelberghs, E., Buys, R., Goetschlckx, K., Pattyn, N., Vanhees, L. & Cornelissen, V. 2016.
426 Test-Retest Reliability of Maximal and Submaximal Gas Exchange Variables in Patients
427 With Coronary Artery Disease. *J Cardiopulm Rehabil Prev,* 36**,** 263-9.
428 Dai, H., Zhang, Q., Much, A. A., Maor, E., Segev, A., Beinart, R., Adawi, S., Lu, Y., Bragazzi,
429 N. L. & Wu, J. 2021. Global, regional, and national prevalence, incidence, mortality, and
430 risk factors for atrial fibrillation, 1990-2017: results from the Global Burden of Disease
431 Study 2017. *Eur Heart J Qual Care Clin Outcomes,* 7**,** 574-582.

Elborn, J. S., Stanford, C. F. & Nicholls, D. P. 1990. Reproducibility of cardiopulmonary parameters during exercise in patients with chronic cardiac failure. The need for a preliminary test. *Eur Heart J,* 11**,** 75-81.

Faricier, R., Keltz, R., Hartley, T., Mckelvie, R., Suskin, N., Prior, P. & Keir, D. 2023. Quantifying improvement in V̇O2peak and exercise thresholds in cardiovascular disease using reliable change indices. *Journal of Cardiopulmonary Rehabilitattion and Prevention,* Accepted.

Garnvik, L. E., Malmo, V., Janszky, I., ELLEKJAER, H., WISLOFF, U., LOENNECHEN, J. P. & NES, B. M. 2020. Physical activity, cardiorespiratory fitness, and cardiovascular outcomes in individuals with atrial fibrillation: the HUNT study. *Eur Heart J,* 41**,** 1467-1475.

Hannan, A. L., Hing, W., Simas, V., Climstein, M., Coombes, J. S., Jayasinghe, R., Byrmes, J. & Furness, J. 2018. High-intensity interval training versus moderate-intensity continuous training within cardiac rehabilitation: a systematic review and meta-analysis. *Open Access J Sports Med,* 9**,** 1-17.

Iannetta, D., De Almeida Azevedo, R., Ingram, C. P., Keir, D. A. & Murias, J. M. 2020a. Evaluating the suitability of supra-PO(peak) verification trials after ramp-incremental exercise to confirm the attainment of maximum O(2) uptake. *Am J Physiol Regul Integr Comp Physiol,* 319**,** R315-R322.

Iannetta, D., De Almeida Azevedo, R., Keir, D. A. & Murias, J. M. 2019. Establishing the V̇O2 versus constant-work-rate relationship from ramp-incremental exercise: simple strategies for an unsolved problem. *J Appl Physiol (1985),* 127**,** 1519-1527.

454 Iannetta, D., Inglis, E. C., Mattu, A. T., Fontana, F. Y., Pogliaghi, S., Keir, D. A. & Murias, J. M.
455 2020b. A Critical Evaluation of current methods for exercise prescription in women and
456 men. *Med Sci Sports Exerc,* 52**,** 466-473.
457 Keir, D. A., Iannetta, D., Mattioni Maturana, F., Kowlchuk, J. M. & Murias, J. M. 2022.
458 Identification of non-invasive exercise thresholds: methods, strategies, and an online app.
459 *Sports Med,* 52**,** 237-255.
460 Keir, D. A., Paterson, D. H., Kowalchuk, J. M. & Murias, J. M. 2018. Using ramp-incremental
461 VO(2) responses for constant-intensity exercise selection. *Appl Physiol Nutr Metab,* 43**,**
462 882-892.
463 Kokkinos, P., Faselis, C., Samuel, I. B. H., Lavie, C. J., Zhang, J., Vargas, J. D., Pittaras, A.,
464 Doumas, M., Karasik, P., Moore, H., Heimal, M. & Myers, J. 2023. Changes in
465 cardiorespiratory fitness and survival in patients with or without cardiovascular disease. *J*
466 *Am Coll Cardiol,* 81**,** 1137-1147.
467 Koo, T. K. & Li, M. Y. 2016. A guideline of selecting and reporting intraclass correlation
468 coefficients for reliability research. *J Chiropr Med,* 15**,** 155-63.
469 Kornej, J., Borschel, C. S., Benjamin, E. J. & Schnabel, R. B. 2020. Epidemiology of atrial
470 fibrillation in the 21st century: novel methods and new insights. *Circ Res,* 127**,** 4-20.
471 Lip, G. Y., Fauchier, L., Freedman, S. B., Van Gelder, I., Natale, A., Gianni, C., Nattel, S.,
472 Potpara, T., Rienstra, M., Tse, H. F. & Lane, D. A. 2016. Atrial fibrillation. *Nat Rev Dis*
473 *Primers,* 2**,** 16016.
474 MaCinnis, M. J. & Gibala, M. J. 2017. Physiological adaptations to interval training and the role
475 of exercise intensity. *J Physiol,* 595**,** 2915-2930.

476 Malmo, V., Nes, B. M., Amundsen, B. H., Tjonna, A. E., Stoylen, A., Rossvoll, O., Wisloff, U.
477 & Loennechen, J. P. 2016. Aerobic interval training reduces the burden of atrial
478 fibrillation in the short term: a randomized trial. *Circulation,* 133**,** 466-73.
479 Mandic, S., Walker, R., Stevens, E., Nye, E. R., Body, D., Barclay, L. & Williams, M. J. 2013.
480 Estimating exercise capacity from walking tests in elderly individuals with stable
481 coronary artery disease. *Disabil Rehabil,* 35**,** 1853-8.
482 Mentel, A., Quinne, T. J., Cameron, A. C., Lees, K. R. & Abdul-rahim, A. H. 2020. The impact
483 of atrial fibrillation type on the risks of thromboembolic recurrence, mortality and major
484 haemorrhage in patients with previous stroke: A systematic review and meta-analysis of
485 observational studies. *Eur Stroke J,* 5**,** 155-168.
486 Mertens, D. J. & Kavanagh, T. 1996. Exercise training for patients with chronic atrial
487 fibrillation. *J Cardiopulm Rehabil,* 16**,** 193-6.
488 Mezzani, A., Agostoni, P., Cohen-Solal, A., Corra, U., Jegier, A., Kouidi, E., Mazic, S., Meurin,
489 P., Piepoli, M., Simon, A., Laethem, C. V. & Vanhees, L. 2009. Standards for the use of
490 cardiopulmonary exercise testing for the functional evaluation of cardiac patients: a
491 report from the Exercise Physiology Section of the European Association for
492 Cardiovascular Prevention and Rehabilitation. *Eur J Cardiovasc Prev Rehabil,* 16**,** 249-
493 67.
494 O'Connell, M. E., Suskin, N. & Prior, P. L. 2021. Measuring true change in individual patients:
495 reliable change indices of cardiac rehabilitation outcomes, and implications for quality
496 indicators. *CJC Open,* 3**,** 1139-1148.
497 Reed, J. L., Terada, T., Vidal-Almela, S., Tulloch, H. E., Mistura, M., Birnie, D. H., Wells, G.
498 A., Nair, G. M., Hans, H., Way, K. L., Chirico, D., O'Neill, C. D. & Pipe, A. L. 2022.

Effect of high-intensity interval training in patients with atrial fibrillation: a randomized clinical trial. *JAMA Netw Open,* 5**,** e2239380.

Rison, S. S., Zwisler, A. D., Rasmussen, T. B., Siblitz, K. L., Madsen, T. L., Svendsen, J. H., Gluud, C., Lindschou, J., Winkel, P. & Berg, S. K. 2016. Cardiac rehabilitation versus usual care for patients treated with catheter ablation for atrial fibrillation: Results of the randomized CopenHeart(RFA) trial. *Am Heart J,* 181**,** 120-129.

Ruddox, V., Sandven, I., Munkhaugen, J., Skattebu, J., Edvardsen, T. & Otterstad, J. E. 2017. Atrial fibrillation and the risk for myocardial infarction, all-cause mortality and heart failure: A systematic review and meta-analysis. *Eur J Prev Cardiol,* 24**,** 1555-1566.

Schnabel, R. B., Yin, X., Gona, P., Larson, M. G., Beiser, A. S., Mcmanus, D. D., Newton-Cheh, C., Lubitz, S. A., Magnani, J. W., Ellinor, P. T., Seshadri, S., Wolf, P. A., Vasan, R. S., Benjamin, E. J. & Levy, D. 2015. 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. *Lancet,* 386**,** 154-62.

Shukla, A. & Curtis, A. B. 2014. Avoiding permanent atrial fibrillation: treatment approaches to prevent disease progression. *Vasc Health Risk Manag,* 10**,** 1-12.

Sullivan, M., Genter, F., Savvides, M., Roberts, M., Myers, J. & Froelicher, V. 1984. The reproducibility of hemodynamic, electrocardiographic, and gas exchange data during treadmill exercise in patients with stable angina pectoris. *Chest,* 86**,** 375-82.

Van Hooren, B., Souren, T. & Bongers, B. C. 2023. Accuracy of respiratory gas variables, substrate, and energy use from 15 CPET systems during simulated and human exercise. *Scand J Med Sci Sports*.

521 Weir, J. P. 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and
522 the SEM. *J Strength Cond Res,* 19**,** 231-40.
523 Wolf, P. A., Abbott, R. D. & Kannel, W. B. 1991. Atrial fibrillation as an independent risk factor
524 for stroke: the Framingham Study. *Stroke,* 22**,** 983-8.
525
526

527 **Figure legends**

528 **Figure 1**. Baseline cardiopulmonary exercise testing (CPET) protocols.

529 Broken line represents $CPET_{research}$ and solid line represents $CPET_{diag}$, respectively.

530

531 **Figure 2**. Test-retest reliability of baseline cardiopulmonary exercise testing

532 $CPET_{diag}$: cardiopulmonary exercise test conducted by diagnostics; $CPET_{research}$: cardiopulmonary

533 exercise testing conducted by the study group; HR: heart rate; PO_{peak}: peak power output; RPE:

534 rating of perceived exertion; $\dot{V}O_{2peak}$: peak oxygen consumption; VT_1: first ventilatory threshold

535 (gas exchange threshold)

536 $ICC_{3,1}$: intraclass correlation coefficient based on a two-way random-effects model and type

537 consistency using average measures

538 Significantly different from $CPET_{diag}$: *p<0.05 and †p<0.001

539

540 **Figure 3.** Bland-Altman plots:

541 CPET: cardiopulmonary exercise testing; $CPET_{diag}$: CPET performed by diagnostics;

542 $CPET_{research}$: CPET performed by the study group; HR: heart rate; HR_{peak}: peak heart rate; PO_{peak}:

543 peak power output; LOA: limits of agreement; RER: respiratory exchange ratio; RPE: rating of

544 perceived exertion; $\dot{V}O_{2peak}$: peak oxygen consumption; VT_1: first ventilatory threshold (gas

545 exchange threshold)

546

Table 1. Participant characteristics

	Test-retest reliability	Exercise interventions		
	(N=18)	HIIT (n=13)	M-VICT (n=10)	HIIT vs. M-VICT, p-value
Female, n (%)	6 (33)	3 (23)	4 (40)	0.382
Persistent AF, n (%)	6 (33)	6 (46)	3 (30)	0.431
Permanent AF, n (%)	12 (67)	7 (54)	7 (70)	
Time in AF, %	96 (12)	95 (15)	99 (3)	0.426
Age, years	69 (7)	67 (4)	71 (9)	0.215
Height, cm	172.7 (7.6)	174.6 (4.3)	174.8 (9.5)	0.947
Body mass, kg	91.5 (19.4)	96.8 (20.6)	95.6 (25.5)	0.901
Body mass index, kg/m^2	30.6 (5.7)	31.8 (7.0)	31.2 (8.6)	0.844
Fat mass, %	28.9 (7.5)	30.3 (8.0)	30.0 (7.0)	0.944
Systolic blood pressure, mmHg	126 (19)	123 (15)	133 (20)	0.164
Diastolic blood pressure, mmHg	80 (11)	77 (10)	84 (9)	0.117
Resting heart rate, bpm	73 (11)	70 (13)	71 (9)	0.769
Physical activity level				
Light, min/day	66 (45)	69 (54)	92 (88)	0.456
Moderate, min/day	13 (11)	19 (17)	16 (15)	0.670
Vigorous, min/day	1 (1)	1 (1)	1 (1)	0.359
Medical condition, n (%)				
Hypertension	10 (56)	7 (54)	5 (50)	0.855
Dyslipidemia	6 (33)	3 (23)	2 (20)	0.859
Coronary artery disease	2 (11)	2 (15)	0 (0)	0.194
Diabetes	4 (22)	1 (8)	3 (30)	0.162
Medication, n (%)				
ACE inhibitor	5 (28)	4 (31)	1 (10)	0.231
ARB	3 (17)	1 (8)	4 (40)	0.063
Statin	9 (50)	5 (39)	4 (40)	0.940
Beta-blocker	12 (67)	11 (85)	5 (50)	0.074

Data are presented as mean (SD) unless indicated otherwise. AF: atrial fibrillation; ARB: angiotensin receptor blockers

Figure 1

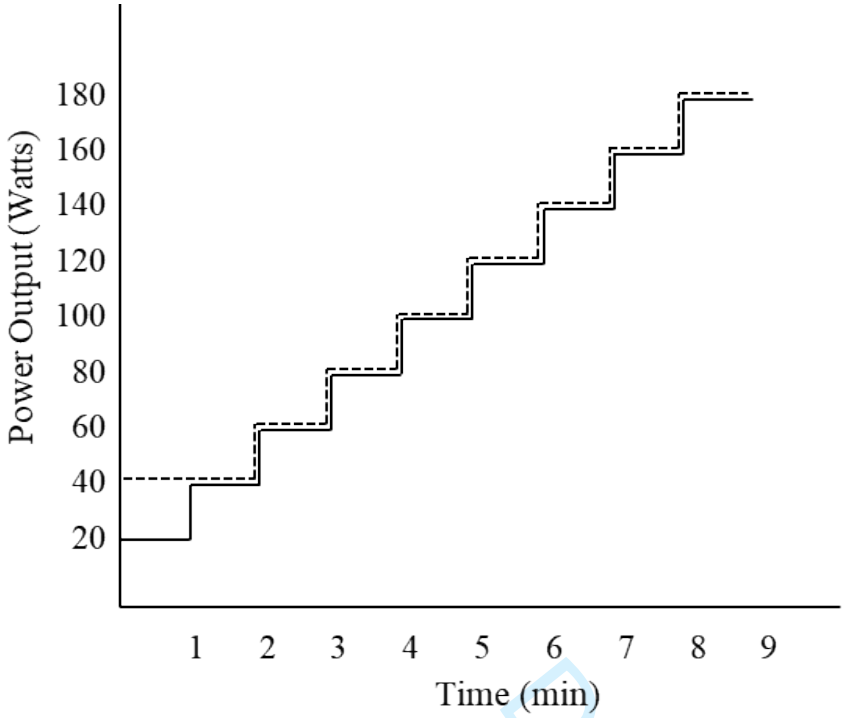
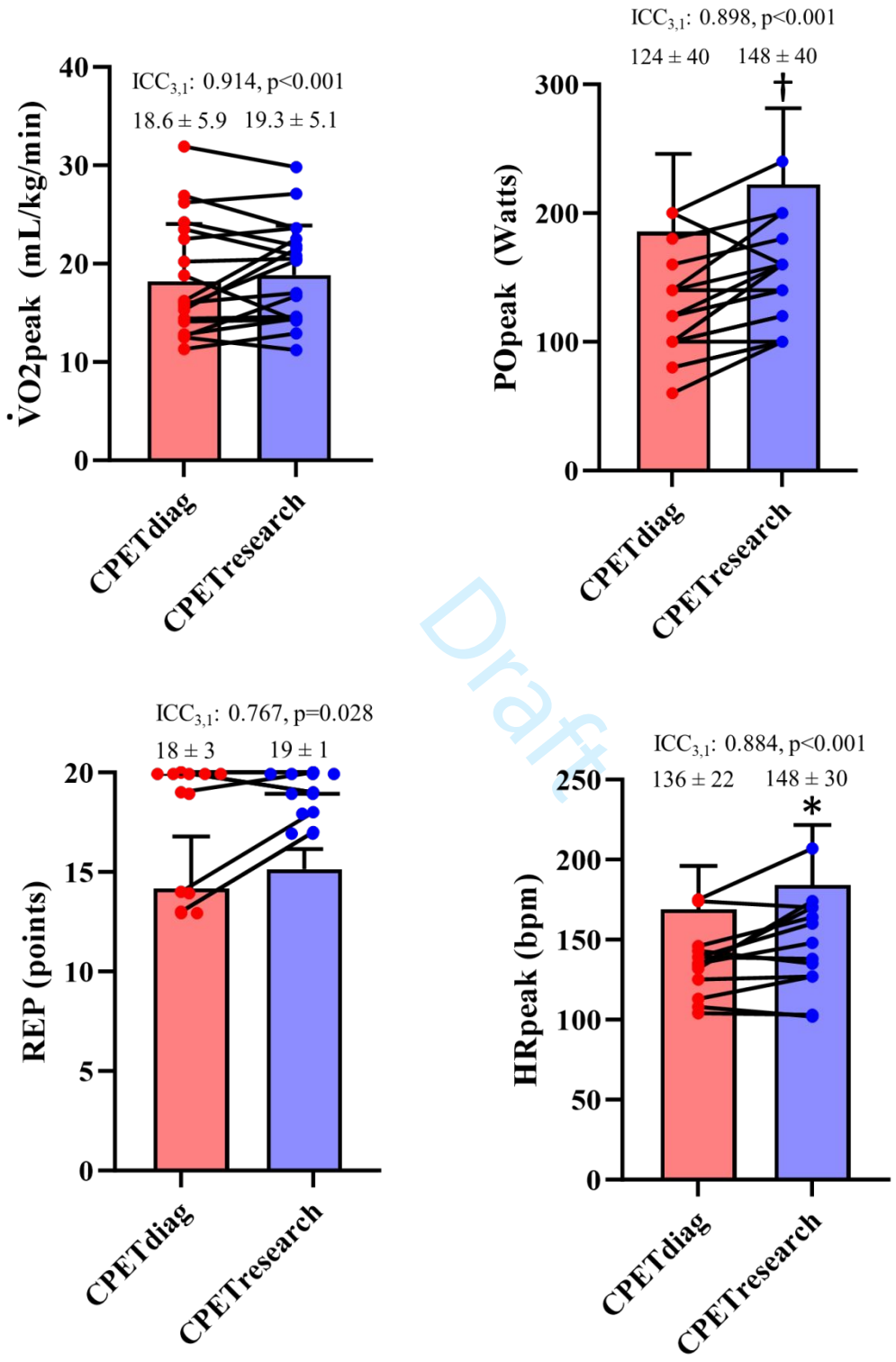


Figure 2



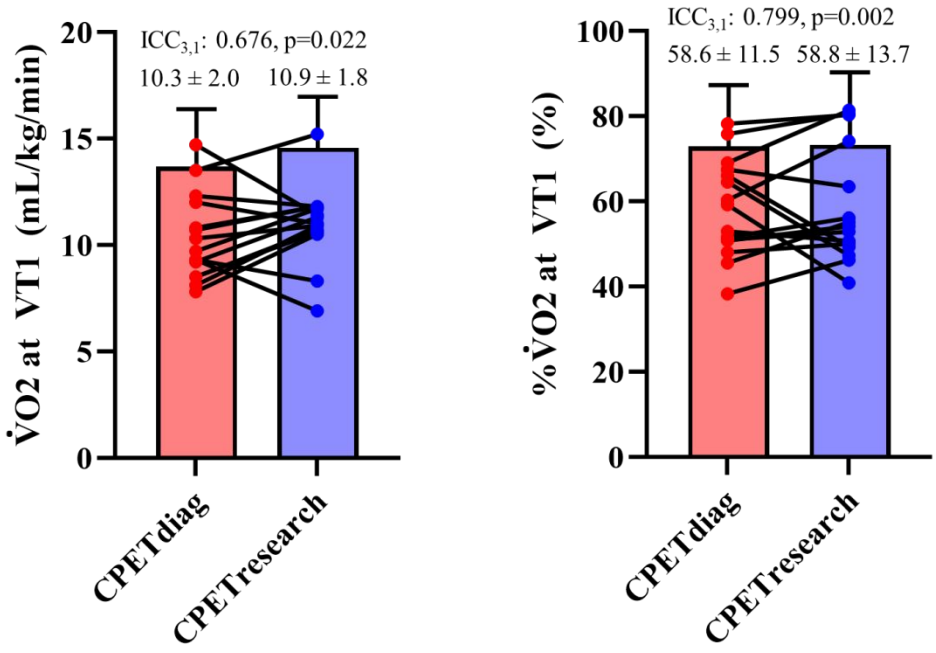


Figure 3