Volume 22 Issue 11 November 2022 ISSN 1746-1391

European Journal of Sport Science

European Journal of Sport Science
The Official Journal of the European College of Sport Science

Routledge
Taylor & Francis Group

Routledge

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tejs20

Exploring interindividual differences in fasting and postprandial insulin sensitivity adaptations in response to sprint interval exercise training

Richard S. Metcalfe, Brendon J. Gurd & Niels B. J. Vollaard

To cite this article: Richard S. Metcalfe, Brendon J. Gurd & Niels B. J. Vollaard (2022): Exploring interindividual differences in fasting and postprandial insulin sensitivity adaptations in response to sprint interval exercise training, European Journal of Sport Science, DOI: 10.1080/17461391.2022.2124385

To link to this article: https://doi.org/10.1080/17461391.2022.2124385

EUROPEAN JOURNAL OF SPORT SCIENCE
https://doi.org/10.1080/17461391.2022.2124385

Routledge
Taylor & Francis Group

Exploring interindividual differences in fasting and postprandial insulin sensitivity adaptations in response to sprint interval exercise training

Richard S. Metcalfe [a], Brendon J. Gurd[b] and Niels B. J. Vollaard[c]

[a]Applied Sports, Technology, Exercise and Medicine (A-STEM) Research Centre, Faculty of Science and Engineering, Swansea University, Swansea, UK; [b]School of Kinesiology and Health Studies, Queen's University, Kingston, Canada; [c]Faculty of Health Sciences and Sport, University of Stirling, Stirling, UK

ABSTRACT

Previous studies have concluded that wide variance in changes in insulin sensitivity markers following exercise training demonstrates heterogeneity in individual trainability. However, these studies frequently don't account for technical, biological, and random within-subject measurement error. We used the standard deviation of individual responses (SD_{IR}) to determine whether interindividual variability in trainability exists for fasting and postprandial insulin sensitivity outcomes following low-volume sprint interval training (SIT). We pooled data from 63 untrained participants who completed 6 weeks of SIT ($n = 49$; VO_2max: 35 (7) $mL \cdot kg^{-1} \cdot min^{-1}$) or acted as no-intervention controls ($n = 14$; VO_2max: 34 (6) $mL \cdot kg^{-1} \cdot min^{-1}$). Fasting and oral glucose tolerance test (OGTT)-derived measures of insulin sensitivity were measured pre- and post-intervention. SD_{IR} values were positive and exceeded a small effect size threshold for changes in fasting glucose ($SD_{IR} = 0.27$ [95%CI 0.07,0.38] $mmol \cdot L^{-1}$), 2-h OGTT glucose ($SD_{IR} = 0.89$ [0.22,1.23] $mmol \cdot L^{-1}$), glucose area-under-the-curve ($SD_{IR} = 66.4$ [−81.5,124.3] $mmol \cdot L^{-1} \cdot 120min^{-1}$) and The Cederholm Index ($SD_{IR} = 7.2$ [−16.0,19.0] $mg \cdot l^2 \cdot mmol^{-1} \cdot mU^{-1} \cdot min^{-1}$), suggesting meaningful individual responses to SIT, whilst SD_{IR} values were negative for fasting insulin, fasting insulin resistance and insulin AUC. For all variables, the 95% CIs were wide and/or crossed zero, highlighting uncertainty about the existence of true interindividual differences in exercise trainability. Only 2–22% of participants could be classified as responders or non-responders with more than 95% certainty. Our findings demonstrate it cannot be assumed that variation in changes in insulin sensitivity following SIT is attributable to inherent differences in trainability, and reiterate the importance of accounting for technical, biological, and random error when examining heterogeneity in health-related training adaptations.

KEYWORDS
Exercise; training; health; metabolism; physiology

Highlights

- This study tested whether true interindividual variability exists for changes in insulin sensitivity and glyceamic control following 6-weeks of low volume sprint interval training (SIT).
- The high level of technical, biological, and random error associated with repeated measurements of insulin sensitivity and glycaemic control, means we can neither confidently conclude that there is evidence of true interindividual differences in the trainability of these outcomes following SIT, nor confidently identify responders or non-responders for such parameters.
- Researchers contrasting responders vs. non-responders for a given parameter, either to understand mechanisms of adaptation and/or develop physiological/genetic/epigenetic predictors of response, need to be aware that identification of responders and non-responders with sufficient certainty may not be achievable for parameters with a high level of technical, biological, and random error.

1. Introduction

Exercise training improves various biomarkers of cardiometabolic health on average, but at an individual level, the magnitude of observed change can be highly variable (Wilmore et al., 2001). For example, the HERITAGE family study reported a 10% mean increase in postprandial insulin sensitivity following 24 weeks of aerobic training, yet ~40% of individuals showed either no change or a numerical *decrease* of varying magnitude

CONTACT Richard S. Metcalfe r.s.metcalfe@swansea.ac.uk Applied Sports, Technology, Exercise and Medicine (A-STEM) Research Centre, Faculty of Science and Engineering, Swansea University, Swansea SA1 8EN, UK

(Boulé et al., 2005). Similar variability was observed for changes in blood pressure, blood lipids, fasting insulin sensitivity, and body composition (Wilmore et al., 2001). These observations have often been taken to infer the presence of inherent interindividual differences in exercise "trainability" (i.e. how much an individual improves in response to exercise) and used as a justification to explore its physiological determinants (Sparks, 2017). However, many studies that present and evaluate individual responses to exercise interventions fail to account for technical, biological, and random within-subject variability in their outcome through the inclusion of a no-intervention control group (Williamson et al., 2017). The limitations of making inferences on interindividual variability in response to an intervention in the absence of a control condition are now well established and, therefore, for many outcomes, the presence of true interindividual responses to exercise training requires reappraisal (Atkinson & Batterham, 2015; Bonafiglia et al., 2019; Williamson et al., 2017).

Establishing whether there are true interindividual differences in the response to different types of exercise is important because it provides the basis (or lack of) for further studies aiming to contrast responders vs. non-responders to interventions to elucidate mechanisms of adaptation and to develop physiological/genetic/epigenetic predictors and/or moderators/mediators of individual response (Atkinson & Batterham, 2015). This in turn will be a vital step to enable precision exercise, which aims to prescribe individually tailored exercise interventions to optimise improvements in health (Atkinson et al., 2018). If there is no evidence of individual responses beyond the variation caused by technical, biological, and random error, then such objectives will not be appropriate or achievable. It is possible to quantify the likelihood that a given individual participant in a study did or did not respond to the intervention, but if too many individuals are incorrectly labelled as responders or nonresponders then this will invalidate the analysis. Furthermore, true interindividual variation in response to exercise on a group level can be established by calculating the standard deviation of individual responses (SD_{IR}), which compares the variation (i.e. SD) in response to an (exercise) intervention with the "natural" variation of change in a control condition where no intervention has been implemented (Atkinson & Batterham, 2015; Hopkins, 2015). A positive SD_{IR} indicates that variability in response to the intervention exceeds the variation of change in the control condition (Atkinson & Batterham, 2015; Hopkins, 2015). If the 95% CI of this SD_{IR} does not span zero, this indicates that true interindividual responses to the intervention may exist. Conversely, small positive or negative SD_{IR} values with trivial effect sizes reflect the absence of a meaningful difference between the variability in the control group and the intervention group, and therefore offer no support for the existence of true interindividual variability in training response. A negative SD_{IR} with a 95% CI that does not span zero may either indicate a potential "homogenising" effect of the intervention, or additional sources of error in the control group that were not – or were to a lesser extent – present in the experimental group (Atkinson & Batterham, 2015; Hopkins, 2015).

The SD_{IR} has previously been employed to provide evidence of true interindividual differences in trainability of mean arterial blood pressure and maximal oxygen uptake following high-intensity interval or sprint interval training (Phillips et al., 2017; Metcalfe & Vollaard, 2021). On the other hand, recent meta-analyses have not been able to find evidence of meaningful individual differences in the trainability of either maximal oxygen uptake (Bonafiglia et al., 2021) or blood pressure (Kelley et al., 2020) following aerobic exercise training. In addition, a mixture of mostly trivial/small positive and negative SD_{IR} values have been reported for changes in body composition (Hammond et al., 2019; Walsh et al., 2019) and, overall, meta-analyses find no meaningful interindividual variation in weight and fat loss after training (Bonafiglia et al., 2021; Kelley et al., 2021, 2020; Williamson et al., 2018). Similarly, Bonafiglia et al. (2021) were unable to demonstrate consistent interindividual differences for various molecular adaptations in skeletal muscle.

An improvement in insulin sensitivity is a classic health-related adaptation observed following various modes of exercise including low volume sprint interval training (Babraj et al., 2009; Gillen et al., 2016; Hawley & Gibala, 2012; Metcalfe et al., 2012). However, although several studies have alluded to interindividual differences in changes in insulin sensitivity following exercise training (Álvarez et al., 2017; Brennan et al., 2020; Gray et al., 2018), only one study has utilised the SD_{IR} to explore whether there is evidence of *true* interindividual variability in response beyond observed technical, biological, and random within-subjects variability, and this was limited to fasting glucose concentrations (Walsh et al., 2019). No studies have employed the SD_{IR} to explore changes in postprandial measures of insulin sensitivity and glycaemic control, and no studies have attempted to quantify or report the likelihood of individual participants being responders or non-responders for these parameters. Furthermore, no study has examined interindividual responses for insulin sensitivity indices in response to sprint interval training (SIT) interventions with very low exercise volume. Thus, the aim of this paper was to quantify interindividual variability in

changes of fasting and postprandial insulin sensitivity in response to 6 weeks of low-volume, time-efficient SIT.

2. Methods

2.1. Participants/sample

The sample for this exploratory analysis was derived from two previously published studies conducted in our laboratories (Metcalfe et al., 2016, 2012). In total, this included 49 pooled training participants (23 male, 26 female; age: 31 (10) y; Body Mass Index (BMI): 24 (3) $kg \cdot m^{-2}$; VO_2max: 34.9 (7.1) $mL \cdot kg^{-1} \cdot min^{-1}$) who underwent 6 weeks of supervised laboratory-based SIT and 14 no intervention control participants (6 male, 8 female; age: 20 (2) y; BMI: 24 (4) $kg \cdot m^{-2}$, VO_2max: 33.8 (5.5) $mL \cdot kg^{-1} \cdot min^{-1}$) who completed two oral glucose tolerance tests (OGTTs) over a similar time frame. The control participants were drawn from one study (Metcalfe et al., 2012) whilst the training participants were pooled from both studies (Metcalfe et al., 2016, 2012). The inclusion/exclusion criteria were similar across both studies. All participants were low/moderately active based on the International Physical Activity Questionnaire (Craig et al., 2003). Anyone with contraindications to high-intensity exercise were excluded. Ethical approval was granted for both studies (approval references are available in the original articles (Metcalfe et al., 2016, 2012)) and all participants provided informed consent.

2.2. Oral glucose tolerance testing (OGTT)

Indices of insulin sensitivity and glycaemic control were assessed pre- and post-training using an OGTT. For participants in the exercise training group, the pre-training OGTTs were performed two weeks prior to the first training session, whilst post-training OGTTs took place 3 days following the final training session. For participants in the no-intervention control group, the two OGTTs were performed approximately 8 weeks apart and they were asked to maintain their current diet and physical activity patterns in between. A gap of 8 weeks was left between the two OGTTs to try and ensure that female participants in both groups would be in a similar stage of their menstrual cycle; however, this was not formally assessed.

To remove the confounding effect of acute physical activity on insulin action, participants were asked to avoid moderate-to-vigorous physical activity in the 3 days prior to the OGTT. In addition, they were asked to avoid alcohol for 1 d prior. The OGTTs were performed in the morning between 07:30 and 09:30 am following an overnight fast from 10:00 pm the previous evening. On arrival to the lab, participants rested quietly for 10 min, and then a fasting venous blood sample was obtained from an antecubital vein. Further venous blood samples were then drawn 60 and 120 min after participants consumed a drink containing 75 g of glucose. Although further blood samples at 30 and 90 min were collected and are reported in one study (Metcalfe et al., 2016), these were excluded here to ensure consistency in the calculation of outcome measures such as area-under-the-curves (AUCs) across participants in this analysis. Blood samples were collected into precooled plastic tubes containing ethylenediaminetetraacetic acid, centrifuged to separate the plasma, frozen and stored, and then analysed for plasma glucose and insulin concentrations. The full protocol of the OGTTs is available in the original manuscripts (Metcalfe et al., 2016, 2012). We derived several different indices of glycaemic control and insulin sensitivity from the OGTTs. Fasting-based measures were fasting glucose, fasting insulin, and the homeostasis model of insulin resistance (HOMA2-IR) (Wallace et al., 2004). Postprandial or dynamic measures were the total AUCs for the glucose and insulin response to the glucose load (calculated using the trapezoid rule), the OGTT 2-h glucose concentration, and the Cederholm Index of insulin sensitivity (Cederholm & Wibell, 1990).

2.3. Sprint interval training

Training group participants completed a 6-week supervised cycling-based SIT intervention (reduced-exertion high-intensity interval training (REHIT)) with a frequency of 3 sessions per week (18 sessions in total). Each training session consisted of a 3-min low intensity (60 W or unloaded pedalling) warm up, followed by one (first session) or two (all other sessions) all-out sprints interspersed with low-intensity recovery. The all-out sprints were performed against a fixed resistance of 5% (Metcalfe et al., 2016) or 7.5% (Metcalfe et al., 2012) of participants' pre-training body mass and increased in duration from 10 s in week 1, to 15 s in weeks 2 and 3, and finally to 20 s in the final 3 weeks. For each sprint, participants increased their cadence to their maximal in the 3 s prior, the resistance was applied to the flywheel, and then they maintained the highest possible cadence for the duration of the sprint. The sessions were performed on a Monark friction-braked cycle ergometer (Monark, Vansbro, Sweden).

2.4. Statistical analysis

To assess the group level changes in all insulin sensitivity and glycaemic control variables following SIT, a two-way

repeated measures analysis of variance (ANOVA) (group × time) was performed in Graphpad Prism 8 for macOS (Version 8.4.2, San Diego, Cal, USA). The main statistic of interest was the group × time interaction. Alpha was set at $p<0.05$.

The SD_{IR} was calculated from the square root of the difference between the square of the SD of the change in the exercise training group (SDex) and the control group (SDcon) (Atkinson & Batterham, 2015):

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

The 95% confidence intervals (CIs) for the SD_{IR} were calculated using the method described by Hopkins (2015). In instances where the SD_{CON} was greater than the SD_{EX}, the SD_{IR} equation was reversed and the SD_{IR} is reported as a negative value (Hopkins, 2015). Standardised effect sizes were calculated by dividing the SD_{IR} by the standard deviation of the baseline scores in the control condition (Hopkins, 2015). The thresholds for interpreting the effect size were 0.1 (small), 0.3 (moderate), 0.6 (large), 1.0 (very large) and 2.0 (extremely large) (Hopkins, 2015).

Individual responses were classified using 95% CIs derived from the typical error of measurement, which was calculated using the formula recommended by Swinton et al. (2018):

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

Responses were then classified against a small effect size of 0.2 x baseline SD of the control group. If the entire 95% CI lay beyond a small beneficial or a small harmful effect size then they were classified as a "responder" or "adverse responder", respectively (Bonafiglia et al., 2018). If the full 95% CI lay within the cut point for a small beneficial effect (and didn't lie beyond a small adverse effect) then they were classified as a "non-responder", and if it crossed the cut point for a small beneficial effect the response was classified as "uncertain" (Bonafiglia et al., 2018).

3. Results

3.1. Group effects

For fasting glucose, fasting insulin, HOMA2-IR and the OGTT 2 h glucose, there were no differences in the change over time between the SIT group and the control group ($p>0.05$ for all group×time interaction effects; Table 1 and Figure 1). For glucose AUC, insulin AUC and the Cederholm Index, there were significant group×time interaction effects suggesting an improved response in the SIT group compared to the control group (all $p<0.05$; Table 1 and Figure 1).

3.2. Interindividual responses

We observed a positive SD_{IR} for fasting glucose, the OGTT 2 h glucose, glucose AUC and the Cederholm Index; however, in each case, the 95% confidence intervals were wide and for glucose AUC and the Cederholm index they spanned zero (Table 1 and Figure 1). The effect size of the SD_{IR} would be considered *small* for glucose AUC and the Cederholm Index ($d = 0.29$ and 0.27, respectively), *moderate* for OGTT 2 h glucose ($d = 0.58$), and *large* for fasting glucose ($d = 0.73$) (Table 1 and Figure 1). On the other hand, we observed a negative SD_{IR} for fasting insulin, HOMA2-IR, and insulin AUC (Table 1 and Figure 1), demonstrating that the variability in the control group responses exceeded the variability in the SIT group responses. The inferences for fasting insulin and HOMA2-IR were unaltered when a potentially influential control participant with a −80.3 pmol·L^{-1} change in fasting insulin was removed from the analysis: the SD_{IR} remained negative for both fasting insulin ($SD_{IR} = -6.34$, 95% CI: −16.3,13.8 pmol·L^{-1}, $d = -0.17$) and HOMA2-IR ($SD_{IR} = -0.10$, 95% CI: −0.33,0.31 arbitrary units, $d = -0.12$).

For all variables, only a very small number of participants could confidently be classified as either a "responder" or a "non-responder", with the majority of responses being classified as "uncertain" (Figure 2). The proportion of "uncertain" responses was 78% for fasting glucose, 81% for 2-h glucose, 92% for glucose AUC, 94% for insulin AUC, 96% for the Cederholm Index, 98% for HOMA2-IR and 98% for fasting insulin (Figure 2).

4. Discussion

This study examined whether meaningful interindividual variability in changes of fasting and OGTT-derived insulin sensitivity occurs in response to low-volume, time-efficient SIT. We found a negative SD_{IR} for changes in fasting insulin, fasting insulin resistance, and for the insulin response to the OGTT – evidence that no meaningful interindividual variability in training response exists for these outcomes. On the other hand, we found a positive SD_{IR} for fasting glucose, the glucose response to the OGTT, and postprandial insulin sensitivity. However, in each case, the 95% confidence intervals were wide and/or spanned zero, meaning that a high level of caution is warranted about the occurrence of additional variability in the SIT group compared to the control group. This is particularly true for the variables where we found no evidence of a mean training effect following SIT (fasting glucose, OGTT 2 h glucose), as true *variability* in training response in the absence of

Table 1. Mean changes and interindividual variability for changes in fasting and postprandial insulin sensitivity and glycaemic control.

Outcome	SIT Baseline	Control Delta	*p*-value (interaction) Baseline	SD_{IR} [95% CI limits] Delta	Effect size [95% CI limits]		
Fasting Glucose ($mmol \cdot L^{-1}$)	5.03 (0.45)	0.00 (0.36)	5.22 (0.38)	0.09 (0.24)	0.394	0.27 [0.07, 0.38]	0.73 [0.19, 1.01]
Fasting Insulin ($pmol \cdot L^{-1}$)	44.5 (29.6)	0.95 (14.87)	84.6 (37.4)	−9.47 (25.62)	0.058	−20.9 [−30.8, 8.8]	−0.56 [−0.82, 0.23]
HOMA2-IR (AU)	0.98 (0.63)	0.02 (0.33)	1.83 (0.81)	−0.19 (0.54)	0.069	−0.43 [−0.64, 0.21]	−0.53 [−0.79, 0.26]
OGTT 2-h Glucose ($mmol \cdot L^{-1}$)	5.43 (1.27)	0.02 (1.18)	6.02 (1.54)	0.61 (0.79)	0.096	0.89 [0.22, 1.23]	0.58 [0.14, 0.80]
Glucose AUC ($mmol \cdot L^{-1} \cdot 120min^{-1}$)	749.7 (183.1)	−30.7 (126.8)	753.7 (232.7)	75.1 (108.0)	0.006*	66.4 [−81.45, 124.3]	0.29 [−0.35, 0.53]
Insulin AUC ($pmol \cdot L^{-1} \cdot 120min^{-1}$)	42348 (29706)	−4851 (18495)	45530 (25644)	7366 (21804)	0.041*	−11548 [−22883, 16028]	−0.45 [−0.89, 0.63]
Cederholm Index ($mg \cdot L^{-2} \cdot mmol^{-1} \cdot mU^{-1} \cdot min^{-1}$)	71.7 (27.0)	3.3 (19.9)	63.4 (26.7)	−9.9 (18.5)	0.029*	7.2 [−16.0, 19.0]	0.27 [−0.60, 0.71]

Data for baseline and Δ SIT and Δ CTRL is presented as Mean (SD).
* Denotes $p<0.05$ for group × time interaction effect.

evidence for a *mean* training response would require the (unlikely) existence of similar proportions of true adverse and favourable responders to exercise. It was only possible to classify between 0-22% of participants as either a responder or non-responder to the intervention with more than 95% certainty, i.e. for all outcomes the responder status of most participants was uncertain. Taken together, the high level of technical, biological, and random error associated with performing repeated measurements of both fasting and OGTT-derived insulin sensitivity over time, means that we can neither confidently conclude that there is evidence of true inter-individual differences in the trainability of these outcomes following low-volume SIT, nor confidently identify individual participants who did or did not respond to the intervention, which are pre-requisites for being able to progress to studies aiming to contrast responders vs. non-responders to elucidate mechanisms of adaptation or develop predictors of the response to exercise.

The indicators of insulin sensitivity measured in our study have been used in several studies that report inter-individual responses to various types of exercise training (Álvarez et al., 2017; Álvarez et al., 2018; Brennan et al., 2020; Gray et al., 2018; Lannoy et al., 2017; Wilmore et al., 2001). Without exception these studies have ascribed the variability in their exercise training arm to inherent differences in individual trainability and/or reported rates of responders/non-responders to the training intervention (Álvarez et al., 2017; Álvarez et al., 2018; Brennan et al., 2020; Gray et al., 2018; Lannoy et al., 2017; Wilmore et al., 2001). Our findings suggest these conclusions may be premature. High levels of measurement error associated with repeated measures of insulin sensitivity inflate the variability in responses in the intervention group, and this variability should always be corrected by comparing it to the variability present in a suitable control group. To illustrate this point, we reanalysed the publicly available dataset from a paper by Lannoy and colleagues (Lannoy et al., 2017). This study in 171 overweight/sedentary adults examined variability in the change in fasting insulin, 2 h OGTT glucose concentrations, and the insulin AUC during an OGTT, following 6 months of different doses of aerobic exercise training (low amount, low intensity (LALI); high amount, low intensity (HALI); and high amount, high intensity (HAHI)) and in a no-intervention control group (Lannoy et al., 2017). The intensity of LALI and HALI were matched (∼50% HR_{max}) and different from HAHI (∼75% HR_{max}), whilst the volume of HALI and HAHI were matched (∼360 kcals for females and ∼600 kcals for males) and approximately double that of the LALI group (∼180 kcals for females and ∼300 kcals for males) (Lannoy et al., 2017). All groups were prescribed a frequency of 5 sessions/week (Lannoy et al., 2017). The SD_{IR} and 95% CIs (standardised to baseline variation in the controls) for each group are shown in Figure 3 with our low-volume SIT data shown for comparison. The data show consistently negative SD_{IR} scores for fasting insulin, whilst for insulin AUC and OGTT 2 h glucose concentrations there is a mixture of positive and negative SD_{IR} estimates with no clear pattern in favour of one of the exercise conditions (Figure 3). Crucially, for all variables in all exercise conditions, the confidence limits around the SD_{IR} are wide and in many cases cross zero (Figure 3).

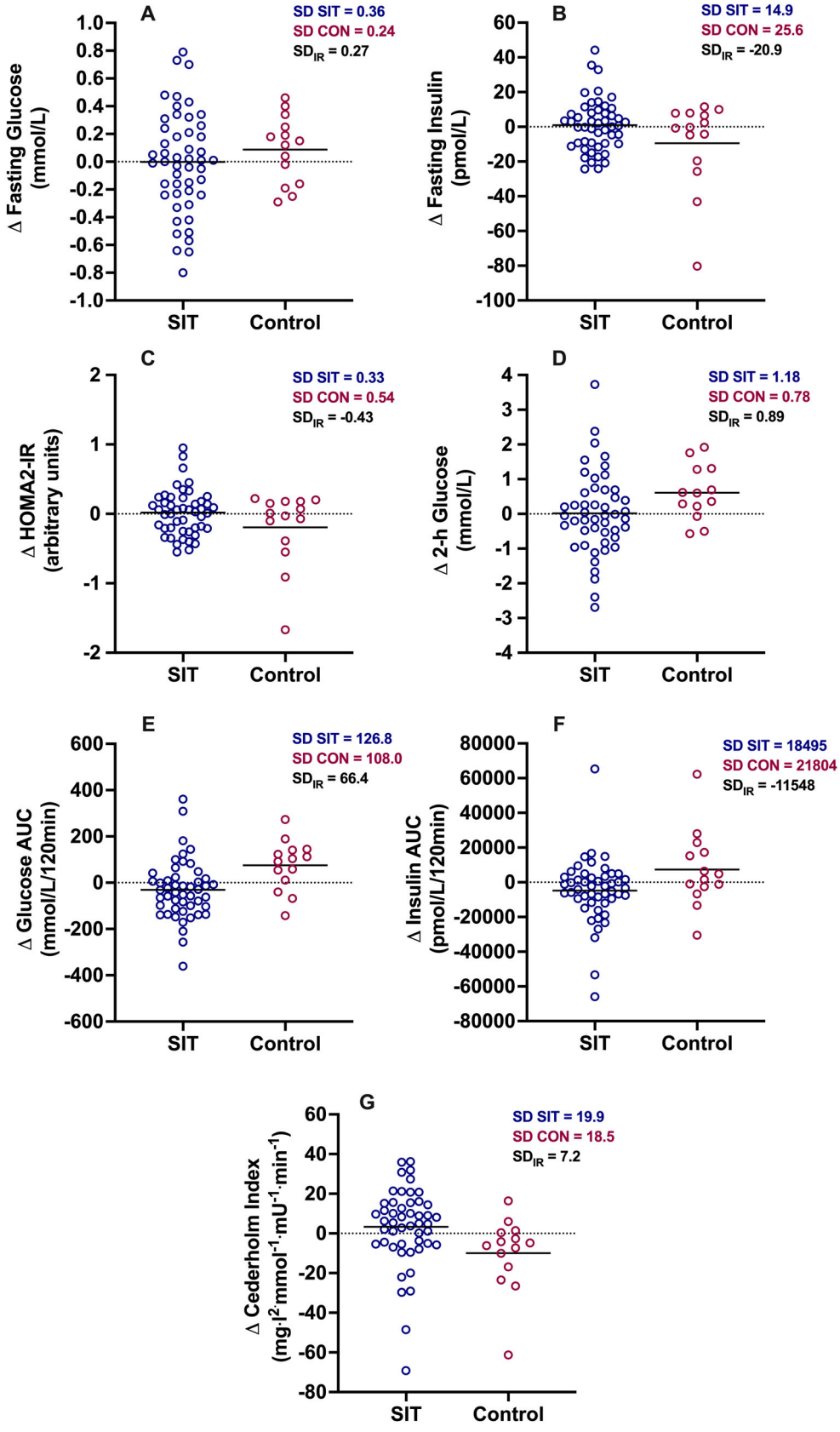

Figure 1. Individual changes in fasting glucose (A), fasting insulin (B), HOMA2-IR (C), OGTT 2-h glucose concentration (D), glucose AUC (E), insulin AUC (F) and The Cederholm Index (G) following 6-weeks of SIT (blue dots) or no-intervention control (dark red dots). Each dot represents the individual training response in original measured units (post training score minus pre training score). The solid lines represent the mean change.

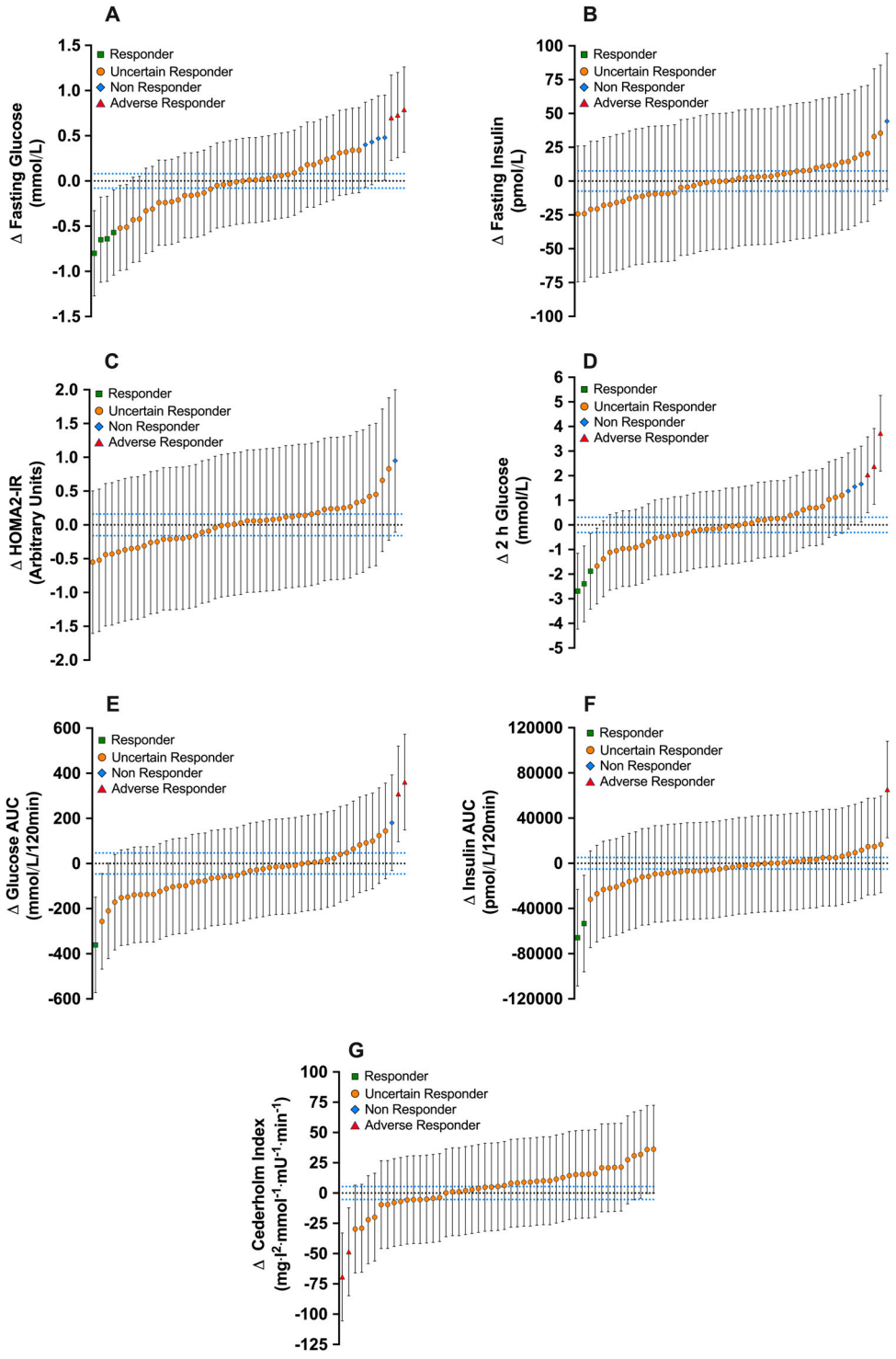

Figure 2. Classification of responders (green squares), non-responders (blue diamonds), uncertain responders (orange circles), and adverse responders (red triangles) following SIT for changes in fasting glucose (A), fasting insulin (B), HOMA2-IR (C), 2 h OGTT glucose concentrations (D), glucose AUC (E), insulin AUC (F) and The Cederholm Index (G). Dots are individual changes and error bars are 95% confidence intervals derived from the technical error. Responses are classified against a small effect size of 0.2 × the baseline variation in the control condition (blue dotted lines). A "responder" was classified when the entire CI lay beyond a small beneficial effect; "adverse responder" was classified when the entire CI lay beyond a small harmful effect; a "non-responder" was classified when the 95% lay within a small beneficial effect (but not beyond a small harmful effect); if the 95% CI lay across the cut point for a small beneficial effect, then the response was classified as "uncertain".

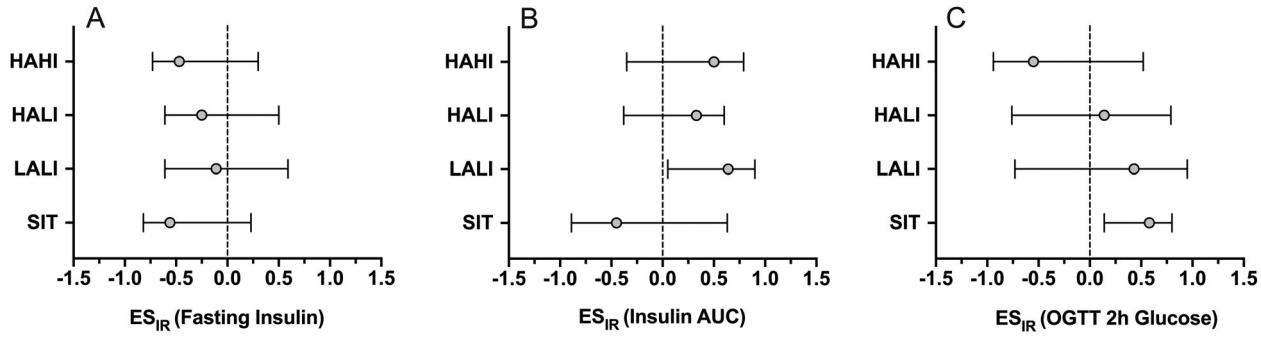

Figure 3. SD$_{IR}$ estimates for changes in fasting insulin (A), insulin AUC during an OGTT (B) and 2 h OGTT glucose concentrations (C) in response to 6 months of different intensities and volumes of aerobic exercise training (secondary analysis of Lannoy et al. (2017)) and following 6 weeks of SIT (present study). SD$_{IR}$ scores are presented as effect sizes standardised to baseline variation in the respective control condition (grey dots) with 95% confidence intervals (error bars). SIT, sprint interval training; LALI, low amount low intensity; HALI, high amount low intensity; HAHI, high amount, high intensity; ESIR, effect size of individual responses.

Accordingly, at most 14% of participants could be classified as a responder or non-responder with 95% certainty for any of the parameters (Figure 4). This contrasts with the terminology used by the authors, who classify many of the unclear responders as non-responders (Lannoy et al., 2017). This re-analysis of previous data demonstrates that our findings following low-volume SIT are not exceptional; even following prolonged, high volume, high frequency aerobic exercise interventions, it is not possible to confidently and consistently detect meaningful interindividual differences in the trainability of indices of fasting or postprandial insulin

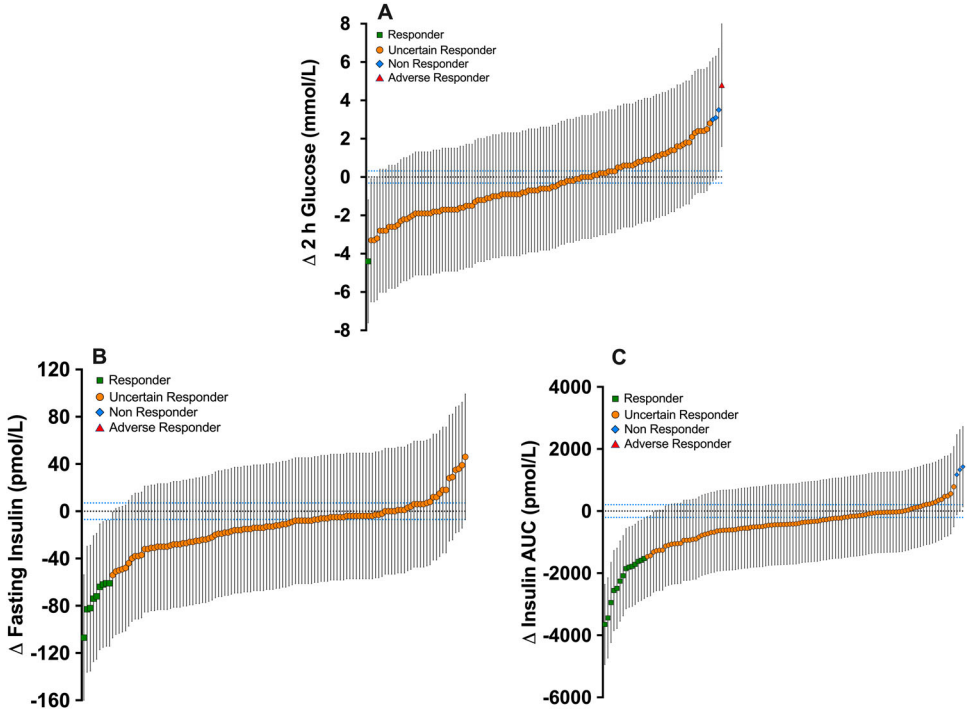

Figure 4. Reanalysis of Lannoy et al. (2017) showing classification of responders (green squares), non-responders (blue diamonds), uncertain responders (orange circles), and adverse responders (red triangles) following 24 weeks of aerobic exercise for changes in 2 h OGTT glucose (A), fasting insulin (B) and insulin AUC (C). Dots are individual changes and error bars are 95% confidence intervals derived from the technical error. Responses are classified against a small effect size of 0.2 x the baseline variation in the control condition (blue dotted lines). A "responder" was classified when the entire CI lay beyond a small beneficial effect; "adverse responder" was classified when the entire CI lay beyond a small harmful effect; a "non-responder" was classified when the 95% lay within a small beneficial effect (but not beyond a small harmful effect); if the 95% CI lay across the cut point for a small beneficial effect, then the response was classified as "uncertain". Note that data from three different exercise training doses (low amount low intensity; high amount low intensity; high amount, high intensity) have been combined in one figure for simplicity.

sensitivity, or to identify clear responders and non-responders.

We observed a positive SD_{IR} for glucose AUC and the Cederholm Index with a small effect size, whilst for fasting glucose and the OGTT 2 h glucose concentration, the effect sizes were large and moderate, respectively (Hopkins, 2015). This could be interpreted as evidence of additional variability in the SIT group compared to the control group. Yet, notwithstanding the high level of uncertainty surrounding these SD_{IR} estimates, it should also be highlighted that the standardised effect sizes used to interpret the magnitude of the SD_{IR} scores are not anchored to any clinical outcome (e.g. reduced disease risk) and, even if we assume that the positive SD_{IR} reflects true differences in the effect of low-volume SIT on these variables, it remains unclear how relevant the observed interindividual differences would be for long term health. Interpreting the magnitude of the SD_{IR} against an evidence-based minimal clinically important difference (MCID) may be informative, but we are unaware of appropriate thresholds for these indices of insulin sensitivity in our population of young, healthy but inactive adults, or indeed, any other population. Moreover, even for parameters with a large effect size for group-level interindividual differences, it may not be possible to confidently identify individuals with a response smaller or greater than a given MCID.

It is curious that, except for the Cederholm Index, we observed positive SD_{IR} values for glucose-related variables, whereas we found negative SD_{IR} estimates for variables more dependent on insulin concentrations. This is particularly true given the effect size for the negative SD_{IR} estimates would be considered moderate (Hopkins, 2015). If there is indeed no difference in the variability in the SIT group and the control group, then SD_{IR} scores with trivial effect sizes would be more likely to occur, whereas a meaningful and moderate effect size would be less likely and therefore could be inferred to (weakly) suggest an additional source of variability in the control group compared to the training group. However, a more likely explanation in our study is the smaller sample size in the control group compared to the SIT group, combined with an overall higher measurement error for insulin-related variables over time. With smaller sample sizes, the SD will be more heavily influenced by extreme individual change scores, which appear to be more likely for insulin compared with glucose indices. Our sensitivity analysis demonstrates this point: when we removed the most extreme change score for fasting insulin and HOMA2-IR from the control group, in each case the SD_{IR} remained negative, but the effect size became borderline trivial/small.

One of the main goals of studying interindividual variability in training response is to try and pinpoint a specific individuals' likelihood of responding (or not responding) to training. Indeed, several previous studies have classified individuals as responders/non-responders for indices of insulin sensitivity in the absence of a non-exercise control group (Álvarez et al., 2017; Brennan et al., 2020; Gray et al., 2018). However, this approach may be erroneous, and these analyses should be treated with caution (Atkinson & Batterham, 2015). Using the methods employed in our analysis it is possible to estimate the proportions of responders, non-responders, and unclear responders, taking into account the technical error of measurement. The proportion of unclear responders will always be high for parameters with high technical error (such as insulin-related parameters). Thus, studies aiming to contrast responders vs. non-responders should be limited to parameters with low technical error.

There are important limitations to the current analysis. Firstly, the SD_{IR} assumes that, apart from the exercise group undergoing exercise training, all other sources of variability between the exercise and control groups are identical (Atkinson & Batterham, 2015). This is a pooled analysis of two training studies with minor methodological differences, with the (smaller) control group drawn from only one of these studies (Metcalfe et al., 2012) and the SIT group pooled from both (Metcalfe et al., 2016, 2012). Although this pooling approach is common in the individual response literature to achieve larger sample sizes (Gurd et al., 2016; Metcalfe & Vollaard, 2021; Williams et al., 2019), it is possible that the validity of the SD_{IR} estimate could be affected. Our reanalysis of the Lannoy et al. (2017) randomised controlled trial, which provided support for our conclusions following SIT (Figures 3 and 4), can partly assuage these concerns but, nevertheless, our analysis should be treated with some caution. Secondly, as pointed out by Bonafiglia et al. (2021, 2019), the lack of treatment blinding in randomised controlled trials of exercise training may introduce differential behavioural changes in exercise and control group participants and this could also impact the validity of the SD_{IR}. Thirdly, we did not perform a formal sample size calculation because of the analysis approach and the lack of a meaningful, evidence based MCID. However, *post hoc* power analysis demonstrates that to detect a difference between the SD_{EX} and SD_{CON} using a conventional independent sample t-test with effect sizes of 0.3, 0.6, or 1.0 would require sample sizes of 352, 90, or 34 respectively. Although our study may therefore be underpowered, the consistent findings between our study ($n = 61$) and the reanalysis of Lannoy et al. (2017) ($n = 171$), indicates

that the key conclusions from our study (e.g. quantification of "true" interindividual variability and identification of responders and non-responders with sufficient certainty may not be achievable for parameters with high error) would be the same for a study with a larger sample size.

In conclusion, this study demonstrates that variability in training-induced changes in both fasting and OGTT-derived insulin sensitivity following low-volume SIT (and potentially exercise training in general) is largely reflective of high levels of measurement error. These findings add to the growing body of literature highlighting that the observed variability in an exercise training group is not necessarily reflective of inherent differences in the individual responsiveness to training (Kelley et al., 2021, 2020; Walsh et al., 2019; Williamson et al., 2018). Future research investigating heterogeneity in insulin sensitivity and other health-related adaptations to exercise training should include a no-intervention control condition and adopt appropriate statistical frameworks that account for technical, biological, and random noise.

Acknowledgements

We would like to thank all participants who took part in this research.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Richard S. Metcalfe http://orcid.org/0000-0003-0980-2977

References

Álvarez, C., Ramírez-Campillo, R., Ramírez-Vélez, R., & Izquierdo, M. (2017). Effects and prevalence of nonresponders after 12 weeks of high-intensity interval or resistance training in women with insulin resistance: A randomized trial. *Journal of Applied Physiology*, *122*(4), 985–996. https://doi.org/10.1152/japplphysiol.01037.2016

Álvarez, C., Ramírez-Vélez, R., Ramirez-Campillo, R., Shigenori, I., Celis-Morales, C., García-Hermoso, A., … Izquierdo, M. (2018). Inter-individual responses to different exercise stimuli among insulin-resistant women. *Scandinavian Journal of Medicine & Science in Sports*, *28*, 2052–2065. https://doi.org/10.1111/sms.13213.

Atkinson, G., & Batterham, A. M. (2015). True and false interindividual differences in the physiological response to an intervention: True and false interindividual differences. *Experimental Physiology*, *100*(6), 577–588. https://doi.org/10.1113/EP085070

Atkinson, G., Williamson, P., & Batterham, A. M. (2018). Exercise training response heterogeneity: Statistical insights. *Diabetologia*, *61*(2), 496–497. https://doi.org/10.1007/s00125-017-4501-2

Babraj, J. A., Vollaard, N. B., Keast, C., Guppy, F. M., Cottrell, G., & Timmons, J. A. (2009). Extremely short duration high intensity interval training substantially improves insulin action in young healthy males. *BMC Endocrine Disorders*, *9*(3). https://doi.org/10.1186/1472-6823-9-3.

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. *Physiological Reports*, *7*(14), e14163. https://doi.org/10.14814/phy2.14163.

Bonafiglia, J. T., Islam, H., Preobrazenski, N., Ma, A., Deschenes, M., Erlich, A. T., … Gurd, B. J. (2021). Examining interindividual differences in select muscle and whole-body adaptations to continuous endurance training. *Experimental Physiology*, *101*(11), 2168–2176. https://doi.org/10.1113/EP089421.

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. *Physiological Reports*, *6*(22), e13928. https://doi.org/10.14814/phy2.13928

Boulé, N. G., Weisnagel, S. J., Lakka, T. A., Tremblay, A., Bergman, R. N., Rankinen, T., Leon, A. S., Skinner, J. S., Wilmore, J. H., Rao, D. C., Bouchard, C., & Study, H. F. (2005). Effects of exercise training on glucose homeostasis: The HERITAGE family study. *Diabetes Care*, *28*(1), 108–114. https://doi.org/10.2337/diacare.28.1.108

Brennan, A. M., Standley, R. A., Yi, F., Carnero, E. A., Sparks, L. M., & Goodpaster, B. H. (2020). Individual response variation in the effects of weight loss and exercise on insulin sensitivity and cardiometabolic risk in older adults. *Frontiers in Endocrinology*, *11*, 632. https://doi.org/10.3389/fendo.2020.00632

Cederholm, J., & Wibell, L. (1990). Insulin release and peripheral sensitivity at the oral glucose tolerance test. *Diabetes Research and Clinical Practice*, *10*(2), 167–175. https://doi.org/10.1016/0168-8227(90)90040-z

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J. F., & Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise*, *35*(8), 1381–1395. https://doi.org/10.1249/01.MSS.0000078924.61453.FB

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*(4), e0154075. https://doi.org/10.1371/journal.pone.0154075

Gray, S. R., Aird, T. P., Farquharson, A. J., Horgan, G. W., Fisher, E., Wilson, J., Hopkins, G. E., Anderson, B., Ahmad, S. A., Davis, S.

R., & Drew, J. E. (2018). Inter-individual responses to sprint interval training, a pilot study investigating interactions with the sirtuin system. *Applied Physiology, Nutrition, and Metabolism*, *43*(1), 84–93. https://doi.org/10.1139/apnm-2017-0224

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. *Applied Physiology, Nutrition, and Metabolism*, *41*(3), 229–234. https://doi.org/10.1139/apnm-2015-0449

Hammond, B. P., Stotz, P. J., Brennan, A. M., Lamarche, B., Day, A. G., & Ross, R. (2019). Individual variability in waist circumference and body weight in response to exercise. *Medicine & Science in Sports & Exercise*, *51*(2), 315–322. https://doi.org/10.1249/MSS.0000000000001784

Hawley, J. A., & Gibala, M. J. (2012). What's new since hippocrates? Preventing type 2 diabetes by physical exercise and diet. *Diabetologia*, *55*(3), 535–539. https://doi.org/10.1007/s00125-012-2460-1

Hopkins, W. G. (2015). Individual responses made easy. *Journal of Applied Physiology*, *118*(12), 1444–1446. https://doi.org/10.1152/japplphysiol.00098.2015

Kelley, G. A., Kelley, K. S., & Pate, R. R. (2020). Are there inter-individual differences in Fat mass and percent body Fat as a result of aerobic exercise training in overweight and obese children and adolescents? A meta-analytic perspective. *Childhood Obesity*, *16*(5), 301–306. https://doi.org/10.1089/chi.2020.0056

Kelley, G. A., Kelley, K. S., & Pate, R. R. (2021). Inter-individual differences in body mass index were not observed as a result of aerobic exercise in children and adolescents with overweight and obesity. *Pediatric Obesity*, *16*(1), e12692. https://doi.org/10.1111/ijpo.12692

Lannoy, L. de., Clarke, J., Stotz, P.J., Ross, R. (2017). Effects of intensity and amount of exercise on measures of insulin and glucose: Analysis of inter-individual variability. *PLOS ONE*, *12*(5), e0177095. https://doi.org/10.1371/journal.pone.0177095.

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. *European Journal of Applied Physiology*, *112*(7), 2767–2775. https://doi.org/10.1007/s00421-011-2254-z

Metcalfe, R. S., Tardif, N., Thompson, D., & Vollaard, N. B. J. (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. *Applied Physiology, Nutrition, and Metabolism*, *41*(11), 1117–1123. https://doi.org/10.1139/apnm-2016-0253

Metcalfe, R. S., & Vollaard, N. B. J. (2021). Heterogeneity and incidence of non-response for changes in cardiorespiratory fitness following time-efficient sprint interval exercise training. *Applied Physiology, Nutrition, and Metabolism*, *46*(7), 735–742. https://doi.org/10.1139/apnm-2020-0855

Phillips, B., Kelly, B. M., Lilja, M., Ponce-Gonzalez, 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. *Frontiers in Endocrinology*, *8*, 229. https://doi.org/10.3389/fendo.2017.00229

Sparks, L. M. (2017). Exercise training response heterogeneity: Physiological and molecular insights. *Diabetologia*, *60*(12), 2329–2336. https://doi.org/10.1007/s00125-017-4461-6

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. *Frontiers in Nutrition*, *5*, 41. https://doi.org/10.3389/fnut.2018.00041

Wallace, T. M., Levy, J. C., & Matthews, D. R. (2004). Use and abuse of HOMA modeling. *Diabetes Care*, *27*(6), 1487–1495. https://doi.org/10.2337/diacare.27.6.1487

Walsh, J. J., Bonafiglia, J. T., Goldfield, G. S., Sigal, R. J., Kenny, G. P., Doucette, S., Hadjiyannakis, S., Alberga, A. S., Prud'homme, D., & Gurd, B. J. (2019). Interindividual variability and individual responses to exercise training in adolescents with obesity. *Appl. Physiol. Nutr. Metab.* apnm-2019-0088. https://doi.org/10.1139/apnm-2019-0088

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., … Coombes, J. S. (2019). A multi-center comparison of O2peak trainability between interval training and moderate intensity continuous training. *Frontiers in Physiology*, *10*, 19. https://doi.org/10.3389/fphys.2019.00019

Williamson, P. J., Atkinson, G., & Batterham, A. M. (2017). Inter-Individual responses of maximal oxygen uptake to exercise training: A critical review. *Sports Medicine*, *47*(8), 1501–1513. https://doi.org/10.1007/s40279-017-0680-8

Williamson, P. J., Atkinson, G., & Batterham, A. M. (2018). Inter-individual differences in weight change following exercise interventions: A systematic review and meta-analysis of randomized controlled trials: Individual variance in weight loss response: A systematic review. *Obesity Reviews*, *19*(7), 960–975. https://doi.org/10.1111/obr.12682

Wilmore, J. H., Green, J. S., Stanforth, P. R., Gagnon, J., Rankinen, T., Leon, A. S., Rao, D. C., Skinner, J. S., & Bouchard, C. (2001). Relationship of changes in maximal and submaximal aerobic fitness to changes in cardiovascular disease and non[ndash]insulin-dependent diabetes mellitus risk factors with endurance training: The HERITAGE family study. *Metabolism*, *50*(11), 1255–1263. https://doi.org/10.1053/meta.2001.27214