Heterogeneity and incidence of non-response for changes in cardiorespiratory fitness following time-efficient sprint interval exercise training

Richard S. Metcalfe[1][*] & Niels B.J. Vollaard[2]

[1] Applied Sports Science Technology and Medicine Research Centre (A-STEM), Swansea University, Swansea, SA1 8EN, UK. Email: r.s.metcalfe@swansea.ac.uk

[2] Faculty of Health Sciences and Sport, University of Stirling, Stirling, FK9 4LA, UK. Email: n.vollaard@stir.ac.uk

[*] **Corresponding Author:**

Dr Richard S Metcalfe

Applied Sports, Technology, Exercise and Medicine (A-STEM) Research Centre

College of Engineering

Swansea University

Swansea

Wales, UK

SA1 8EN

Email: r.s.metcalfe@swansea.ac.uk

Abstract

Interindividual variability for training-induced changes in maximal oxygen uptake (VO_2max) is well described following continuous aerobic and high-intensity interval training. Whether similar variability is observed following time-efficient sprint interval training with minimal training volume (i.e., reduced-exertion high-intensity interval training; REHIT) is unknown. We conducted a pooled analysis of n=117 (68 men) training participants (mean±SD: age: 30±10 y; VO_2max: 34.8±7.5 ml·kg^{-1}·min^{-1}), who completed a VO_2max assessment before and 3 days after 6 weeks of REHIT comprising of two 10-20 second 'all-out' cycling sprints per session, and n=40 no-intervention control participants (age: 30±13 y; VO_2max: 31.5±6.5 ml·kg^{-1}·min^{-1}) who completed repeated VO_2max tests over a comparable timeframe. Individual responses estimated using 50% confidence intervals derived from the technical error were interpreted against a smallest worthwhile change (SWC) of 1.75 ml·kg^{-1}·min^{-1}. The standard deviation of individual responses was 2.39 ml·kg^{-1}·min^{-1} demonstrating clinically meaningful heterogeneity in training-induced changes in VO_2max following REHIT that exceed the technical, biological and random within-subjects variability of VO_2max assessment. The likely (75% probability) non-response rate was 18% (21/117), and 49% (57/117) of individuals demonstrated increases in VO_2max likely higher than the SWC. We conclude that the well-described increase in VO_2max following REHIT at the group level is subject to substantial variability in magnitude at an individual level. This has important implications for exercise prescription and can be harnessed to elucidate mechanisms of adaptation.

Novelty:

- There is substantial heterogeneity in VO_2max responses following time-efficient sprint interval training
- Proportion of non-response was 18% and ⌣50% of individuals show clinically meaningful increases in VO_2max

1. Introduction

The maximal attainable rate of oxygen uptake (VO_2max) is amongst the most important physiological traits that determine long term health and longevity. Indeed, a low VO_2max predicts cardiovascular and all-cause mortality to a similar or greater extent compared with other established risk factors, including body mass index, smoking, hypertension and type 2 diabetes (Ross et al., 2016). Although VO_2max is a partly heritable trait (Bouchard et al., 1998), it can also be improved (on average) through regular exercise training (Bouchard et al., 1999; Sisson et al., 2009), and those who are able to improve their VO_2max over several years lower their risk of cardiovascular and all-cause mortality in a dose-dependent manner (Lee et al., 2011).

Whilst VO_2max improves on average in response to both continuous endurance and high-intensity interval exercise training (Bacon et al., 2013; Weston et al., 2014), it has been recognised for over 3 decades that individual measured changes following standardised exercise training can be highly variable and a proportion of people will demonstrate no measurable change (so called, 'non-responders') (Bouchard et al., 1999; Lortie et al., 1984). Even following several months of high-volume aerobic exercise training, measured individual changes in VO_2max can range from decreases of 100 ml·min^{-1} to gains of more than 1100 ml·min^{-1} (Bouchard et al., 1999). This interindividual variability in response is thought to be explained by a range of factors, including random or technical error, the method of standardising relative exercise intensity, and genetic and epigenetic variance (Sarzynski et al., 2017). Some have also argued that 'non-responders' to exercise training in general do not exist and instead are an artefact of an insufficient training stimulus for those individuals (Bacon et al., 2013; Montero and Lundby, 2017). Nonetheless, some individuals may be non-responders to set training interventions (characterised by training intensity, duration, frequency, and mode) that are efficacious at inducing training effects in others.

Over the last 15 years, sprint interval training (SIT) has emerged as an efficacious exercise training stimulus for improving VO_2max in previously inactive individuals (Gillen and Gibala, 2014; Sultana et al., 2019; Vollaard and Metcalfe, 2017; Vollaard et al., 2017). A particularly interesting finding to emerge from SIT research is that the training-induced change in VO_2max does not appear to increase (and possibly decreases) with an increasing

number of sprint repetitions (Vollaard et al., 2017). Indeed, at a group level, improvements in VO_2max have been observed with as few as two or three, 20-second, all-out cycling sprints performed regularly over a 6-12-week training intervention (termed 'reduced-exertion high-intensity interval training' or REHIT) (Gillen et al., 2016; Metcalfe et al., 2016, 2012). These findings have particular relevance in the search for effective, 'real-world', time-efficient exercise strategies to overcome lack of time as a barrier to exercise initiation and adherence in low active individuals (Vollaard and Metcalfe, 2017).

The variability in training-induced changes in VO_2max has been well described in response to continuous endurance training (Bouchard et al., 1999; Sisson et al., 2009) and more recently in response to high-intensity (HIIT) and sprint interval training (SIT) (Astorino and Schubert, 2014; Bonafiglia et al., 2016; Gurd et al., 2016; Islam et al., 2020; Phillips et al., 2017; Williams et al., 2019). However, studies examining interindividual variability in response to SIT to date have involved arduous SIT protocols requiring a relatively high number of sprint repetitions (Astorino and Schubert, 2014; Bonafiglia et al., 2016; Gurd et al., 2016; Islam et al., 2020; Williams et al., 2019), have combined HIIT and SIT protocols together (Williams et al., 2019) or, in some cases, the analysis has been limited by not utilising relevant information from a time-matched control condition (Astorino and Schubert, 2014). No study has characterised the heterogeneity in response or the incidence of non-responders to a genuinely time-efficient SIT exercise protocol such as REHIT. Given the low overall dose of exercise involved with REHIT (a total of <10 minutes of sprint exercise within 3 hours of training time over a 6-week training period, e.g. (Metcalfe et al., 2016, 2012)), alongside suggestions that 'non-responders' may be an artefact of an insufficient training stimulus (Bacon et al., 2013; Montero and Lundby, 2017), it is reasonable to question whether the incidence of non-response would be high following this training intervention, and what proportion of individuals (if any) are likely to show a change that would be considered clinically meaningful. Individual variability in the training-induced change in VO_2max in response to REHIT has been alluded to (Metcalfe et al., 2016), but not definitively demonstrated using an adequate sample size, or appropriate experimental and statistical methods. The inclusion of data from no-exercise control group is particularly important when assessing individual responses to exercise training in order to account for the variance caused by technical error, day-to-day biological and random within-subjects variability (Atkinson and Batterham, 2015; Bonafiglia et al., 2019). Thus, the aim of this study was firstly to establish whether true individual variability in changes in VO_2max in

response to REHIT exists and, if so, to characterise the heterogeneity of response and incidence of non-responders to this extremely low-volume and time-efficient exercise intervention.

2. Methods

2.1. Participants / Sample

We conducted a secondary analysis of six published studies conducted in our laboratories between 2012 and 2020 (Metcalfe et al., 2020, 2016, 2012; Nalçakan et al., 2018; Songsorn et al., 2016; Thomas et al., 2020). This included a sample of 117 pooled training participants (68 male, 49 female) who underwent an almost identical SIT intervention (described in full below), and 40 pooled no-intervention control participants (16 male, 24 female) from three of these studies who underwent two assessments of maximal oxygen uptake either ~4 (n=14; (Songsorn et al., 2016)) or ~6 (n=26; (Metcalfe et al., 2020, 2012)) weeks apart. Training participants from two of the included studies who underwent substantially different SIT interventions, either involving single 20-s sprints (Songsorn et al., 2016) or reduced sprint duration (Nalçakan et al., 2018), were excluded because these studies demonstrated that these interventions either do not alter VO_2max (single 20-s sprints; (Songsorn et al., 2016)) or results in a significantly lower mean increase (reduced sprint duration; (Nalçakan et al., 2018)). The inclusion and exclusion criteria were similar across all six included studies. All participants were classified as either sedentary or moderately physically active on enrolment onto the study according to the criteria of the International Physical Activity Questionnaire. Participants with any contraindication to exercise based on a self-report health history questionnaire and an assessment of high resting blood pressure (>140/90 mmHg) or high resting heart rate (>100 bpm) were excluded. The pooled participant characteristics are shown in **Table 1**. Ethical approval was obtained for all included experiments (details and approval references are available in the original articles (Metcalfe et al., 2020, 2016, 2012; Nalçakan et al., 2018; Songsorn et al., 2016; Thomas et al., 2020)) and all participants provided their written consent to take part after they received information about the study both verbally and in writing. All experiments were conducted in accordance with the Declaration of Helsinki.

Table 1 Baseline Participant Characteristics

	Training (n=117)	Control (n=40)	p-value

Male / Female (n)	68/49	16/24	-
Age (y)	30 ± 10	30 ± 13	0.98
Height (m)	1.72 ± 0.09	1.69 ± 0.09	0.17
Body Mass (kg)	74.7 ± 15.5	70.4 ± 16.1	0.14
BMI ($kg \cdot m^{-2}$)	25.2 ± 4.2	24.4 ± 4.6	0.30
VO_2max ($ml \cdot kg^{-1} \cdot min^{-1}$)	34.8 ± 7.5	31.5 ± 6.5	0.01

Data are presented as mean $\pm$ SD. p-values derived from independent t-test.

2.2. Assessment of Maximal Oxygen Uptake

All participants underwent an incremental cycling test to their limit of tolerance to determine maximal oxygen uptake. In participants in the SIT groups, these were conducted in the week prior to training and then ~3 days following the final training session. The specific protocols have been described previously (Metcalfe et al., 2020, 2016, 2012; Nalçakan et al., 2018; Songsorn et al., 2016; Thomas et al., 2020). Although there were slight variations in the protocol and equipment used across studies, the mean and standard deviation of the change in VO_2max are strikingly similar (**Table 3**) suggesting these are 'typical' for studies of this nature and hence appropriate for a pooled analysis. In the majority (n=103) of training participants and all (n=40) of the control participants, breath by breath measurements of pulmonary gas exchange were captured continuously during the test using a metabolic cart. Breath by breath measurements of oxygen uptake (VO_2) were converted into 15-breath rolling averages and VO_2max was taken as the highest value for a 15-breath rolling average achieved during the test. This method was used in all but one of the included experiments which was reanalysed for the present analysis (Metcalfe et al., 2012). For a small sample of training participants (n=14) from one study (Metcalfe et al., 2016), pulmonary gas exchange was measured using the Douglas Bag technique. It was considered appropriate to retain these participants in this analysis because previous independent studies have reported similar test-retest reliability for VO_2max measured using Douglas Bags or breath by breath methods (Katch et al., 1982; Phillips et al., 2017). The following secondary criteria were used to verify a maximal effort: 1) volitional exhaustion, 2) a plateau in VO_2 despite increasing workload, 3) RER >1.15 and 4) a maximal heart rate within 10 beats of age predicted maximum (220-age). All participants achieved two or more of these criteria.

2.3. Sprint Interval Training and No-Intervention Control

All training participants underwent a 6-week cycling based REHIT intervention with only small differences across studies (**Table 2**). The majority of participants (n=104) completed

this supervised in an exercise physiology laboratory on a mechanically braked cycle ergometer (Monark, Vansbro, Sweden) (Metcalfe et al., 2016, 2012; Nalçakan et al., 2018; Thomas et al., 2020). Participants (n=13) in one study completed the intervention unsupervised on a commercially available electronically braked cycle ergometer (CAROL[TM], Integrated Health Partners Ltd, London, UK) (Metcalfe et al., 2020). Each SIT session lasted ~10 min and consisted of low intensity cycling interspersed with two 'all-out' sprints against a fixed resistance (between 10 and 20 s). In the 2-3 seconds prior to each 'all-out' sprint, participants increased their pedal cadence to their maximal speed, the braking resistance was then applied to the bike, and participants cycled as fast as they could for the duration of the sprint. The majority of participants completed 3 sessions/week (n=76) but a subset completed either 2 (n=29) or 4 (n=12) sessions/week. We recently demonstrated that the mean change in VO_2max with SIT was not different across training frequencies of 2, 3 or 4 sessions/week (Thomas et al., 2020), so it was deemed appropriate to pool them together in this analysis. All no-intervention control participants from the 3 separate studies were given the same instructions to maintain their current physical activity levels and dietary patterns for the duration of the study.

Table 2 *Training interventions applied in included studies*

	Metcalfe 2012	Metcalfe 2016	Nalcakan 2018	Thomas 2020	Metcalfe 2020
Intervention Duration (weeks)	6	6	6	6	6
Frequency (sessions / week)	3	3	3	2, 3 or 4	2
Total Session Duration (mins)	10	10	10	10	8:20-8:40
Sprints Per Session (n)	2	2	2	2	2
Braking Mass (% BM)	7.5	5	7.5	7.5	5
Sprint Duration Week 1 (s)	10	10	10	10	10
Sprint Duration Week 2 (s)	15	15	15	15	15
Sprint Duration Week 3 (s)	15	15	20	20	20
Sprint Duration Week 4 (s)	20	20	20	20	20
Sprint Duration Week 5 (s)	20	20	20	20	20
Sprint Duration Week 6 (s)	20	20	20	20	20
Warm-Up Duration (s)	180	180	120	100-110	120
Recovery Duration (s)	200-220	200-220	200-220	200-220	180
Cool Down Duration (s)	180	180	240	240	180
Intensity of Warm-Up, Recovery and Cool Down (W)	~60 W	Unloaded	Unloaded	Unloaded	~25 W

Note: in the Metcalfe et al (2012 and 2016) studies, the first session only contained 1 x 10 s sprint. Where a range of warm-up and recovery durations are presented, this is due to the increase in sprint duration during initial training weeks.

2.4. Statistical Analysis

2.4.1. Group Responses

The effect of SIT on VO_2max was analysed using a two-way repeated measures analysis of variance (group x time) performed in Graphpad Prism 8 for macOS (Version 8.4.2, San Diego, CA, USA). Alpha was set at $p<0.05$ and effect size was calculated using Cohens *d* (Hopkins et al., 2009).

2.4.2. Individual Variability

To characterise whether true individual responses to SIT were present, the standard deviation (SD) of individual responses (SD_{IR}) was calculated from the square root of the difference between the square of the SD of the change (post *minus* pre) in the exercise training group (SD_{ex}) and the control group (SD_{con}) (Atkinson and Batterham, 2015):

$$SD_{IR} = \sqrt{(SD_{EX})^2 - (SD_{CON})^2}$$

The SD_{IR} was subsequently interpreted against thresholds for standardised mean change of 0.1 (small effect), 0.3 (moderate effect) and 0.6 (large effect) (Hopkins, 2015), and against the smallest clinically worthwhile change (see below).

2.4.3. Classification of Non-, Uncertain and Positive Responders to SIT

The statistical procedures recommended by Swinton et al (Swinton et al., 2018) were followed to classify individual changes in VO_2max following SIT using 50% confidence intervals (CIs) that were calculated using the typical error for repeated measurements of VO_2max from the control participants. The typical error (TE) was calculated using the formula:

$$TE = SD_{CON} / \sqrt{2}$$

Typical error for relative VO_2max was 1.30 $ml·kg^{-1}·min^{-1}$. For completeness, we classified individual responses against a zero and clinically relevant response thresholds as this is more likely to be informative of individual responsiveness in short term (e.g. 6 weeks) exercise training studies where adaptations may still be accumulating (Islam et al., 2020). However, when interpreting and discussing the data we take the more conservative approach of defining

responses against the smallest worthwhile change, alluding to responses against the zero threshold where relevant. Responses were considered against two clinically relevant thresholds of change to provide additional information on the proportion of individual responses at different magnitudes: thresholds of 1.75 ml·kg^{-1}·min^{-1} (equivalent to ~0.5 METs) and 3.5 ml·kg^{-1}·min^{-1} (equivalent to ~1 MET) were applied based on the data of Lee *et al* who demonstrated an ~8% and ~16% decrease in relative risk of CVD, respectively, over ~11 years of follow up (Lee et al., 2011). Both of these thresholds have been applied in previous studies on this topic (Bonafiglia et al., 2018). A responder was classified if the entire 50% CI lay above the specified response threshold. In these instances, the use of a 50% CI means that there is a 75% probability of a response for this individual, i.e. 'likely' (Hopkins, 2015). If the 50% CI crossed the response threshold then this was classified as an 'uncertain response', whilst a 'non-responder' was defined if the entire 50% CI lay below the response threshold (Bonafiglia et al., 2018). All analysis of individual responses was performed in Microsoft Excel and Graphpad Prism 8 for macOS (Version 8.4.2, San Diego, CA, USA).

3. Results

3.1. Group Effects and SD_{IR}

For relative VO_2max, there were main effects of group ($F=13.2$, $p<0.001$) and time ($F=35.7$, $p<0.001$) as well as a group x time interaction ($F=39.8$, $p<0.001$): mean relative VO_2max increased in the SIT group compared to the control group (**Table 3 and Figure 1,** $d=0.43$). The SD_{IR} for the change in relative VO_2max following SIT compared to the control group was 2.39 $ml·kg^{-1}·min^{-1}$ with an effect size of 0.32 ('moderate'). The SD_{IR} was not substantially altered (2.30 $ml·kg^{-1}·min^{-1}$) when one potentially influential SIT participant with a -5.99 $ml·kg^{-1}·min^{-1}$ decrease in VO_2max was removed from the analysis ($d=0.30$, 'moderate'). Thus, there is evidence of individual differences in VO_2max trainability in response to REHIT of a magnitude that exceeds the smallest clinically worthwhile effect of 1.75 $ml·kg^{-1}·min^{-1}$.

Table 3 Changes in absolute and relative VO_2max following SIT

	Study	Absolute (L·min⁻¹)			Relative (ml·kg⁻¹·min⁻¹)		
		Pre	Post	Delta	Pre	Post	Delta
Training	Metcalfe 2012	2.19±0.57	2.48±0.60	0.29±0.13	33.5±4.9	38.3±4.8	4.7±2.7
	Metcalfe 2016	2.54±0.65	2.78±0.67	0.24±0.23	35.0±7.8	38.1±7.9	3.0±3.3
	Nalcakan 2018	2.77±0.75	3.04±0.75	0.27±0.28	39.0±6.9	42.3±6.5	3.3±3.3
	Thomas 2020	2.77±0.77	3.01±0.82	0.24±0.24	35.4±7.2	38.5±7.5	3.1±2.8
	Metcalfe 2020	2.25±0.75	2.42±0.82	0.17±0.21	28.0±6.7	29.8±7.6	1.8±2.5
Control	Metcalfe 2012	2.32±0.64	2.37±0.77	0.05±0.18	33.8±5.5	34.2±6.0	0.3±2.3

Songsorn 2015		2.07±0.69	2.08±0.68	0.01±0.10	32.0±5.8	32.2±6.0	0.2±1.6
Metcalfe 2020		2.26±0.70	2.20±0.71	-0.06±0.09	28.1±7.3	27.4±7.2	-0.7±1.4
Pooled	**Training**	**2.59±0.73**	**2.83±0.77**	**0.24±0.23**	**34.8±7.5**	**37.9±7.9**	**3.1±3.0**
	Control	**2.21±0.67**	**2.22±0.71**	**0.00±0.14**	**31.5±6.5**	**31.5±6.8**	**-0.1±1.8**

Data are presented as mean ± SD

3.2. Classification of Non-, Uncertain and Positive Responders to SIT

When considered against a threshold of zero, 5/117 (4%) of participants were classified as likely non-responders, 29/117 (25%) were classified as uncertain responders, and 83/117 (71%) of the participants showed an increase in VO_2max that was likely higher than zero (**Figure 2A**). When considered against the minimal clinically relevant threshold, 18% (21/117) of participants were likely non-responders, 39/117 (33%) were classified as uncertain responders, and 57/117 (49%) showed an increase likely higher than the minimal clinically relevant threshold of 1.75 ml·kg^{-1}·min^{-1} (**Figure 2B**). Furthermore, 33 out of those 57 (i.e. 33/117; 28%) of those participants showed an increase likely higher than 3.5 ml·kg^{-1}·min^{-1} (**Figure 2B**).

4. Discussion

Individual variability in the change in VO_2max following exercise training has important implications for optimising and personalising exercise prescriptions to improve health. This is the first study to report the heterogeneity of response and the incidence of non-response to a genuinely time-efficient SIT protocol that has been shown to be efficacious at a group level in both supervised lab (Metcalfe et al., 2016, 2012; Nalçakan et al., 2018; Thomas et al., 2020) and unsupervised real-world (Metcalfe et al., 2020) settings. Importantly, for the first time, we demonstrate statistically that inter-individual differences in training-induced changes in VO_2max following REHIT exceed the technical, day-to-day biological, and random within-subjects variability of VO_2max assessment over a similar time frame. Of particular note, we report a non-response rate of 18% and that 49% of individuals demonstrate an increase in VO_2max that is likely higher than the smallest clinically worthwhile difference.

Several previous studies have quantified non-response rates to a high dose of aerobic exercise training and to HIIT and SIT protocols involving a greater number/duration of high-intensity or all-out sprint efforts (Astorino and Schubert, 2014; Bonafiglia et al., 2016; Gurd et al., 2016; Islam et al., 2020; Phillips et al., 2017; Sisson et al., 2009; Williams et al., 2019). Comparisons of non-response rates between different studies and exercise protocols (in similar populations) requires consideration of the mean change, the SD_{IR} and the threshold applied to quantify non-response (Atkinson et al., 2019). This makes direct comparisons between the current data and these previous HIIT and SIT studies somewhat challenging because they have either not quantified the SD_{IR} and/or different thresholds have been applied to quantify non-response, including a change <1xTE (i.e. <0.86 ml·kg^{-1}·min^{-1};

(Phillips et al., 2017)), <2xTE (i.e. <1.74 ml·kg^{-1}·min^{-1}; (Gurd et al., 2016)), <1 coefficient of variation (CV) of repeated measurement (i.e. <~1.3 ml·kg^{-1}·min^{-1}; (Astorino and Schubert, 2014)), or <SWC minus 1 x TE (i.e. <1.80 ml·kg^{-1}·min^{-1}; (Williams et al., 2019)). In the largest pooled analysis of HIIT (n=299) and SIT (n=116) to date, Williams et al reported a mean change and non-response rate of +3.3 ml·kg^{-1}·min^{-1} and 35% for HIIT, and +2.5 ml·kg^{-1}·min^{-1} and 52% for SIT, respectively, using a threshold of <1.8 ml·kg^{-1}·min^{-1} to define a non-responder (Williams et al., 2019). If we apply the same criteria to the current study this would yield a non-response rate of 38% (with a pooled mean change of +3.1 ml·kg^{-1}·min^{-1}). Thus, it is possible to conclude that the proportion of likely non-responders observed with REHIT is similar to and not greater than other SIT and HIIT protocols involving a greater number/duration of high-intensity or all-out sprint efforts (Williams et al., 2019).

The fact that similar rates of non-response are observed when applying HIIT and SIT protocols with varying numbers and durations of sprints implies that non-response for VO$_2$max is not an artefact caused by an insufficient dose of exercise training. If this were the case, then the non-response rate would be expected to decrease with an increased number/duration of sprints and hence a greater training 'stimulus'. The application of sprints of an 'all-out' intensity adds further weight to this argument because this likely overcomes any issues regarding the standardisation of exercise intensity (or more specifically the homeostatic disruption) across individuals during each acute training session. Whilst individual differences in the homeostatic disturbance can be expected with MICT standardised using a % of HRmax or VO$_2$max, and this may have subsequent implications for the adaptive response (Jamnick et al., 2020; Montero and Lundby, 2017; Preobrazenski et al., 2018), the disruption in homeostasis with SIT / REHIT is always likely to be severe and hence sufficient to switch on signalling pathways that underpin adaptation (if possible for that individual). Indeed, the fact that group level increases in VO$_2$max are similar between SIT protocols applying 2-3 sprints compared with protocols applying 6-8 sprints (Vollaard et al., 2017) implies that the adaptive signalling pathways responsible for increasing VO$_2$max in response to SIT become 'saturated' with only a small number of acute 'all-out' sprint efforts (Vollaard and Metcalfe, 2017). Taken together, we contend that non responders for VO$_2$max in response to exercise training are a real physiological phenomenon and our data show they are observed in response to REHIT/SIT. That said, it remains unknown whether non-responders to REHIT/SIT for changes in VO$_2$max are also non-responders for other health markers (i.e. universal non responders), but previous work looking at interindividual

differences in response to aerobic exercise training suggests that this may not be the case (Vollaard et al., 2009). Future studies need to clearly establish whether non-responders to REHIT/SIT would demonstrate adaptations to different types of exercise, such as MICT. Although previous studies have attempted to examine this (Bonafiglia et al., 2016), the small sample size limits the conclusions that can be made at this stage.

Our data show that ~half of individuals are likely to show an increase in VO_2max that is greater than the smallest worthwhile change after only 6 weeks of training. This is a striking observation given that REHIT involves less than 10 min of sprint exercise and a total exercise time of 3 hours over that 6-week period. Nevertheless, it is worth noting that this represents a relatively short-term training intervention. At a group level, the magnitude of increase in VO_2max following SIT that occurs from 6 to 12 weeks is similar to that observed from 0 to 6 weeks of the intervention (Gillen et al., 2016) and this raises important questions about individual differences in the rate at which adaptations are accrued. Indeed, in our analysis, a proportion of individuals (26/117; 22%) demonstrated a change that was likely >0 $ml \cdot kg^{-1} \cdot min^{-1}$ but not of a magnitude which exceeded the minimal clinically important difference threshold of 1.75 $ml \cdot kg^{-1} \cdot min^{-1}$. Whilst it would be expected that non-responders would continue to show limited adaptation to a longer intervention, it will be important to determine whether individuals showing a slower rate of adaptation at 6 weeks would continue to accrue adaptations (at this lower rate) and would demonstrate a clinically meaningful magnitude of response at a later time point.

The mechanisms that explain the individual variation in response to REHIT are unclear and cannot be determined from the present study. As the training interventions applied were unable (due to logistical reasons) to control the majority of environmental factors outside of the training intervention (e.g., each participant's pattern of nutrition, sleep, stress etc.), it can be expected that some of the variability in adaptation between individuals is explained by (unquantified) environment-training interactions. As an example, the dose, type and timing of nutrition can have a powerful impact on the skeletal muscle signalling response to acute exercise (Cluberton et al., 2005; Guerra et al., 2010; Stocks et al., 2018) and, hence, can modify adaptation to exercise training. If such variables are not controlled across individuals, then it can be expected that this will introduce a level of individual variability in adaptation. This may also explain, at least in part, the poor reproducibility of individual responses to

repeated (identical) exercise training interventions in the same sample of participants (Del Giudice et al., 2020).

On the other hand, it is also clear that there is a heritable component to exercise trainability (Bouchard, 2019; Bouchard et al., 1999; Sarzynski et al., 2017; Timmons et al., 2010). The majority of this evidence comes from studies of the training response to aerobic and resistance exercise and the relevance of this information to SIT remains unknown. Indeed, we still do not know whether REHIT/SIT enhances VO_2max through mechanisms that are similar to or distinct from MICT (Gibala and Little, 2019; Vollaard and Metcalfe, 2017). However, the inter-individual variability observed in the present study can be used by future investigations interested in identifying (molecular) predictors of response. If responders vs. non-responders to a specific intervention (e.g. SIT/REHIT) can be identified, then contrasting traits / molecular signatures in groups of responders vs. non-responders provides a strong approach to identify potential physiological / genetic / epigenetic factors that determine the interindividual variability in training response (Keller et al., 2007). Such studies are needed to enable personalised medicine, for example, to enable personalised advice on effective interventions (Keller et al., 2007; Timmons et al., 2010) and can be a powerful way to elucidate molecular mechanisms of training adaptations (Keller et al., 2011, 2007). An improved understanding of the molecular mechanisms of adaptation to SIT would be invaluable in the effort to optimise SIT protocols to enable the greatest adaptations with minimal required effort and time-commitment.

There are a number of limitations to the current analysis that should be considered. Firstly, and most importantly, whilst the large sample size is a strength of this study, this is a pooled dataset from six independent studies and there were minor differences in training protocols, testing procedures, and the duration of the control intervention (4-weeks for n=14 and 6-weeks for n=26) between some of the studies (described in full in the methods). It is possible that these differences may affect the validity of the SD_{IR} estimate, which assumes that all sources of variability are equal between the exercise and control groups except that the exercise group underwent exercise training (Atkinson and Batterham, 2015; Bonafiglia et al., 2019). However, two pieces of information can help to mitigate these concerns. Firstly, the SD of the training-induced change in VO_2max was comparable between studies (**Table 3**). Secondly, we performed a sensitivity analysis to examine how excluding the 4-week control participants impacts the SD_{IR} calculation. We found a similar SD_{IR} estimate of 2.30 ml·kg⁻

$^1 \cdot min^{-1}$ compared to 2.39 $ml \cdot kg^{-1} \cdot min^{-1}$ when all control participants were included. Similarly, a comparable SD_{IR} of 2.36 $ml \cdot kg^{-1} \cdot min^{-1}$ was found when training participants undergoing VO_2max assessment with Douglas Bags (n=14) were excluded from the analysis. Thus, these differences in methodology across the pooled independent studies to do not appear to greatly impact the validity of our findings.

Another limitation is that most of the studies involved supervised, lab-based exercise, so it remains unclear whether the results would be the same in real-world settings and this will be important to address in future studies. Furthermore, this analysis is also largely limited to young sedentary but healthy men and women and it is not possible to determine whether different populations (e.g., lean vs overweight, young vs old, men vs women) may show different levels of response / non-response. Finally, it should also be noted that other important health markers were not considered in this analysis and so at present it remains unknown whether non-responders for VO_2max in response to SIT would be able to improve markers of cardiometabolic health.

In conclusion, we demonstrate for the first time that the well described increase in VO_2max observed following REHIT at the group level, is subject to substantial variability in magnitude at an individual level. This is an important observation with potential future implications for prescribing SIT/REHIT as an intervention for improving health and can be harnessed by future studies aiming to elucidate the mechanisms by which REHIT improves VO_2max.

References

Astorino, T.A., Schubert, M.M., 2014. Individual responses to completion of short-term and chronic interval training: a retrospective study. PLoS One 9, e97638. https://doi.org/10.1371/journal.pone.0097638

Atkinson, G., Batterham, A.M., 2015. True and false interindividual differences in the physiological response to an intervention: True and false interindividual differences. Exp. Physiol. 100, 577–588. https://doi.org/10.1113/EP085070

Atkinson, G., Williamson, P., Batterham, A.M., 2019. Issues in the determination of 'responders' and 'non responders' in physiological research. Exp. Physiol. 104, 1215–1225. https://doi.org/10.1113/EP087712

Bacon, A.P., Carter, R.E., Ogle, E.A., Joyner, M.J., 2013. VO2max trainability and high intensity interval training in humans: a meta-analysis. PloS One 8, e73182. https://doi.org/10.1371/journal.pone.0073182

Bonafiglia, J.T., Brennan, A.M., Ross, R., Gurd, B.J., 2019. An appraisal of the SD $_{IR}$ as an estimate of true individual differences in training responsiveness in parallel arm exercise randomized controlled trials. Physiol. Rep. 7. https://doi.org/10.14814/phy2.14163

Bonafiglia, J.T., Nelms, M.W., Preobrazenski, N., LeBlanc, C., Robins, L., Lu, S., Lithopoulos, A., Walsh, J.J., Gurd, B.J., 2018. Moving beyond threshold-based dichotomous classification to improve the accuracy in classifying non-responders. Physiol. Rep. 6, e13928. https://doi.org/10.14814/phy2.13928

Bonafiglia, J.T., Rotundo, M.P., Whittall, J.P., Scribbans, T.D., Graham, R.B., Gurd, B.J., 2016. Inter-Individual Variability in the Adaptive Responses to Endurance and Sprint Interval Training: A Randomized Crossover Study. PLOS ONE 11, e0167790. https://doi.org/10.1371/journal.pone.0167790

Bouchard, C., 2019. DNA Sequence Variations Contribute to Variability in Fitness and Trainability. Med. Sci. Sports Exerc. 51, 1781–1785. https://doi.org/10.1249/MSS.0000000000001976

Bouchard, C., An, P., Rice, T., Skinner, J.S., Wilmore, J.H., Gagnon, J., Pérusse, L., Leon, A.S., Rao, D.C., 1999. Familial aggregation of VO(2max) response to exercise training: results from the HERITAGE Family Study. J Appl Physiol 1985 87, 1003–8.

Bouchard, C., Daw, E., Rice, T., Pérusse, L., Gagnon, J., Province, M., Leon, A., Rao, D., Skinner, J., Wilmore, J., 1998. Familial resemblance for ˙VO2maxin the sedentary state: the HERITAGE family study. Med. Sci. Sports Exerc. 30, 252–258.

Cluberton, L.J., McGee, S.L., Murphy, R.M., Hargreaves, M., 2005. Effect of carbohydrate ingestion on exercise-induced alterations in metabolic gene expression. J. Appl. Physiol. Bethesda Md 1985 99, 1359–1363. https://doi.org/10.1152/japplphysiol.00197.2005

Del Giudice, M., Bonafiglia, J.T., Islam, H., Preobrazenski, N., Amato, A., Gurd, B.J., 2020. Investigating the reproducibility of maximal oxygen uptake responses to high-intensity interval training. J. Sci. Med. Sport 23, 94–99. https://doi.org/10.1016/j.jsams.2019.09.007

Gibala, M.J., Little, J.P., 2019. Physiological basis of brief vigorous exercise to improve health. J. Physiol. JP276849. https://doi.org/10.1113/JP276849

Gillen, J.B., Gibala, M.J., 2014. Is high-intensity interval training a time-efficient exercise strategy to improve health and fitness? Appl Physiol Nutr Metab 39, 409–12. https://doi.org/10.1139/apnm-2013-0187

Gillen, J.B., Martin, B.J., MacInnis, M.J., Skelly, L.E., Tarnopolsky, M.A., Gibala, M.J., 2016. Twelve Weeks of Sprint Interval Training Improves Indices of Cardiometabolic Health Similar to Traditional Endurance Training despite a Five-Fold Lower Exercise Volume and Time Commitment. PLoS One 11, e0154075. https://doi.org/10.1371/journal.pone.0154075

Guerra, B., Guadalupe-Grau, A., Fuentes, T., Ponce-González, J.G., Morales-Alamo, D., Olmedillas, H., Guillén-Salgado, J., Santana, A., Calbet, J.A.L., 2010. SIRT1, AMP-activated protein kinase phosphorylation and downstream kinases in response to a single bout of sprint exercise: influence of glucose ingestion. Eur. J. Appl. Physiol. 109, 731–743. https://doi.org/10.1007/s00421-010-1413-y

Gurd, B.J., Giles, M.D., Bonafiglia, J.T., Raleigh, J.P., Boyd, J.C., Ma, J.K., Zelt, J.G.E., Scribbans, T.D., 2016. Incidence of nonresponse and individual patterns of response following sprint interval training. Appl. Physiol. Nutr. Metab. Physiol. Appl. Nutr. Metab. 41, 229–234. https://doi.org/10.1139/apnm-2015-0449

Hopkins, W.G., 2015. Individual responses made easy. J. Appl. Physiol. 118, 1444–1446. https://doi.org/10.1152/japplphysiol.00098.2015

Hopkins, W.G., Marshall, S.W., Batterham, A.M., Hanin, J., 2009. Progressive statistics for studies in sports medicine and exercise science. Med Sci Sports Exerc 41, 3–13. https://doi.org/10.1249/MSS.0b013e31818cb278

Islam, H., Siemens, T.L., Matusiak, J.B.L., Sawula, L., Bonafiglia, J.T., Preobrazenski, N., Jung, M.E., Gurd, B.J., 2020. Cardiorespiratory fitness and muscular endurance responses immediately and 2 months after a whole-body Tabata or vigorous-intensity continuous training intervention. Appl. Physiol. Nutr. Metab. Physiol. Appl. Nutr. Metab. 45, 650–658. https://doi.org/10.1139/apnm-2019-0492

Jamnick, N.A., Pettitt, R.W., Granata, C., Pyne, D.B., Bishop, D.J., 2020. An Examination and Critique of Current Methods to Determine Exercise Intensity. Sports Med. https://doi.org/10.1007/s40279-020-01322-8

Katch, V.L., Sady, S.S., Freedson, P., 1982. Biological variability in maximum aerobic power. Med. Sci. Sports Exerc. 14, 21–25.

Keller, P., Vollaard, N., Babraj, J., Ball, D., Sewell, D.A., Timmons, J.A., 2007. Using systems biology to define the essential biological networks responsible for adaptation to endurance exercise training. Biochem. Soc. Trans. 35, 1306–1309. https://doi.org/10.1042/BST0351306

Keller, P., Vollaard, N.B.J., Gustafsson, T., Gallagher, I.J., Sundberg, C.J., Rankinen, T., Britton, S.L., Bouchard, C., Koch, L.G., Timmons, J.A., 2011. A transcriptional map of the impact of endurance exercise training on skeletal muscle phenotype. J. Appl. Physiol. Bethesda Md 1985 110, 46–59. https://doi.org/10.1152/japplphysiol.00634.2010

Lee, D.C., Sui, X., Artero, E.G., Lee, I.M., Church, T.S., McAuley, P.A., Stanford, F.C., Kohl, H.W., Blair, S.N., 2011. Long-term effects of changes in cardiorespiratory fitness and body mass index on all-cause and cardiovascular disease mortality in men: the Aerobics Center Longitudinal Study. Circulation 124, 2483–90. https://doi.org/10.1161/CIRCULATIONAHA.111.038422

Lortie, G., Simoneau, J.A., Hamel, P., Boulay, M.R., Landry, F., Bouchard, C., 1984. Responses of Maximal Aerobic Power and Capacity to Aerobic Training. Int. J. Sports Med. 05, 232–236. https://doi.org/10.1055/s-2008-1025911

Metcalfe, R.S., Atef, H., Mackintosh, K., McNarry, M., Ryde, G., Hill, D.M., Vollaard, N.B.J., 2020. Time-efficient and computer-guided sprint interval exercise training for improving health in the workplace: a randomised mixed-methods feasibility study in office-based employees. BMC Public Health 20, 313. https://doi.org/10.1186/s12889-020-8444-z

Metcalfe, R.S., Babraj, J.A., Fawkner, S.G., Vollaard, N.B., 2012. Towards the minimal amount of exercise for improving metabolic health: beneficial effects of reduced-exertion high-intensity interval training. Eur J Appl Physiol 112, 2767–75. https://doi.org/10.1007/s00421-011-2254-z

Metcalfe, R.S., Tardif, N., Thompson, D., Vollaard, N.B., 2016. Changes in aerobic capacity and glycaemic control in response to reduced-exertion high-intensity interval training (REHIT) are not different between sedentary men and women. Appl Physiol Nutr Metab 41, 1117–1123. https://doi.org/10.1139/apnm-2016-0253

Montero, D., Lundby, C., 2017. Refuting the myth of non response to exercise training: 'non responders' do respond to higher dose of training. J. Physiol. 595, 3377–3387. https://doi.org/10.1113/JP273480

Nalçakan, G.R., Songsorn, P., Fitzpatrick, B.L., Yüzbasioglu, Y., Brick, N.E., Metcalfe, R.S., Vollaard, N.B.J., 2018. Decreasing sprint duration from 20 to 10 s during reduced-exertion high-intensity interval training (REHIT) attenuates the increase in maximal aerobic capacity but has no effect on affective and perceptual responses. Appl Physiol Nutr Metab 43, 338–344. https://doi.org/10.1139/apnm-2017-0597

Phillips, B.E., Kelly, B.M., Lilja, M., Ponce-González, J.G., Brogan, R.J., Morris, D.L., Gustafsson, T., Kraus, W.E., Atherton, P.J., Vollaard, N.B.J., Rooyackers, O., Timmons, J.A., 2017. A Practical and Time-Efficient High-Intensity Interval Training Program Modifies Cardio-Metabolic Risk Factors in Adults with Risk Factors for Type II Diabetes. Front. Endocrinol. 8. https://doi.org/10.3389/fendo.2017.00229

Preobrazenski, N., Bonafiglia, J.T., Nelms, M.W., Lu, S., Robins, L., LeBlanc, C., Gurd, B.J., 2018. Does blood lactate predict the chronic adaptive response to training: A comparison of traditional and talk test prescription methods. Appl. Physiol. Nutr. Metab. https://doi.org/10.1139/apnm-2018-0343

Ross, R., Blair, S.N., Arena, R., Church, T.S., Després, J.P., Franklin, B.A., Haskell, W.L., Kaminsky, L.A., Levine, B.D., Lavie, C.J., Myers, J., Niebauer, J., Sallis, R., Sawada, S.S., Sui, X., Wisløff, U., Health, A.H.A.P.A.C. of the C. on L., Cardiometabolic, Cardiology, C. on C., Prevention, C. on E., and Nursing, C. on C., Stroke, Biology, C. on F.G., Translational, Council, S., 2016. Importance of Assessing Cardiorespiratory Fitness in Clinical Practice: A Case for Fitness as a Clinical Vital Sign: A Scientific Statement From the American Heart Association. Circulation 134, e653–e699. https://doi.org/10.1161/CIR.0000000000000461

Sarzynski, M.A., Ghosh, S., Bouchard, C., 2017. Genomic and transcriptomic predictors of response levels to endurance exercise training. J. Physiol. 595, 2931–2939. https://doi.org/10.1113/JP272559

Sisson, S.B., Katzmarzyk, P.T., Earnest, C.P., Bouchard, C., Blair, S.N., Church, T.S., 2009. Volume of exercise and fitness nonresponse in sedentary, postmenopausal women. Med Sci Sports Exerc 41, 539–45. https://doi.org/10.1249/MSS.0b013e3181896c4e

Songsorn, P., Lambeth-Mansell, A., Mair, J.L., Haggett, M., Fitzpatrick, B.L., Ruffino, J., Holliday, A., Metcalfe, R.S., Vollaard, N.B.J., 2016. Exercise training comprising of single 20-s cycle sprints does not provide a sufficient stimulus for improving maximal aerobic capacity in sedentary individuals. Eur. J. Appl. Physiol. 116, 1511–1517. https://doi.org/10.1007/s00421-016-3409-8

Stocks, B., Dent, J.R., Ogden, H.B., Zemp, M., Philp, A., 2018. Postexercise skeletal muscle signaling responses to moderate- to high-intensity steady-state exercise in the fed or fasted state. Am. J. Physiol.-Endocrinol. Metab. 316, E230–E238. https://doi.org/10.1152/ajpendo.00311.2018

Sultana, R.N., Sabag, A., Keating, S.E., Johnson, N.A., 2019. The Effect of Low-Volume High-Intensity Interval Training on Body Composition and Cardiorespiratory Fitness: A Systematic Review and Meta-Analysis. Sports Med. 49, 1687–1721. https://doi.org/10.1007/s40279-019-01167-w

Swinton, P.A., Hemingway, B.S., Saunders, B., Gualano, B., Dolan, E., 2018. A Statistical Framework to Interpret Individual Response to Intervention: Paving the Way for Personalized Nutrition and Exercise Prescription. Front. Nutr. 5, 41. https://doi.org/10.3389/fnut.2018.00041

Thomas, G., Songsorn, P., Gorman, A., Brackenridge, B., Cullen, T., Fitzpatrick, B., Metcalfe, R.S., Vollaard, N.B.J., 2020. Reducing training frequency from 3 or 4 sessions/week to 2 sessions/week does not attenuate improvements in maximal aerobic capacity with reduced-exertion high-intensity interval training (REHIT). Appl. Physiol. Nutr. Metab. Physiol. Appl. Nutr. Metab. 1–3. https://doi.org/10.1139/apnm-2019-0750

Timmons, J.A., Knudsen, S., Rankinen, T., Koch, L.G., Sarzynski, M., Jensen, T., Keller, P., Scheele, C., Vollaard, N.B., Nielsen, S., Akerström, T., MacDougald, O.A., Jansson, E., Greenhaff, P.L., Tarnopolsky, M.A., van Loon, L.J., Pedersen, B.K., Sundberg, C.J., Wahlestedt, C., Britton, S.L., Bouchard, C., 2010. Using molecular classification to predict gains in maximal aerobic capacity following endurance exercise training in humans. J Appl Physiol 1985 108, 1487–96. https://doi.org/10.1152/japplphysiol.01295.2009

Vollaard, N.B., Constantin-Teodosiu, D., Fredriksson, K., Rooyackers, O., Jansson, E., Greenhaff, P.L., Timmons, J.A., Sundberg, C.J., 2009. Systematic analysis of adaptations in aerobic capacity and submaximal energy metabolism provides a unique insight into determinants of human aerobic performance. J Appl Physiol 1985 106, 1479–86. https://doi.org/10.1152/japplphysiol.91453.2008

Vollaard, N.B., Metcalfe, R.S., 2017. Research into the Health Benefits of Sprint Interval Training Should Focus on Protocols with Fewer and Shorter Sprints. Sports Med 47, 2443–2451. https://doi.org/10.1007/s40279-017-0727-x

Vollaard, N.B.J., Metcalfe, R.S., Williams, S., 2017. Effect of Number of Sprints in an SIT Session on Change in V̇O2max: A Meta-analysis. Med Sci Sports Exerc 49, 1147–1156. https://doi.org/10.1249/MSS.0000000000001204

Weston, M., Taylor, K.L., Batterham, A.M., Hopkins, W.G., 2014. Effects of low-volume high-intensity interval training (HIT) on fitness in adults: a meta-analysis of controlled and non-controlled trials. Sports Med 44, 1005–17. https://doi.org/10.1007/s40279-014-0180-z

Williams, C.J., Gurd, B.J., Bonafiglia, J.T., Voisin, S., Li, Z., Harvey, N., Croci, I., Taylor, J.L., Gajanand, T., Ramos, J.S., Fassett, R.G., Little, J.P., Francois, M.E., Hearon, C.M., Sarma, S., Janssen, S.L.J.E., Van Craenenbroeck, E.M., Beckers, P., Cornelissen, V.A., Pattyn, N., Howden, E.J., Keating, S.E., Bye, A., Stensvold, D., Wisloff, U., Papadimitriou, I., Yan, X., Bishop, D.J., Eynon, N., Coombes, J.S., 2019. A Multi-Center Comparison of O2peak Trainability Between Interval Training and Moderate Intensity Continuous Training. Front. Physiol. 10, 19. https://doi.org/10.3389/fphys.2019.00019

Figure Legends:

Figure 1 Changes in VO2max following SIT (panel A) and following a no-intervention control (panel B). Data is presented as mean and SD (bars) on the primary y-axis, or as individual change scores (black dots) on the secondary y-axis.

Figure 2 Individual changes in VO2max following SIT classified against either a zero (panel A) or clinically relevant thresholds of 1.75 ml·kg^{-1}·min^{-1} (light blue dashed and dotted line) and 3.5 ml·kg^{-1}·min^{-1} (dark blue dashed line) (panel B). Dots are individual changes and error bars are 50% confidence intervals. Red square = likely non-responder, orange diamond = uncertain responder, light blue circle = likely responder (>0 ml·kg^{-1}·min^{-1} in panel A and >1.75 ml·kg^{-1}·min^{-1} in panel B). In panel B, dark blue circle = likely responder >3.5 ml·kg^{-1}·min^{-1}. Pie charts show the absolute proportion (n) of participants in each category.

Conflict of Interest Statement: