BONAFIGLIA, J.T., SWINTON, P.A., ROSS, R., JOHANNSEN, N.M., MARTIN, C.K., CHURCH, T.S., SLENTZ, C.A., ROSS, L.M., KRAUS, W.E., WALSH, J.J., KENNY, G.P., GOLDFIELD, G.S., PRUD'HOMME, D., SIGAL, R.J., EARNEST, C.P. and GURD, B.J. 2022. Interindividual differences in trainability and moderators of cardiorespiratory fitness, waist circumference, and body mass responses: a large-scale individual participant data meta-analysis. *Sports medicine* [online], 52(12), pages 2837-2851. Available from: https://doi.org/10.1007/s40279-022-01725-9

# Interindividual differences in trainability and moderators of cardiorespiratory fitness, waist circumference, and body mass responses: a large-scale individual participant data meta-analysis.

BONAFIGLIA, J.T., SWINTON, P.A., ROSS, R., JOHANNSEN, N.M., MARTIN, C.K., CHURCH, T.S., SLENTZ, C.A., ROSS, L.M., KRAUS, W.E., WALSH, J.J., KENNY, G.P., GOLDFIELD, G.S., PRUD'HOMME, D., SIGAL, R.J., EARNEST, C.P. and GURD, B.J.

2022

**Interindividual differences in trainability and moderators of cardiorespiratory fitness, waist circumference, and body mass responses: A large-scale individual participant data meta-analysis**

Jacob T. Bonafiglia [1], Paul A. Swinton [2], Robert Ross [1], Neil M. Johannsen [3,4], Corby K. Martin [4], Timothy S. Church [4], Cris A. Slentz [5], Leanna M. Ross [5], William E. Kraus [5,6], Jeremy J. Walsh [7], Glen P. Kenny [8,9], Gary S. Goldfield [8,10,11,12], Denis Prud'homme [8,13], Ronald J. Sigal [8,9,14], Conrad P. Earnest [15†], and Brendon J. Gurd [1*]

[*] Corresponding Author; email: gurdb@queensu.ca

[1] School of Kinesiology and Health Studies, Queen's University, Kingston, ON, Canada

[2] School of Health Sciences, Robert Gordon University, Aberdeen, United Kingdom

[3] School of Kinesiology, Louisiana State University, Baton Rouge, LA, USA

[4] Pennington Biomedical Research Center, Baton Rouge, LA, USA

[5] Duke Molecular Physiology Institute, Duke University School of Medicine, Durham, NC, USA

[6] Division of Cardiology, Duke University School of Medicine, Durham, NC, USA

[7] Department of Kinesiology, McMaster University, Hamilton, ON, Canada

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

[9] Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada

[10] Healthy Active Living and Obesity Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada

[11] Department of Pediatrics, University of Ottawa, Ottawa, ON, Canada

[12] School of Psychology, University of Ottawa, Ottawa, ON, Canada

[13] Institut du Savoir Montfort, Ottawa, ON, Canada

[14] Departments of Medicine, Cardiac Sciences and Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[15] Health and Kinesiology, Texas A & M University, College Station, TX, USA

[†] Deceased. Dr. Earnest made substantial contributions to the conception, design, and data collection of the trials conducted at the Pennington Biomedical Research Center. Dr. Earnest was unable to edit or approve the final version of this manuscript.

**Abstract**

Although many studies have assumed variability reflects variance caused by exercise training, few studies have examined whether interindividual differences in trainability are present following exercise training. The present individual participant data (IPD) meta-analysis sought to: 1) investigate the presence of interindividual differences in trainability for cardiorespiratory fitness (CRF), waist circumference, and body mass; and 2) examine the influence of exercise training and potential moderators on the probability that an individual will experience clinically important differences. The IPD meta-analysis combined data from 1,879 participants from eight previously-published randomized controlled trials. We implemented a Bayesian framework to: 1) test the hypothesis of interindividual differences in trainability by comparing variability in change scores between exercise and control using Bayes factors; and 2) compare posterior predictions of control and exercise across a range of moderators (baseline BMI and exercise duration, intensity, amount, mode and adherence) to estimate the proportions of participants expected to exceed minimum clinically important differences (MCIDs) for all three outcomes. Bayes factors demonstrated a lack of evidence supporting a high degree of variance attributable to interindividual differences in trainability across all three outcomes. These findings indicate that interindividual variability in observed changes are likely due to measurement error and external behavioural factors, not interindividual differences in trainability. Additionally, we found that a larger proportion of exercise participants were expected to exceed MCIDs compared with controls for all three outcomes. Moderator analyses identified that larger proportions were associated with a range of factors consistent with standard exercise theory and were driven by mean changes. Practitioners should prescribe exercise interventions known to elicit large mean

changes to increase the probability that individuals will experience beneficial changes in CRF, waist circumference, and body mass.

**Running Title –** Exercise individual response: IPD meta-analysis

**Key Points**

- For the purposes of this meta-analysis, we define "trainability" as the change in a given variable directly attributable to an effect of exercise training free of measurement error and confounding factors.
- Larger exercise doses and other prescription factors consistent with standard exercise theory and larger mean changes were associated with larger proportions of individuals experiencing clinically meaningful changes in cardiorespiratory fitness, waist circumference, and body mass.
- Regardless of whether individuals respond differently as a result of exercise training *per se*, clinicians should prescribe exercise doses known to elicit large mean changes in order to increase the probability that individuals experience clinically meaningful improvements in cardiorespiratory fitness, waist circumference, and body mass.

## 1 - Introduction

Many exercise training studies have interpreted wide ranges of observed changes in physiological outcomes as evidence that individuals demonstrate varying degrees of trainability – the change in a given variable directly attributable to an effect of exercise training *per se* ([1–3]). However, these interpretations ignore the confounding influence of measurement error and/or variability introduced by changes in behavioural/environmental factors not related to exercise training including changes in sleep, diet, stress, etc. [4]. The confounding influences of behavioural and environmental factors are collectively referred to “within-subject variability”, and recognizing this source of variation challenges the assumption that interindividual differences in trainability exist following ostensibly the same exercise training stimulus [5,6]. Rather than assuming its existence, several studies [7–12] have estimated the presence of interindividual differences in trainability by determining whether the variability of change scores is larger in exercise compared with control groups [5]. Only some of these studies reported larger variability in exercise groups [7–12], and this inconsistency may be explained by small sample sizes (range: 26 to 181) leading to imprecise estimates, or by heterogeneity in the outcomes examined across these studies. It therefore remains unclear the extent to which variability in observed changes reflects interindividual differences in trainability.

Analyses pooling data from the same outcome across multiple studies can offer greater precision for determining the presence of interindividual differences in trainability. Recent aggregate data meta-analyses – with sample sizes ranging from 1,185 to 1,500 participants – have reported a lack of clinically-important [13,14] or no [15] evidence of interindividual differences in trainability in body mass and body composition parameters. An alternative to aggregate data meta-analyses are individual participant data (IPD) meta-analyses, which involve obtaining and analyzing raw participant data. Compared with aggregate data meta-analyses, IPD meta-analyses permit the ability to investigate potential moderators, provide more precise estimates, and enable greater flexibility in statistical modelling by unrestricting assumptions of the distribution of underlying change scores [16]. We [17] recently compiled a large dataset of 1,879 participants across eight RCTs that investigated the effects of different doses of exercise training on various health outcomes including cardiorespiratory fitness (CRF), waist circumference, and body mass. This dataset presents an opportunity to perform an IPD meta-analysis to assess the extent to which interindividual differences exist in trainability of CRF and body composition parameters.

Despite limited evidence supporting the presence of interindividual differences in trainability, there is an abundance of evidence (reviewed in: [1–3]) demonstrating individual differences in observed changes in outcomes after completing ostensibly the same exercise training intervention. For example, individual changes in relative cardiorespiratory fitness (CRF) following 24 weeks of standard aerobic training ranged from ~ -3 to +16 mL/kg/min [18], and this range of change scores (~19 mL/kg/min) substantially exceeded both a clinically-meaningful CRF change (*e.g.* 3.5 mL/kg/min [19]) and the variation that is equivalent to measurement error alone (~2.31 mL/kg/min). That is, although the relative contribution of trainability to observed changes in outcomes is unclear, it is clear participants with the largest observed change scores had a higher probability of experiencing clinically meaningful CRF improvements than participants with the lowest observed change scores. Exploring potential moderators of observed change scores may elucidate exercise prescription strategies for maximizing the probability that an individual experiences a meaningful change. Employing a Bayesian framework that enables flexible modelling and generation of subjective probabilities

[20] provides an effective method for interpreting change scores not simply on mean values in the measured units, but more applied and clinically relevant interpretations such as the expected proportions to exceed relevant thresholds. Conducting a Bayesian IPD meta-analysis with our large dataset [17] provides the scope to examine the role of potential moderators such as exercise adherence, intensity, duration, and mode on the probability that an individual will experience a meaningful change in CRF, waist circumference, or body mass.

Accordingly, the objectives of this large dataset (n = 1,879 participants) IPD meta-analysis were to: 1) investigate the presence of interindividual differences in trainability for CRF, waist circumference, and body mass, and 2) examine the influence of exercise training and potential moderators on the probability that an individual will experience benefit in these three outcomes. We also estimated the influence of exercise training and potential moderators on the distribution (*i.e.* standard deviation) of CRF, waist circumference, and body mass change scores.

## 2 - Methods

The present study is an IPD meta-analysis of CRF, waist circumference, and body mass data from eight previously published exercise intervention RCTs. Table 1 summarizes the participant characteristics, total sample sizes, and training protocols, with full study details published elsewhere [21–28]. Each study received ethics approval at their respective institutions, conformed to the Declaration of Helsinki, and obtained written informed consent from each participant prior to commencing data collection.

### *2.1 - Outcomes*

Although outcome assessment protocols varied slightly across studies (full details elsewhere: [18,22–25,27–32]), all 8 studies used similar methods to measure CRF, waist circumference, and body mass. Briefly, CRF was determined as the maximum level of oxygen consumption, measured via gas exchange using a metabolic cart, during an incremental exercise test to exhaustion and expressed in relative (mL/kg/min) units. Waist circumference was manually assessed using tape measures (expressed in centimeters) and body mass was measured using scales (expressed in kilograms). We focused on these three outcomes because they were included in all 8 studies and because they are clinically relevant due to their association with all-cause morbidity and mortality [19,33,34]. Our analyses (described in 2.2 and 2.4) estimated the proportion of individuals that would be expected to exceed minimal clinically important differences (MCID), which were +3.5mL/kg/min for CRF, -2cm for waist circumference, and -2kg for body mass as we [11,35] and others [9] have used previously. The analysis approach was selected for multiple reasons. Firstly, the proportion of individuals that exceed an MCID provides an easy-to-understand outcome that communicates the effectively the practical relevance of an intervention. Secondly, the difference in proportion of individuals that exceed the MCID between exercise and control, or due to change in a moderator provides an informative and clinically relevant perspective. Thirdly, the results of each of these large reviews have been published previously where analyses have already focussed on standard analyses such as mean change.

*2.2 - Bayesian framework*

The majority of meta-analysis (examples: [13–15]) follow a frequentist framework whereby parameters (*e.g.* means and standard deviations [SDs]) are objectively estimated from the data and uncertainty is expressed with confidence intervals. A limitation with confidence intervals is their inability to provide distributional information, such that there is no direct sense for whether a parameter estimate in the middle of the interval is more probable of representing the true value than any other value within the interval [20]. In other words, a 90% confidence interval centered around a mean CRF change of 3 mL/kg/min and ranging from 1 to 5 mL/kg/min should be interpreted as: 90% of similarly sized intervals (*i.e.* ranging 4 mL/kg/min) obtained from repeatedly completing the trial will contain the true mean change [20]. However, researchers often misinterpret confidence intervals [36] as (in keeping with the previous example): there being a 90% chance that the true change in CRF is between 1 and 5 mL/kg/min. Although the latter interpretation is perhaps more intuitive and desirable when trying to estimate a given parameter (*e.g.* true mean change in CRF), this interpretation cannot be made within a traditional-frequentist framework [20].

Instead of implementing a frequentist approach, we implemented a Bayesian framework for our IPD meta-analysis. Rather than estimating parameters from the data alone, Bayesian frameworks combine prior beliefs and the data to estimate the most plausible parameter values (*e.g.* mean change in CRF). Bayesian frameworks are therefore considered subjective because researchers can incorporate their *a priori* expectations when estimating parameters. For example, a researcher could use information from several large-scale, rigorous meta-analyses to develop an expected mean change in CRF, and then combine this information with their actual data to derive the most plausible estimate for the true mean change in CRF. In Bayesian analysis, prior beliefs refer to the probability of obtaining parameter values (*e.g.* mean change in CRF) given a specific data generating model (*e.g.* normal distribution), and are written as:

$$p(\Theta|M)$$

where $p$ is the probability, $\Theta$ are the parameters of the model (*e.g.* mean change in CRF and standard deviation), the vertical dash means given, and $M$ is the model (*e.g.* normal distribution). The prior is combined with the likelihood, which refers to the probability of obtaining the data (*e.g.* dataset of raw CRF change scores) given specific parameter values and the specified model. The likelihood is written as:

$$p(y|\Theta, M)$$

The prior and likelihood are then combined and scaled to obtain a posterior distribution reflecting updates of beliefs in the light of the data and written as:

$$p(\Theta|y, M)$$

Intervals known as credible intervals (CrIs) can also be constructed from the posterior distributions and quantify the probability of containing the actual parameter value (*e.g.* a 90% chance of containing the true mean change in CRF). It is important to emphasize that credible intervals represent *subjective* probabilities because they are built using prior beliefs. Nevertheless, if prior beliefs are well justified (*e.g.* established using relevant data), then credible intervals permit more intuitive, and arguably more useful, interpretations compared with confidence intervals [20]. Finally, different moderator values can be entered into models (e.g. exercise intervention of 4, 6 and 8 months) to simulate new data $\tilde{y}$ and estimate proportions of individuals expected to exceed thresholds such as the MCID.

We conducted our IPD meta-analysis by fitting Bayesian hierarchical distributional regression models which modeled the mean and variance parameters. All models comprised random intercepts to account for systematic differences across studies, and models with group (exercise *vs.* control) and moderators (defined below) included these variables as fixed effects.

The subsequent methods sections provide specific details for how we used these Bayesian models to investigate interindividual differences in trainability and proportions of participants exceeding MCIDs.

*2.3 - IPD Meta-Analysis: Interindividual differences in trainability*

We fit initial base models of our IPD meta-analysis that included the mean and variance parameters across three different types of distributions: normal, skew normal, and t-distributions. The most appropriate distribution type for each outcome was determined using the Watanabe-Akaike information criterion, and these identified distribution types were then used in all subsequent analyses for each outcome.

To investigate the presence of interindividual differences in trainability, we first conducted analyses to obtain Bayes factors. Bayes factors are denoted as:

$$\left(\frac{p(y|M_1)}{p(y|M_2)}\right)$$

because they are obtained by estimating the probability ($p$) of obtaining the data ($y$) given two different models: $M_1$ represents a model that included group as a fixed effect of the variance parameters (*i.e.* exercise *vs.* control), whereas the $M_2$ model did not contain a group factor for variance parameters (*i.e.* all data combined as coming from one large group). That is, the $M_1$ model allowed us to estimate the probability that the variance in exercise change scores exceeded the variance in control change scores – an observation indicating the presence of interindividual differences in trainability [5]. Conversely, the $M_2$ model estimated the probability of the null hypothesis (*i.e.* variance in exercise not greater than variance in control). A Bayes factor greater than 1.0 would indicate that $M_1$ was a better fit, which would then indicate the presence of interindividual differences in trainability because the probability of the variance in exercise exceeding control was higher than the probability of the null [37]. Bayes factors less than 1.0 would therefore indicate a lack of interindividual differences in trainability [37]. The strength of evidence in favour of either model ($M_1$ or $M_2$) was evaluated according to a previously defined scale [37]. As described above, Bayesian frameworks require incorporating prior beliefs. Given limited pre-existing data to justify appropriate priors, we created “local” priors using our dataset. Specifically, we developed priors from randomly created “training sets” that consisted of 1/3 of the total dataset, meaning that Bayes factors were calculated on the remaining 2/3 of the dataset. Due to stability issues with calculating Bayes factors [38], we repeated these steps four times (*i.e.* creating five different priors each containing 1/3 of the data) and calculated an average Bayes factor for each outcome. As a final check, we calculated Bayes factors with weakly informative priors, which returned values close to the average Bayes factors calculated with local priors.

*2.4 - IPD Meta-Analysis: Posterior predictions for proportions and distributions of change scores*

To investigate the proportion of individuals in exercise and control exceeding the MCID, we used the posterior samples $p(\theta|y, M)$ from the best fit distributional base model to generate posterior predictions $p(\tilde{y}|\theta, M)$ (n=1000) and calculated the proportion of samples exceeding the MCID. To compare variances in both exercise and control, the $M_1$ model was used. Given the heterogeneous nature of the data with regards to participant (sex, age and diabetes status) and exercise (aerobic, resistance or combined) characteristics, individual subgroup analyses were conducted and are presented in Supplemental Tables 1-3. Moderator analyses were then investigated through a similar process, first obtaining posterior samples, and then generating posterior predictions. Moderator fixed effects were included for the mean and variance parameters. As mentioned above, an additional advantage of Bayesian analysis is the flexibility in fitting models when pooling data in IPD meta-analyses [39]. For instance, although only one trial included measures at four months [40], we were able to include this time point in our duration moderator analysis through simulation and subsequently estimate proportions exceeding MCID and standard deviations at four months.

We evaluated six moderators: 1) intervention duration (4, 6 or 8 months); 2) exercise adherence (number of calories expended during aerobic exercise training relative to the amount prescribed; categorized as ≥ or < 70% for "high" or "low" adherence, respectively); 3) exercise mode (aerobic, resistance, or combined); 4) exercise intensity (aerobic exercise only – including binary low/high with cut-offs comprising 60% of maximum CRF, heart rate, or $VO_2$ reserve); 5) exercise amount (aerobic exercise only – low: less than 500kcal per session; mid: between 500-1000kcal per session; high: greater that 1000kcal per session); and 6) baseline BMI (trinary as mean or beyond ± 1 SD). We only evaluated exercise adherence for groups that followed aerobic or combined aerobic and resistance training as exercise expended calories were not used to characterize adherence to resistance training. Because Bayesian analyses estimates *subjective* probabilities, we *subjectively* interpreted differences in proportions across moderators rather than identifying influential moderators with *objective* cut-offs. For example, because confidence intervals do not provide any distributional information (*e.g.* unclear whether most likely proportion is at the center or outskirts of the confidence interval), a frequentist approach using confidence intervals may limit us to identifying moderators as being influential only if confidence intervals do not overlap (*e.g.* high intensity confidence interval lay fully above low intensity confidence interval). However, this conservative approach is unwarranted with

Bayesian analyses because each proportion represents the most probable estimate (*i.e.* the center of the credible interval is indeed the most likely proportion). Therefore, our *subjective* interpretations looked for patterns in proportions across levels (*e.g.* proportions increasing from 4 to 6 to 8 months) and noted whether results were consistent with standard exercise theory (*e.g.* higher exercise dose resulting in larger proportions [17]). It is important to note that the proportion of individuals exceeding the MCID was based on a modelling approach of the change distributions and not dichotomisation of individual results (*e.g.* direct calculation of proportion from the sample) which substantially reduces the amount of information available and fails to account for uncertainty in individual measurements. We therefore did not use the terms "responder" or "non-responder" when interpreting our results.

Weakly informative Student-t prior and half-t priors with 3 degrees of freedom and scale parameter equal to 2.5 were used for intercept and variance parameters for the hierarchical distributional models [41]. All analyses were performed using the R wrapper package brms interfaced with Stan to perform sampling [42] and the R package bridgesampling to calculate Bayes factors. Convergence of parameter estimates was obtained for all models with Gelman-Rubin R-hat values below 1.1 [43].

## 3 - Results

### *3.1 – Cardiorespiratory fitness*

The best model fit for CRF change scores (Figure 1) was obtained using a t-distribution (expected log predictive density [elpd] difference: t-distribution vs. normal skew = 3.0 times standard error; t-distribution vs. normal = 4.0 times standard error). The base IPD model estimated a mean change of 2.2 ml/kg/min [90%CrI: 1.5 to 3.0] for exercise and -0.29 ml/kg/min

[90%CrI: -1.0 to 0.6] for control. The base IPD model also estimated a standard deviation of change scores of 3.4 [90%CrI: 2.9 to 3.9] and 3.5 [90%CrI: 2.9 to 4.2] for exercise and control. The average Bayes factor was less than 1.0 and identified moderate evidence (average Bayes factor = 0.11, range: 0.01 to 0.15) supporting the $M_2$ l model, thereby refuting the presence of interindividual differences in trainability. Substantive overlap of standard deviation of change scores across all subgroups (Supplemental Table 1) provides additional support refuting the presence of interindividual differences in trainability. Table 2 presents the estimated proportions of participants exceeding the MCID of 3.5 mL/kg/min and estimated standard deviations of change scores with 90% CrI denoting the subjective probabilities. Exercise training had a higher estimated proportion of participants (estimated proportion, 30% [90% CrI:21 to 41%]) exceeding the MCID of 3.5 mL/min/kg compared with control (11% [90% CrI:5 to 19%]). Several moderators appeared to increase estimated proportions of participants exceeding the CRF MCID in the exercise group consistent with standard exercise theory (Table 2): 1) longer exercise durations, 2) higher exercise adherence, 3) higher exercise intensity, 4) combined aerobic and resistance, which was prescribed at a higher exercise dose than aerobic or resistance training alone [23,24,28], and 5) higher exercise amount. Interestingly, larger mean changes likely explained larger proportions because proportions increased within a given group (exercise or control) and within some moderators (duration, baseline BMI and exercise mode) despite larger estimates of standard deviation of change scores (Table 2).

*3.2 - Body Composition Parameters*

The best model fit for both waist circumference (Figure 2) and body mass (Figure 3) was obtained using a t-distribution (elpd difference: t-distribution vs. normal skew = 2.3 to 2.6 times standard error; t-distribution vs. normal = 3.9 to 5.0 times standard error). The base IPD model estimated a mean waist circumference change of -2.5 cm [90%CrI: -3.2 to -1.9] for exercise and -0.04 cm [90%CrI: -0.8 to 0.6] for control, and a mean body mass change of -1.4 kg [90%CrI: -2.2 to -0.8] for exercise and -0.02 kg [90%CrI: -0.8 to 0.6] for control. The base IPD model also estimated a standard deviation of waist circumference change scores of 4.9 cm [90%CrI: 4.2 to 5.6] for exercise and 5.7 [90%CrI: 4.6 to 7.9] for control, and a standard deviation of body mass change scores of 4.1 kg [90%CrI: 3.5-5.0] for exercise and 4.6 [90%CrI: 3.7 to 6.4] for control. The average Bayes factor was less than 1.0 for both outcomes and identified "anecdotal" evidence supporting the $M_2$ model (waist circumference: average Bayes factor = 0.47, range: 0.41 to 0.56; body mass: average Bayes factor = 0.39, range: 0.22 to 0.68). Similar to changes in CRF, substantive overlap of standard deviation of change scores across all subgroups (Supplemental Tables 2 and 3) provides additional support refuting the presence of interindividual differences in trainability. Tables 3 and 4 present the estimated proportions of participants exceeding MCIDs of -2cm and -2kg as well as estimated mean and standard deviations of change with 90% CrIs denoting subjective probabilities for waist circumference and body mass, respectively. Both outcomes had higher estimated proportion of participants exceeded MCIDs in exercise (waist circumference: 54% [90% CrI: 48 to 61%]; body mass: 42% [90% CrI: 34 to 50%]) compared with control groups (waist circumference: 30% [90% CrI: 23 to 38%]; body mass: 26% [90% CrI: 18 to 35%).

Several moderators appeared to increase estimated proportions of participants exceeding the waist circumference MCID in the exercise group consistent with standard exercise theory (Table 3): 1) higher exercise adherence, 2) higher exercise intensity, 3) combined aerobic and resistance compared with aerobic or resistance training alone, and 4) higher exercise amount. However, longer exercise durations beyond 4 months did not appear to increase proportions exceeding the MCID for waist circumference. Several moderators also appeared to increase estimated proportions of participants exceeding the body mass MCID in the exercise group consistent with standard exercise theory (Table 4): 1) higher exercise adherence, 2) higher exercise intensity, 3) combined aerobic and resistance training compared with aerobic or resistance training alone, and 4) higher exercise amounts (low *vs.* high). Interestingly, longer exercise durations appeared to decrease the proportions of participants exceeding the body mass MCID. Additionally, our results indicated an inconsistent pattern with baseline BMI as both lower (-1SD) and higher (+1SD) levels were associated with larger proportions than mean levels (± 1SD). Similar to CRF, many of the most probable estimates of standard deviations of change scores were larger as proportions increased within a given group (exercise or control) and within some moderators for waist circumference (exercise duration, mode, and amount; Table 3) and body mass (baseline BMI, adherence, and exercise amount; Table 4).

## 4 - Discussion

This was the first IPD meta-analysis to investigate the presence of interindividual differences in trainability and estimate proportions of participants expected to experience meaningful benefit in CRF, waist circumference, and body mass. Our results revealed four key findings: 1) large between-subject variability in observed change scores in both exercise and control groups; 2) consistent evidence of a lack of interindividual differences in trainability; 3) a higher proportion of participants exceeding MCIDs following exercise training compared with control for all three outcomes; and 4) several moderators consistent with standard exercise theory including higher exercise adherence, intensity, amount, and combined aerobic and resistance training were associated with higher proportions of participants exceeding MCIDs for all three outcomes. Collectively, our results indicate that over periods of 4 to 8 months, individuals can experience relatively large changes in observed CRF, waist circumference, and body mass. The variation in these changes is consistent between exercise and control groups, negating the notion that interindividual differences in trainability explains why individuals appear to differentially benefit following exercise training. However, compared with control, exercise results in larger mean changes causing systematic shifts in change score distributions centered around the mean change. This shift has a substantive effect on the proportion of individuals expected to experience clinically meaningful benefits in CRF, waist circumference, and body mass. Accordingly, exercise prescriptions that elicit larger mean changes – such as increasing exercise amount [18,30,32] – can also shift the overall change distribution and thus further increase the likelihood of clinically meaningful benefits.

Our findings add to the growing body of work questioning the assumption that variability in observed responses to exercise training reflects interindividual differences in trainability [12,14,15,44,45]. Among the meta-analyses questioning this assumption [13–15], we believe the present IPD meta-analysis provides the most powerful evidence for several reasons: 1) we included a very large sample size (n = 1,879) gathered from 8 methodologically-robust RCTs [21–28], 2) we obtained consistent findings across multiple outcomes; 3) we included flexible and detailed analysis frameworks that assessed the distribution of change scores (e.g. a t-distribution with wider tails than Gaussian such that more than 5% of participants lay beyond 2 standard deviations), and 4) we demonstrated consistent variances between exercise and control even when including moderators such as duration and baseline BMI. In addition, the present IPD meta-analysis extends previous meta-analyses [13–15] by contextualizing the practical significance (*i.e.* proportions exceeding MCIDs) of shifted but similar spread change score distributions between exercise and control, and across different levels of common exercise moderators (Tables 2 – 4). However, this assumption may be inappropriate as the inability to blind group assignment in exercise RCTs may lead to some participants initiating behavioural changes based on their preference toward their assigned group [46], which in turn can lead to unequal within-subject variability between groups [6]. There are additional differences between groups within an RCT that can lead to differences in variance between groups such as (non)-compliance or pre-randomization susceptibility to adaptation [47]. A within-subjects design in which participants are repeatedly exposed to both control and exercise interventions avoids this assumption by directly quantifying error and within-subject variability [48,49]. However, these study designs are costly, labour intensive, and may introduce additional confounding variables (*e.g.* carryover effects) [50]. Therefore, at present, the exercise training literature has yet to conclusively demonstrate the presence of interindividual differences in trainability.

Although we did not observe evidence of variability caused by exercise training *per se*, we did obtain large most probable estimates of standard deviation of change scores (Tables 2 – 4). For instance, the standard deviation of change scores for both exercise and control groups exceeded the typical errors of measurement reported in the literature (~1-2 mL/kg/min for CRF [18,51]; ~0.5 cm for waist circumference [52,53], and ~0.5 kg for body mass [52]). Our findings therefore indicate that individuals experienced real physiological differences in changes in CRF, waist circumference, and body mass, and that behavioural factors (*e.g.* sleep, stress, external physical activity, etc. [4]) may underlie this variance rather than exercise *per se*. Future work is needed to investigate the contribution of various behavioural factors on observed changes following standardized and controlled exercise interventions.

Regardless of whether a group of individuals respond differently to exercise training, practitioners in clinical and applied settings remain faced with the challenge of prescribing exercise at the individual level. Our analyses first found that a higher proportion of exercise participants were expected to exceed MCIDs for CRF, waist circumference, and body mass compared with controls, which is consistent with the well-established effect of exercise training on important health outcomes [54]. Additionally, several moderators consistent with standard exercise theory – higher exercise amounts, intensities, adherence, and combined aerobic and resistance training – resulted in higher proportions for all three outcomes. Because standard deviation of change scores did not shrink with increasing proportions (Tables 2 - 4), larger mean changes likely explained why certain moderators (*e.g.* higher exercise amounts) increased proportions of participants exceeding MCIDs. Thus, although we only explored six potential moderators, these findings suggest that mean changes would also explain why other moderators impact response proportions; however, future work is needed to confirm this speculation. We recently demonstrated that larger mean changes, not reduced interindividual variability, explain why higher doses of exercise training increase CRF response rates [17]. The present Bayesian analysis supports our recent finding [17], and suggests that practitioners should prescribe exercise doses known to elicit large mean changes in order to increase the probability that an individual experiences a meaningful change in CRF, waist circumference, and body mass. Whilst substantive imbalances in exercise and control sample sizes were obtained across all analyses, these imbalances are unlikely to have influenced the findings. Lower sample sizes in control groups resulted in wider credible intervals for estimates of change score standard deviations, however, overlap in central estimates were considerable across all analyses leading to very consistent findings regardless of the outcome variable or moderator investigated.

*4.1 - Limitations*

There are several limitations with the present analysis. First, our Bayes factor analysis supports the notion that variability in observed changes is confounded by the totality of the effects of measurement error and variation in behavioural/environmental factors. Our study design, and the designs of the included trials, did not allow us to determine the extent to which certain individual behavioural/environmental factors contributed to within-subject variability. The evidence that subtle changes in sleep quality, stress levels, or other behavioural/environmental factors impact training adaptations is indirect at best [4], warranting the need for future designed to test the effects of individual behavioural/environmental on observed variability. Second, we unfortunately do not have measures of measurement error, such as coefficients of variation, for CRF, WC, or body mass for each trial and it is possible that measurement errors varied across trial sites. Given that many previous studies have similarly reported a lack of interindividual differences to exercise training [10–13,55], we do not believe potential differences in measurement error across trial sites would have a major impact on our Bayes factor results. Nevertheless, when possible, future studies should consider incorporating site-specific measurement error into statistical models for pooled analyses. Third, although our subgroup analyses revealed a consistent lack of interindividual differences in trainability across various participant characteristics, all included trails recruited overweight, obese, or diabetic participants suggesting that our findings are not generalizable to other populations such as lean and healthy adults. In our recent systematic review [56] we did not identify any study statistically investigating the presence of interindividual differences in trainability in lean, healthy adults, thus highlighting another area for future work. Fourth, it is important to acknowledge that comparing our results in Tables 2-4 are likely outcome-dependent as proportions are determined by mean changes [17] and outcome-specific MCIDs [57]. These results should therefore be interpreted independently for each outcome and should not be used to compare proportions across CRF, WC, and body mass. Finally, it is important to recognize that the use of MCIDs in the present manuscript represents an effect size justified on associations with clinical outcomes [9,11,35]. Previous discussions have highlighted the limitations of MCIDs such as the inability to delineate regression to the mean from true responses to an intervention [58,59]. The use of MCID in the present analysis was meant to provide an easy-to-understand comparison of proportions between exercise and control groups, and it is important to consider our results in the context of limitations to MCIDs.

## 5 - Conclusion

Despite the widespread assumption that individuals respond differently to exercise, the current IPD meta-analysis provided evidence in favour of no interindividual differences in trainability for CRF, waist circumference, and body mass. Although exercise training *per se* may not explain why individuals differentially benefit from completing ostensibly the same dose of exercise training, completing exercise training will increase the probability that an individual will experience a meaningful change in CRF, waist circumference, and body mass. Moreover, individuals can experience very large changes in these three outcomes following 4 to 9 months of exercise training with large interindividual variability in observed change scores. It is therefore expected that behavioural factors (*e.g.* sleep, nutrition, stress, etc.) can influence whether an individual experiences clinically meaningful improvements, and researchers should seek to better understand which external factors are most influential for observed changes in CRF, waist circumference, or body mass. At present, our results suggest that practitioners should prescribe exercise training doses known to elicit large mean changes in order to increase the probability that an individual will experience meaningful benefits.

## DECLARATIONS

### Contributorship:

All authors, unless otherwise noted (see note regarding Dr. Earnest): 1) made substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data; 2) drafted the work or revised it critically for important intellectual content; 3) approved the version to be published; and 4) agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

### Availability of data and material:

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

### Ethical approval information:

Each study received ethics approval at their respective institutions, conformed to the Declaration of Helsinki, and obtained written informed consent from each participant prior to commencing data collection.

### Acknowledgements:

No additional acknowledgements.

### Conflicts of Interest:

Jacob T. Bonafiglia, Paul A. Swinton, Robert Ross, Neil M. Johannsen, Corby K. Martin, Timothy S. Church, Cris A. Slentz, Leanna M. Ross, William E. Kraus, Jeremy J. Walsh, Glen P. Kenny, Gary S. Goldfield, Denis Prud'homme, Ronald J. Sigal, Conrad P. Earnest and Brendon J. Gurd declare that they have no conflicts of interest relevant to the content of this review.

### Funding, grant and award info:

Jacob T. Bonafiglia and Brendon J. Gurd were supported by the Natural Sciences and Engineering Research Council of Canada.

## References

1. Ross R, Goodpaster BH, Koch LG, Sarzynski MA, Kohrt WM, Johannsen NM, et al. Precision exercise medicine: understanding exercise response variability. Br J Sports Med [Internet]. 2019;53:1141–53. Available from: http://bjsm.bmj.com/

2. Sparks LM. Exercise training response heterogeneity: physiological and molecular insights. Diabetologia [Internet]. Diabetologia; 2017;60:2329–36. Available from: http://link.springer.com/10.1007/s00125-017-4461-6

3. Pickering C, Kiely J. Do Non-Responders to Exercise Exist—and If So, What Should We Do About Them? Sport Med [Internet]. Springer International Publishing; 2019;49:1–7. Available from: http://link.springer.com/10.1007/s40279-018-01041-1

4. Mann TN, Lamberts RP, Lambert MI. High responders and low responders: Factors associated with individual variation in response to standardized training. Sport Med. 2014;44:1113–24.

5. Atkinson G, Batterham AM. True and false interindividual differences in the physiological response to an intervention. Exp Physiol [Internet]. 2015;100:577–88. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25823596

6. Bonafiglia JT, Brennan AM, Ross R, Gurd BJ. 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 [Internet]. 2019;7:e14163. Available from: https://onlinelibrary.wiley.com/doi/abs/10.14814/phy2.14163

7. Steele J, Fisher J, Bruce-Low S, Smith D, Osborne N, Newell D. Variability in Strength, Pain,

and Disability Changes in Response to an Isolated Lumbar Extension Resistance Training Intervention in Participants with Chronic Low Back Pain. Healthcare [Internet]. 2017;5:75. Available from: http://www.mdpi.com/2227-9032/5/4/75

8. Yu F, Salisbury D, Mathiason MA. Interindividual differences in the responses to aerobic exercise in Alzheimer's disease: Findings from the FIT-AD trial. J Sport Heal Sci [Internet]. Elsevier B.V.; 2020; Available from: https://doi.org/10.1016/j.jshs.2020.05.007

9. Hecksteden A, Pitsch W, Rosenberger F, Meyer T. Repeated testing for the assessment of individual response to exercise training. J Appl Physiol [Internet]. 2018;124:1567–79. Available from: http://www.physiology.org/doi/10.1152/japplphysiol.00896.2017

10. Walsh JJ, Bonafiglia JT, Goldfield GS, Sigal RJ, Kenny GP, Doucette S, et al. Interindividual variability and individual responses to exercise training in adolescents with obesity. Appl Physiol Nutr Metab [Internet]. 2020;45:45–54. Available from: http://www.nrcresearchpress.com/doi/10.1139/apnm-2019-0088

11. Hammond BP, Stotz PJ, Brennan AM, Lamarche B, Day AG, Ross R. Individual Variability in Waist Circumference and Body Weight in Response to Exercise. Med Sci Sports Exerc. 2019;51:315–22.

12. Bonafiglia JT, Islam H, Preobrazenski N, Ma A, Deschenes M, Erlich AT, et al. Examining interindividual differences in select muscle and whole-body adaptations to continuous endurance training. Exp Physiol [Internet]. 2021;EP089421. Available from: https://onlinelibrary.wiley.com/doi/10.1113/EP089421

13. Williamson PJ, Atkinson G, Batterham AM. Inter-individual differences in weight change following exercise interventions : A systematic review and meta- analysis of randomized

controlled trials. Obes Rev. 2018;1–16.

14. Kelley GA, Kelley KS, Pate RR. 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. Child Obes [Internet]. 2020;X:chi.2020.0056. Available from: https://www.liebertpub.com/doi/10.1089/chi.2020.0056

15. Kelley GA, Kelley KS, Pate RR. Inter-individual differences in body mass index were not observed as a result of aerobic exercise in children and adolescents with overweight and obesity. Pediatr Obes [Internet]. 2021;16:1–9. Available from: https://onlinelibrary.wiley.com/doi/10.1111/ijpo.12692

16. Cooper H, Patall EA. The relative benefits of meta-analysis conducted with individual participant data versus aggregated data. Psychol Methods [Internet]. 2009;14:165–76. Available from: http://doi.apa.org/getdoi.cfm?doi=10.1037/a0015565

17. Bonafiglia JT, Preobrazenski N, Islam H, Walsh JJ, Ross R, Johannsen NM, et al. Exploring Differences in Cardiorespiratory Fitness Response Rates Across Varying Doses of Exercise Training: A Retrospective Analysis of Eight Randomized Controlled Trials. Sport Med [Internet]. Springer International Publishing; 2021; Available from: https://doi.org/10.1007/s40279-021-01442-9

18. Ross R, de Lannoy L, Stotz PJ. Separate effects of intensity and amount of exercise on interindividual cardiorespiratory fitness response. Mayo Clin Proc [Internet]. Elsevier Inc; 2015;90:1506–14. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0025619615006400

19. Ross R, Blair SN, Arena R, Church TS, Després J-P, Franklin BA, et al. 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 [Internet]. 2016;134:e653–99. Available from: https://www.ahajournals.org/doi/10.1161/CIR.0000000000000461

20. van de Schoot R, Kaplan D, Denissen J, Asendorpf JB, Neyer FJ, van Aken MAG. A Gentle Introduction to Bayesian Analysis: Applications to Developmental Research. Child Dev [Internet]. 2014;85:842–60. Available from: http://doi.wiley.com/10.1111/cdev.12169

21. Morss GM, Jordan AN, Skinner JS, Dunn AL, Church TS, Earnest CP, et al. Dose-Response to Exercise in Women Aged 45-75 yr (DREW): Design and Rationale. Med Sci Sport Exerc [Internet]. 2004;36:336–44. Available from: https://insights.ovid.com/crossref?an=00005768-200402000-00025

22. Myers CA, Johnson WD, Earnest CP, Rood JC, Tudor-Locke C, Johannsen NM, et al. Examination of mechanisms (E-MECHANIC) of exercise-induced weight compensation: study protocol for a randomized controlled trial. Trials [Internet]. 2014;15:212. Available from: https://trialsjournal.biomedcentral.com/articles/10.1186/1745-6215-15-212

23. Alberga AS, Goldfield GS, Kenny GP, Hadjiyannakis S, Phillips P, Prud'homme D, et al. Healthy eating, aerobic and resistance training in youth (HEARTY): Study rationale, design and methods. Contemp Clin Trials [Internet]. 2012;33:839–47. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1551714412000985

24. Church TS, Blair SN, Cocreham S, Johannsen N, Johnson W, Kramer K, et al. Effects of Aerobic and Resistance Training on Hemoglobin A 1c Levels in Patients With Type 2 Diabetes. Jama. 2010;304:2253.

25. Ross R, Hudson R, Day AG, Lam M. Dose-response effects of exercise on abdominal obesity and risk factors for cardiovascular disease in adults: Study rationale, design and methods. Contemp Clin Trials [Internet]. Elsevier B.V.; 2013;34:155–60. Available from: http://dx.doi.org/10.1016/j.cct.2012.10.010

26. Kraus WE, Torgan CE, Duscha BD, Norris J, Brown SA, Cobb FR, et al. Studies of a targeted risk reduction intervention through defined exercise (STRRIDE). Med Sci Sports Exerc [Internet]. 2001;33:1774–84. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed5&NEWS=N&AN=2001365453

27. Slentz CA, Bateman LA, Willis LH, Granville EO, Piner LW, Samsa GP, et al. Effects of exercise training alone vs a combined exercise and nutritional lifestyle intervention on glucose homeostasis in prediabetic individuals: a randomised controlled trial. Diabetologia [Internet]. Diabetologia; 2016;59:2088–98. Available from: http://dx.doi.org/10.1007/s00125-016-4051-z

28. Bateman LA, Slentz CA, Willis LH, Shields AT, Piner LW, Bales CW, et al. Comparison of Aerobic Versus Resistance Exercise Training Effects on Metabolic Syndrome (from the Studies of a Targeted Risk Reduction Intervention Through Defined Exercise - STRRIDE-AT/RT). Am J Cardiol [Internet]. Elsevier Inc.; 2011;108:838–44. Available from: http://dx.doi.org/10.1016/j.amjcard.2011.04.037

29. Duscha BD, Slentz CA, Johnson JL, Houmard JA, Bensimhon DR, Knetzger KJ, et al. Effects of Exercise Training Amount and Intensity on Peak Oxygen Consumption in Middle-Age Men and Women at Risk for Cardiovascular Disease. Chest [Internet]. The American College of Chest Physicians; 2005;128:2788–93. Available from:

http://dx.doi.org/10.1378/chest.128.4.2788

30. Church TS, Earnest CP, Skinner JS, Blair SN. Effects of Different Doses of Physical Activity on Cardiorespiratory Fitness Among Sedentary , Overweight or Obese Postmenopausal. JAMA. 2007;297:2081–91.

31. Martin CK, Johnson WD, Myers CA, Apolzan JW, Earnest CP, Thomas DM, et al. Effect of different doses of supervised exercise on food intake, metabolism, and non-exercise physical activity: The E-MECHANIC randomized controlled trial. Am J Clin Nutr [Internet]. Oxford University Press; 2019;110:583–92. Available from: https://academic.oup.com/ajcn/article/110/3/583/5512180

32. Slentz CA, Duscha BD, Johnson JL, Ketchum K, Aiken LB, Samsa GP, et al. Effects of the Amount of Exercise on Body Weight, Body Composition, and Measures of Central Obesity. Arch Intern Med [Internet]. 2004;164:31. Available from: http://archinte.jamanetwork.com/article.aspx?doi=10.1001/archinte.164.1.31

33. Ross R, Neeland IJ, Yamashita S, Shai I, Seidell J, Magni P, et al. Waist circumference as a vital sign in clinical practice: a Consensus Statement from the IAS and ICCR Working Group on Visceral Obesity. Nat Rev Endocrinol [Internet]. Springer US; 2020;16:177–89. Available from: http://dx.doi.org/10.1038/s41574-019-0310-7

34. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, et al. 2013 AHA/ACC/TOS Guideline for the Management of Overweight and Obesity in Adults. J Am Coll Cardiol [Internet]. 2014;63:2985–3023. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0735109713060300

35. Bonafiglia JT, Ross R, Gurd BJ. The application of repeated testing and monoexponential

regressions to classify individual cardiorespiratory fitness responses to exercise training. Eur J Appl Physiol [Internet]. Springer Berlin Heidelberg; 2019;119:889–900. Available from: http://dx.doi.org/10.1007/s00421-019-04078-w

36. Hoekstra R, Morey RD, Rouder JN, Wagenmakers E-J. Robust misinterpretation of confidence intervals. Psychon Bull Rev [Internet]. 2014;21:1157–64. Available from: http://link.springer.com/10.3758/s13423-013-0572-3

37. Jeffreys H. The Theory of Probability. OUP Oxford; 1939.

38. Gill J. Bayesian Methods for the Social and Behavioral Sciences. Volume 20. Boca Raton, Florida: CRC Press.; 2014.

39. Schad DJ, Nicenboim B, Bürkner P-C, Betancourt M, Vasishth S. Workflow techniques for the robust use of Bayes Factors. arXiv. 2021;

40. Ross R, Hudson R, Stotz PJ, Lam M. Effects of Exercise Amount and Intensity on Abdominal Obesity and Glucose Tolerance in Obese Adults. Ann Intern Med [Internet]. 2015;162:325. Available from: http://annals.org/article.aspx?doi=10.7326/M14-1189

41. Gelman A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian Anal [Internet]. 2006;1:515–34. Available from: https://projecteuclid.org/journals/bayesian-analysis/volume-1/issue-3/Prior-distributions-for-variance-parameters-in-hierarchical-models-comment-on/10.1214/06-BA117A.full

42. Bürkner P-C. brms : An R Package for Bayesian Multilevel Models Using Stan. J Stat Softw [Internet]. 2017;80. Available from: http://www.jstatsoft.org/v80/i01/

43. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DS. Bayesian Data Analysis.

Third. CRC Press; 2013.

44. Islam H, Bonafiglia JT, Del Giudice M, Pathmarajan R, Simpson CA, Quadrilatero J, et al. Repeatability of training-induced skeletal muscle adaptations in active young males. J Sci Med Sport [Internet]. Sports Medicine Australia; 2021;24:494–8. Available from: https://doi.org/10.1016/j.jsams.2020.10.016

45. Del Giudice M, Bonafiglia JTJT, Islam H, Preobrazenski N, Amato A, Gurd BJBJ. Investigating the reproducibility of maximal oxygen uptake responses to high-intensity interval training. J Sci Med Sport [Internet]. Sports Medicine Australia; 2020;23:94–9. Available from: https://doi.org/10.1016/j.jsams.2019.09.007

46. Halpern SD. Evaluating preference effects in partially unblinded, randomized clinical trials. J Clin Epidemiol. 2003;56:109–15.

47. Mills H, Higgins JP, Morris R, Kessler D, Heron J, Wiles N, et al. Detecting heterogeneity of intervention effects using analysis and meta-analysis of differences in variance between arms of a trial. MedRxiv. 2020;

48. Hecksteden A, Kraushaar J, Scharhag-Rosenberger F, Theisen D, Senn S, Meyer T. Individual response to exercise training - a statistical perspective. J Appl Physiol [Internet]. 2015;118:1450–9. Available from: https://www.physiology.org/doi/10.1152/japplphysiol.00714.2014

49. Senn S. Individual Therapy: New Dawn or False Dawn? Drug Inf J [Internet]. 2001;35:1479–94. Available from: http://dij.sagepub.com/lookup/doi/10.1177/009286150103500443%5Cnpapers2://publication/doi/10.1177/009286150103500443

50. Williamson PJ, Atkinson G, Batterham AM. Inter-individual responses of maximal oxygen uptake to exercise training: A critical review. Sport Med [Internet]. Springer International Publishing; 2017;47:1501–13. Available from: http://link.springer.com/10.1007/s40279-017-0680-8

51. Gurd BJ, Giles MD, Bonafiglia JT, Raleigh JP, Boyd JC, Ma JK, et al. Incidence of nonresponse and individual patterns of response following sprint interval training. Appl Physiol Nutr Metab [Internet]. 2016;41:229–34. Available from: http://www.nrcresearchpress.com/doi/10.1139/apnm-2015-0449

52. Barbalho M de SM, Gentil P, Izquierdo M, Fisher J, Steele J, Raiol R de A. There are no no-responders to low or high resistance training volumes among older women. Exp Gerontol [Internet]. Elsevier; 2017;99:18–26. Available from: https://doi.org/10.1016/j.exger.2017.09.003

53. Álvarez C, Ramírez-Campillo R, Cristi-Montero C, Ramírez-Vélez R, Izquierdo M. Prevalence of non-responders for blood pressure and cardiometabolic risk factors among prehypertensive women after long-term high-intensity interval training. Front Physiol. 2018;9:1–13.

54. Ruegsegger GN, Booth FW. Health Benefits of Exercise. Cold Spring Harb Perspect Med [Internet]. 2018;8:a029694. Available from: http://www.ftrdergisi.com/eng/makale/3736/288/Full-Text

55. Leifer ES, Brawner C a., Fleg JL, Kraus WE, Whellan DJ, Piña IL, et al. Are there negative responders to exercise training among heart failure patients? Med Sci Sports Exerc. 2014;46:219–24.

56. Bonafiglia JT, Preobrazenski N, Gurd BJ. A Systematic Review Examining the Approaches

Used to Estimate Interindividual Differences in Trainability and Classify Individual Responses to Exercise Training. Front Physiol [Internet]. 2021;12:1–18. Available from: https://www.frontiersin.org/articles/10.3389/fphys.2021.665044/full

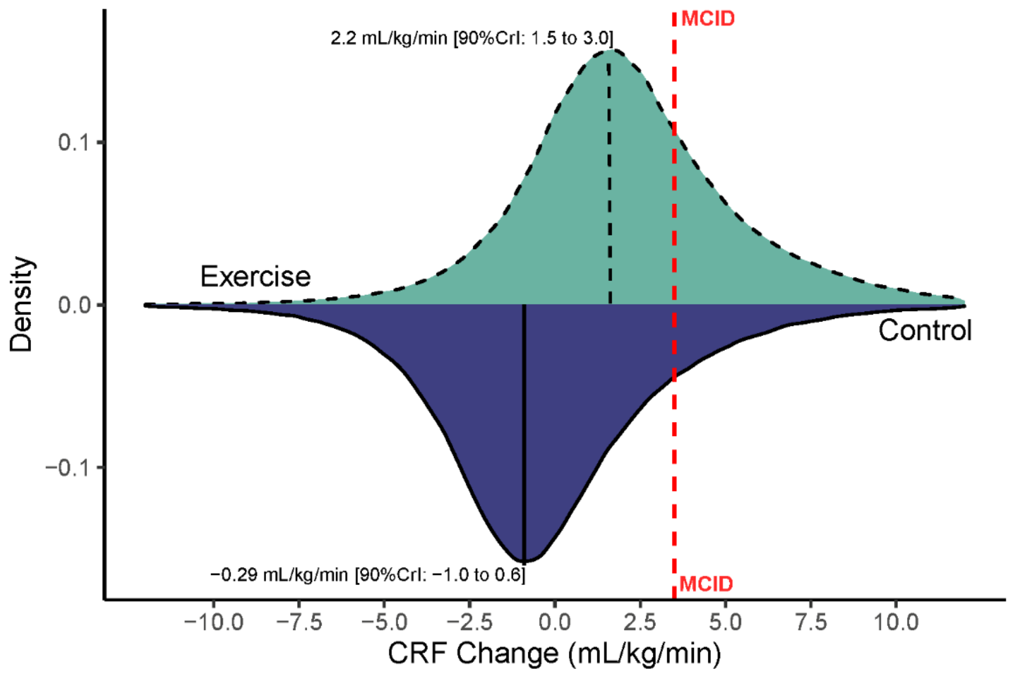
57. Schulhauser KT, Bonafiglia JT, McKie GL, McCarthy SF, Islam H, Townsend LK, et al. Individual patterns of response to traditional and modified sprint interval training. J Sports Sci [Internet]. Routledge; 2020;00:1–11. Available from: https://doi.org/10.1080/02640414.2020.1857507

58. Tenan MS, Simon JE, Robins RJ, Lee I, Sheean AJ, Dickens JF. Anchored Minimal Clinically Important Difference Metrics: Considerations for Bias and Regression to the Mean. J Athl Train [Internet]. 2021;56:1042–9. Available from: https://meridian.allenpress.com/jat/article/56/9/1042/448497/Anchored-Minimal-Clinically-Important-Difference

59. Vigotsky AD, Tiwari SR, Griffith JW, Apkarian AV. What Is the Numerical Nature of Pain Relief? Front Pain Res [Internet]. 2021;2:1–13. Available from: https://www.frontiersin.org/articles/10.3389/fpain.2021.756680/full

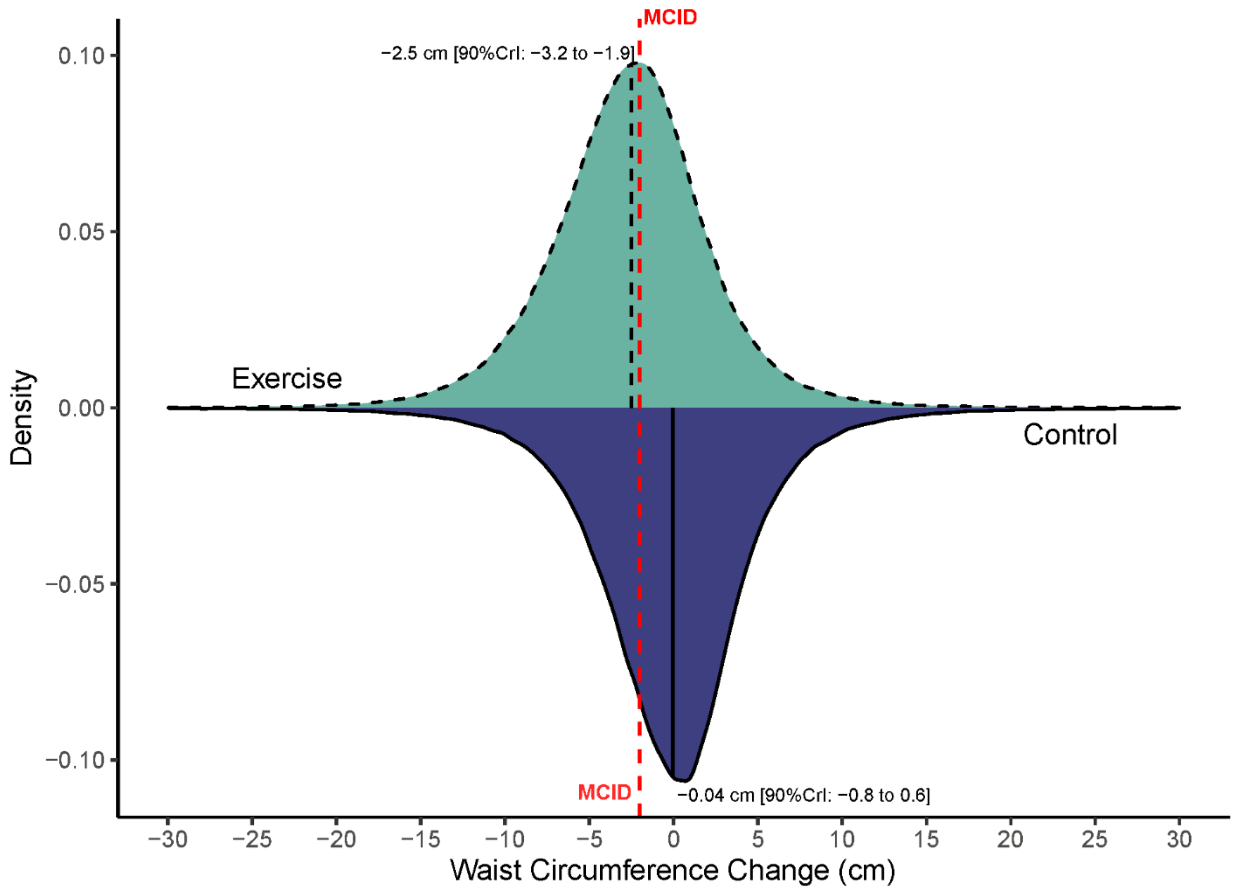
**Figurs**

**Figure 1.** Distribution of change score in cardiorespiratory fitness (CRF) to exercise (green) and control (blue).



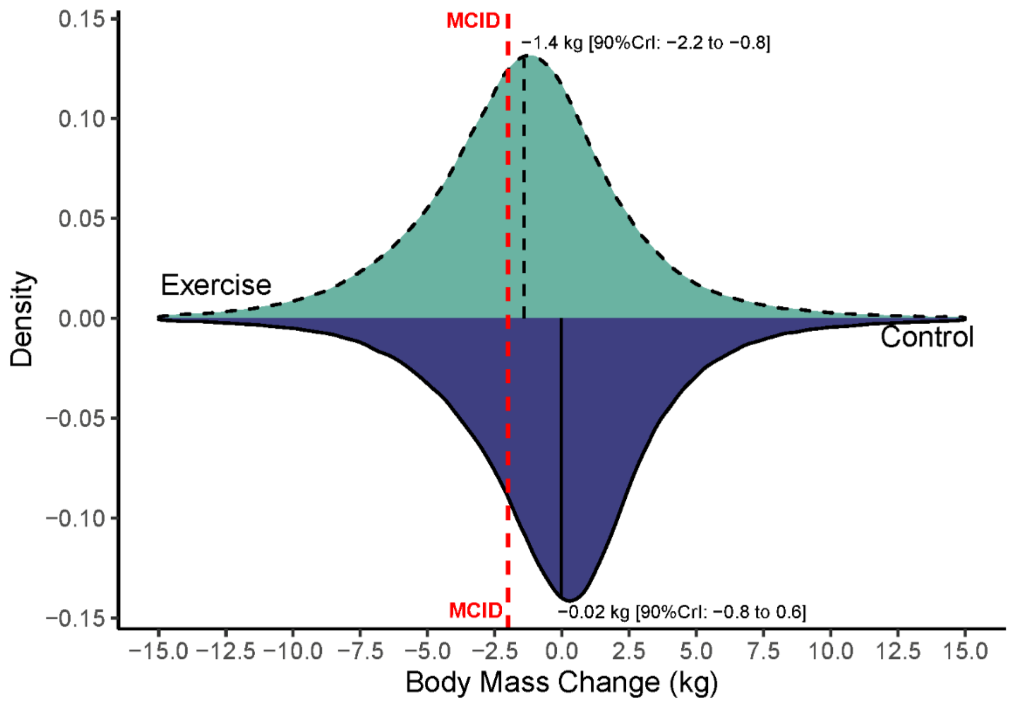

Black vertical lines represent estimated mean changes and the dashed red line represents the minimum clinically important difference of +3.5mL/kg/min. Standard deviations are not reported in figures but are illustrated as the width of the distribution curves. CrI, credible intervals.

**Figure 2.** Distribution of change score in waist circumference to exercise (green) and control (blue).




Black vertical lines represent estimated mean changes and the dashed red line represents the minimum clinically important difference of -2cm. Standard deviations are not reported in figures but are illustrated as the width of the distribution curves. CrI, credible intervals.

**Figure 3.** Distribution of change score in body mass to exercise (green) and control (blue).




Black vertical lines represent estimated mean changes and the dashed red line represents the minimum clinically important difference of -2kg. Standard deviations are not reported in figures but are illustrated as the width of the distribution curves. CrI, credible intervals.

**Table 1.** Participant characteristics, sample sizes, and exercise training amounts from the eight randomized controlled trials included in the present individual participant data meta-analysis.

| Trial | Participant Characteristics | Sample Size | Control Group? | Exercise Group 1 | Exercise Group 2 | Exercise Group 3 |
|---|---|---|---|---|---|---|
| DREW | Females who were inactive and postmenopausal (age: 57.2 ± 6.4) | 465 (0♂ \| 465♀) | Yes | **AT:** 4KKW at 50% $CRF_{max}$ 3-4d/wk for 6mo | **AT:** 8KKW at 50% $CRF_{max}$ 3-4d/wk for 6mo | **AT:** 12KKW at 50% $CRF_{max}$ 3-4d/wk for 6mo |
| E-MECHANIC | Sedentary males and females with overweight or obesity (age: 47.5 ± 12.0) | 195 (51♂ \| 144♀) | Yes | **AT:** 8KKW at 65-85% $CRF_{max}$ 3-5d/wk for 6mo | **AT:** 20KKW at 65-85% $CRF_{max}$ 3-5d/wk for 6mo | - |
| HART-D | Sedentary males and females with type 2 diabetes (age: 55.9 ± 8.8) | 269 (100♂ \| 169♀) | Yes | **AT:** 12KKW at 50-80% $CRF_{max}$ 3-5d/wk for 9mo | **RT:** 9 x (10-12 reps over ~2 sets) at 10-12RM 3d/wk for 9mo | **ATRT:** 10KKW at 50-80% $CRF_{max}$ 3-4d/wk and same RT program for 9mo |
| HEARTY | Inactive postpubertal male and female adolescents with overweight or obesity (age: 15.5 ± 1.3) | 138 (50♂ \| 88♀) | Yes | **AT:** 20-45min at 65-85% $HR_{max}$ 4d/wk for 6mo | **RT:** 7 x (8-15 reps over 2-3 sets) at 8-15RM 4d/wk for 6mo | **ATRT:** Same AT and RT program for 9mo |
| Queen's | Sedentary males and females with overweight or obesity (age: 51.1 ± 8.1) | 267 (91♂ \| 176♀) | Yes | **AT:** 180(F) or 300(M)kcal at 50% $CRF_{max}$ 5d/wk for 6mo | **AT:** 360(F) or 600(M)kcal at 50% $CRF_{max}$ 5d/wk for 6mo | **AT:** 360(F) or 600(M)kcal at 75% $CRF_{max}$ 5d/wk for 6mo |
| STRRIDE | Sedentary males and females with overweight or obesity (age: 52.6 ± 6.5) | 260 (139♂ \| 121♀) | Yes | **AT:** 14KKW at 40-55% $CRF_{max}$ for 7-8mo [a] | **AT:** 14KKW at 65-80% $CRF_{max}$ for 7-8mo [a] | **AT:** 23KKW at 65-80% $CRF_{max}$ for 7-8mo [a] |
| STRRIDE AT/RT | Sedentary males and females with overweight or obesity (age: 49.3 ± 10.2) | 155 (69♂ \| 86♀) | No | **AT:** 14KKW at 65-80% $CRF_{max}$ for 8mo [a] | **RT:** 8 x (8-12 reps over 3 sets) at 8-12RM for 8mo [a] | **ATRT:** Same AT and RT program for 8mo [a] |
| STRRIDE PD | Sedentary males and females with prediabetes (age: 60.5 ± 7.4) | 130 (50♂ \| 80♀) | No [b] | AT: 42KJKW at 50% $VO_2R$ for 6mo [a] | **AT:** 67KJKW at 50% $VO_2R$ for 6mo [a] | **AT:** 67KJKW at 75% $VO_2R$ for 6mo [a] |

Original methods or primary results publications: DREW [21], E-MECHANIC [22], HART-D [24], HEARTY [23], Queen's [25], STRRIDE [26], STRRIDE AT/RT [28], STRRIDE PD [27]. Age is written as mean ± standard deviation years. [a] Each participant could choose their desired exercise frequency. [b] Control group included lifestyle/dietary intervention, thus excluded from current study; AT, aerobic training; RT, resistance training; ATRT, combined aerobic and resistance training; KKW, kcals per kg body mass per week; $CRF_{max}$, maximal cardiorespiratory fitness; $HR_{max}$, maximal heart rate; $VO_2$, reserve oxygen consumption; KJKW, kilojoule per kg body mass per week; ♂, number of male participants; ♀, number of female participants.

**Table 2.** Analysis of relative cardiorespiratory fitness (CRF) change scores and moderator analyses involving exercise *vs.* control and exercise only comparisons.

| | **Exercise** (90% Credible intervals) | | | **Control** (90% Credible intervals) | | |
|---|---|---|---|---|---|---|
| ***Model or moderator*** | **N** | **Proportion ≥ MCID** | **Standard deviation (mL/kg/min)** | **N** | **Proportion ≥ MCID** | **Standard deviation (mL/kg/min)** |
| | | | ***Exercise vs. control*** | | | |
| ***Base Model*** | 1378 | 0.30 (0.21-0.41) | 3.4 (2.9-3.9) | 329 | 0.11 (0.05-0.19) | 3.5 (2.9-4.2) |
| | | | ***Exercise vs. control moderators*** | | | |
| ***Duration*** | | | | | | |
| 4 months | 158 | 0.20 (0.09-0.32) | 3.0 (2.2-3.6) | 23 | 0.07 (0.01-0.15) | 3.0 (2.2-3.8) |
| 6 months | 804 | 0.27 (0.17-0.39) | 3.4 (2.7-4.1) | 237 | 0.11 (0.04-0.20) | 3.5 (2.7-4.4) |
| 8 months | 416[1] | 0.35 (0.25-0.46) | 4.1 (3.3-5.2) | 69 [a] | 0.16 (0.08-0.26) | 4.3 (3.3-5.9) |
| ***Baseline BMI*** | | | | | | |
| -1SD | | 0.29 (0.20-0.40) | 3.4 (2.8-3.9) | | 0.11 (0.06-0.19) | 3.5 (2.9-4.3) |
| Mean | 1376 | 0.31 (0.22-0.42) | 3.5 (2.9-4.1) | 329 | 0.13 (0.07-0.21) | 3.8 (3.0-4.7) |
| +1SD | | 0.27 (0.18-0.39) | 3.2 (2.7-3.8) | | 0.11 (0.05-0.18) | 3.4 (2.7-4.1) |
| | | | ***Exercise only moderators*** | | | |
| ***Exercise Adherence*** | | | | | | |
| Low (<70%) | 73 | 0.21 (0.10-0.34) | 3.6 (3.0-4.2) | | | |
| High (≥70%) | 1252 | 0.30 (0.19-0.44) | 3.6 (2.9-4.1) | | | |
| ***Exercise Intensity*** [b] | | | | | | |
| Low (<60%) | 498 | 0.21 (0.09-0.34) | 3.6 (2.5-4.4) | | | |
| High (≥60%) | 690 | 0.37 (0.22-0.52) | 4.4 (3.3-6.5) | | | |
| ***Exercise Mode*** | | | | | | |
| Aerobic | 1188 | 0.28 (0.17-0.41) | 3.5 (2.9-4.0) | | | |
| Resistance | 97 | 0.24 (0.15-0.36) | 3.4 (2.8-3.9) | | | |
| Combined | 93 | 0.40 (0.31-0.50) | 4.9 (3.8-6.8) | | | |
| ***Exercise Amount*** [c] | | | | | | |
| Low | 145 | 0.22 (0.13-0.33) | 3.9 (3.1-5.2) | | | |
| Mid | 291 | 0.27 (0.16-0.39) | 4.4 (3.5-5.5) | | | |
| High | 749 | 0.36 (0.23-0.48) | 3.7 (2.9-4.5) | | | |

N: Number of individuals included in the IPD model. Proportion > MCID: The proportion estimated to meet or exceed the minimal clinically important clinical difference, with 90% credible intervals denoting Bayesian subjective probabilities. [a] Combines participants from intervention durations of 8 and 9 months. [b] Intensities were prescribed as percentages of different variables across studies (see Table 1 for details). [c] Low, mid, and high exercise amounts categorized as less than 500kcal, between 500 and 1000, and greater than 1000 kcals prescribed per sessions.

**Table 3.** Analysis of relative waist circumference change scores and moderator analyses involving exercise *vs.* control and exercise only comparisons.

| | **Exercise** (90% Credible intervals) | | | **Control** (90% Credible intervals) | | |
|---|---|---|---|---|---|---|
| ***Model or moderator*** | **N** | **Proportion ≥ MCID** | **Standard deviation (cm)** | **N** | **Proportion ≥ MCID** | **Standard deviation (cm)** |
| | | | ***Exercise vs. control*** | | | |
| ***Base Model*** | 1475 | 0.54 (0.48-0.61) | 4.9 (4.2-5.6) | 359 | 0.30 (0.23-0.38) | 5.7 (4.6-7.9) |
| | | | ***Exercise vs. control moderators*** | | | |
| ***Duration*** | | | | | | |
| 4 months | 159 | 0.52 (0.41-0.61) | 4.5 (3.6-5.6) | 31 | 0.26 (0.17-0.35) | 5.6 (4.0-8.9) |
| 6 months | 807 | 0.53 (0.46-0.61) | 4.8 (4.1-5.6) | 248 | 0.29 (0.20-0.37) | 5.7 (4.4-8.1) |
| 8 months | 509[1] | 0.54 (0.48-0.61) | 5.3 (4.4-6.4) | 80 [a] | 0.31 (0.23-0.40) | 5.8 (4.6-8.1) |
| ***Baseline BMI*** | | | | | | |
| -1SD | | 0.54 (0.48-0.61) | 4.8 (4.1-5.6) | | 0.29 (0.22-0.37) | 5.3 (4.4-7.2) |
| Mean | 1475 | 0.53 (0.46-0.61) | 4.4 (3.8-5.2) | 359 | 0.26 (0.19-0.35) | 4.9 (4.0-6.9) |
| +1SD | | 0.56 (0.50-0.62) | 5.2 (4.5-6.1) | | 0.32 (0.24-0.40) | 5.9 (4.8-8.2) |
| | | | ***Exercise only moderators*** | | | |
| ***Exercise Adherence*** | | | | | | |
| Low (<70%) | 98 | 0.39 (0.30-0.47) | 5.1 (4.1-6.1) | | | |
| High (≥70%) | 1325 | 0.56 (0.48-0.63) | 4.9 (4.2-5.7) | | | |
| ***Exercise Intensity*** [b] | | | | | | |
| Low (<60%) | 515 | 0.44 (0.35-0.55) | 4.9 (4.0-5.8) | | | |
| High (≥60%) | 681 | 0.54 (0.46-0.62) | 4.8 (4.0-5.7) | | | |
| ***Exercise Mode*** | | | | | | |
| Aerobic | 1196 | 0.53 (0.45-0.62) | 5.1 (4.3-5.9) | | | |
| Resistance | 140 | 0.46 (0.37-0.56) | 4.8 (4.0-5.9) | | | |
| Combined | 139 | 0.61 (0.54-0.68) | 5.2 (4.4-6.2) | | | |
| ***Exercise Amount*** [c] | | | | | | |
| Low | 142 | 0.43 (0.30-0.61) | 4.5 (3.5-6.5) | | | |
| Mid | 293 | 0.45 (0.35-0.53) | 4.7 (3.8-5.7) | | | |
| High | 759 | 0.56 (0.47-0.62) | 5.1 (4.2-6.1) | | | |

N: Number of individuals included in the IPD model. Proportion > MCID: The proportion estimated to meet or exceed the minimal clinically important clinical difference, with 90% credible intervals denoting Bayesian subjective probabilities. [a] Combines participants from intervention durations of 8 and 9 months. [b] Intensities were prescribed as percentages of different variables across studies (see Table 1 for details). [c] Low, mid, and high exercise amounts categorized as less than 500kcal, between 500 and 1000, and greater than 1000 kcals prescribed per sessions.

**Table 4.** Analysis of relative body mass change scores and moderator analyses involving exercise *vs.* control and exercise only comparisons.

| | **Exercise** (90% Credible intervals) | | | **Control** (90% Credible intervals) | | |
|---|---|---|---|---|---|---|
| ***Model or moderator*** | **N** | **Proportion ≥ MCID** | **Standard deviation (kg)** | **N** | **Proportion ≥ MCID** | **Standard deviation (kg)** |
| | | | ***Exercise vs. control*** | | | |
| ***Base Model*** | 1535 | 0.42 (0.34-0.50) | 4.1 (3.5-5.0) | 375 | 0.26 (0.18-0.35) | 4.6 (3.7-6.4) |
| | | | ***Exercise vs. control moderators*** | | | |
| ***Duration*** | | | | | | |
| 4 months | 159 | 0.47 (0.36-0.57) | 3.9 (3.0-5.1) | 31 | 0.28 (0.18-0.39) | 4.4 (3.2-7.3) |
| 6 months | 823 | 0.43 (0.35-0.51) | 4.1 (3.3-5.0) | 247 | 0.26 (0.18-0.35) | 4.5 (3.5-6.4) |
| 8 months | 553[1] | 0.39 (0.33-0.47) | 4.4 (3.5-5.5) | 97 [a] | 0.25 (0.17-0.33) | 4.7 (3.6-6.8) |
| ***Baseline BMI*** | | | | | | |
| -1SD | | 0.42 (0.35-0.51) | 3.9 (3.3-4.5) | | 0.25 (0.18-0.35) | 4.3 (3.5-5.7) |
| Mean | 1535 | 0.35 (0.27-0.46) | 3.4 (2.9-4.0) | 375 | 0.19 (0.11-0.29) | 3.8 (3.0-5.2) |
| +1SD | | 0.48 (0.41-0.56) | 4.5 (3.8-5.3) | | 0.31 (0.24-0.40) | 4.9 (4.0-6.7) |
| | | | ***Exercise only moderators*** | | | |
| ***Exercise Adherence*** | | | | | | |
| Low (<70%) | 108 | 0.31 (0.22-0.41) | 4.0 (3.2-4.9) | | | |
| High (≥70%) | 1376 | 0.42 (0.33-0.52) | 4.6 (3.6-6.2) | | | |
| ***Exercise Intensity*** [b] | | | | | | |
| Low (<60%) | 550 | 0.37 (0.26-0.48) | 4.0 (3.3-4.9) | | | |
| High (≥60%) | 699 | 0.43 (0.34-0.54) | 4.0 (3.3-4.8) | | | |
| ***Exercise Mode*** | | | | | | |
| Aerobic | 1249 | 0.42 (0.34-0.51) | 4.2 (3.4-5.1) | | | |
| Resistance | 141 | 0.30 (0.22-0.41) | 4.9 (3.8-5.1) | | | |
| Combined | 145 | 0.52 (0.46-0.59) | 4.4 (3.5-5.2) | | | |
| ***Exercise Amount*** [c] | | | | | | |
| Low | 145 | 0.25 (0.14-0.37) | 3.0 (2.6-3.5) | | | |
| Mid | 301 | 0.36 (0.27-0.47) | 3.4 (2.9-4.0) | | | |
| High | 803 | 0.44 (0.33-0.53) | 4.1 (3.5-6.0) | | | |

N: Number of individuals included in the IPD model. Proportion > MCID: The proportion estimated to meet or exceed the minimal clinically important clinical difference, with 90% credible intervals denoting Bayesian subjective probabilities. [a] Combines participants from intervention durations of 8 and 9 months. [b] Intensities were prescribed as percentages of different variables across studies (see Table 1 for details). [c] Low, mid, and high exercise amounts categorized as less than 500kcal, between 500 and 1000, and greater than 1000 kcals prescribed per sessions.

**Interindividual differences in trainability and moderators of cardiorespiratory fitness, waist circumference, and body mass responses: A large-scale individual participant data meta-analysis**

Jacob T. Bonafiglia [1], Paul A. Swinton [2], Robert Ross [1], Neil M. Johannsen [3,4], Corby K. Martin [4], Timothy S. Church [4], Cris A. Slentz [5], Leanna M. Ross [5], William E. Kraus [5,6], Jeremy J. Walsh [7], Glen P. Kenny [8,9], Gary S. Goldfield [8,10,11,12], Denis Prud'homme [8,13], Ronald J. Sigal [8,9,14], Conrad P. Earnest [15†], and Brendon J. Gurd [1*]

## Supplementary tables

**Supplemental Table S1.** Estimated means, standard deviations (SD), and proportions exceeding minimum clinically important difference (MCID) for changes in cardiorespiratory fitness (CRF) by participant characteristic subgroups.

| | Exercise (90% Credible intervals) | | | | Control (90% Credible intervals) | | | |
|---|---|---|---|---|---|---|---|---|
| **Subgroup** | **N** | **Mean (mL/kg/min)** | **SD (mL/kg/min)** | **Proportion ≥ MCID** | **N** | **Mean (mL/kg/min)** | **SD (mL/kg/min)** | **Proportion ≥ MCID** |
| Males | 461 | 3.2 (2.2 to 4.2) | 3.9 (3.3 to 4.4) | 0.43 (0.32 to 0.55) | 90 | -0.62 (-1.6 to 0.50) | 3.4 (2.8 to 4.2) | 0.11 (0.05 to 0.21) |
| Females | 948 | 1.9 (1.4 to 42.6) | 3.2 (2.7 to 3.6) | 0.27 (0.19 to 0.36) | 242 | -0.19 (-0.81 to 0.55) | 3.5 (2.9 to 4.4) | 0.10 (0.06 to 0.18) |
| Non-diabetic | 1102 | 2.8 (2.1 to 3.7) | 3.5 (3.0 to 3.9) | 0.39 (0.28 to 0.49) | 274 | 0.03 (-0.73 to 0.86) | 3.5 (3.0 to 4.3) | 0.14 (0.07 to 0.21) |
| Diabetic or pre-diabetic | 191 | 1.5 (0.91 to 1.8) | 2.6 (2.3 to 3.4) | 0.17 (0.11 to 0.21) | 10 | -0.68 (-2.4 to 0.44) | 3.4 (2.9 to 4.5) | 0.11 (0.08 to 0.17) |
| Adolescents | 85 | 1.4 (0.72 to 2.2) | 3.3 (2.9 to 4.1) | 0.18 (0.11 to 0.22) | 45 | -0.19 (-0.86 to 0.53) | 3.5 (3.0 to 3.9) | 0.16 (0.07 to 0.19) |
| Adult | 1293 | 2.4 (1.7 to 3.3) | 3.4 (2.8 to 3.8) | 0.32 (0.22 to 0.44) | 284 | -0.24 (-0.98 to 0.64) | 3.4 (2.8 to 4.2) | 0.12 (0.05 to 0.20) |
| Aerobic | 1188 | 2.3 (1.5 to 3.0) | 3.4 (2.7 to 3.8) | 0.28 (0.17 to 0.41) | 329 | -0.22 (-0.94 to 0.63) | 3.5 (3.1 to 4.6) | 0.11 (0.06 to 0.20) |
| Resistance | 97 | 2.1 (1.4 to 2.9) | 3.3 (2.7 to 3.8) | 0.26 (0.19 to 0.39) | 329 | -0.22 (-0.96 to 0.61) | 3.5 (3.1 to 4.3) | 0.11 (0.05 to 0.20) |
| Combined | 93 | 3.0 (2.2 to 3.8) | 4.4 (3.6 to 5.2) | 0.39 (0.29 to 0.49) | 329 | -0.24 (-0.95 to 0.59) | 3.6 (3.2 to 4.4) | 0.11 (0.05 to 0.19) |

**Supplemental Table S2.** Estimated means, standard deviations (SD), and proportions exceeding minimum clinically important difference (MCID) for changes in waist circumference by participant characteristic subgroups.

| | Exercise (90% Credible intervals) | | | | Control (90% Credible intervals) | | | |
|---|---|---|---|---|---|---|---|---|
| **Subgroup** | **N** | **Mean (cm)** | **SD (cm)** | **Proportion ≥ MCID** | **N** | **Mean (cm)** | **SD (cm)** | **Proportion ≥ MCID** |
| Males | 430 | -3.0 (-4.8 to -0.84) | 4.9 (4.2 to 5.6) | 0.58 (0.48 to 0.68) | 87 | 0.03 (-1.0 to 1.0) | 4.2 (3.5 to 5.3) | 0.29 (0.20 to 0.40) |
| Females | 1014 | -2.3 (-2.9 to -1.7) | 4.7 (4.1 to 5.5) | 0.52 (0.46 to 0.58) | 269 | -0.09 (-0.79 to 0.60) | 6.0 (4.7 to 7.4) | 0.30 (0.23 to 0.37) |
| Non-diabetic | 1081 | -2.3 (-3.1 to -1.7) | 4.8 (4.1 to 5.7) | 0.52 (0.45 to 0.60) | 272 | -0.01 (-0.85 to 0.75) | 5.9 (4.6 to 7.6) | 0.30 (0.22 to 0.39) |
| Diabetic or pre-diabetic | 308 | -1.5 (-1.9 to -1.3) | 4.1 (3.3 to 4.9) | 0.43 (0.39 to 0.47) | 35 | 0.99 (0.12 to 2.1) | 3.5 (3.2 4.1) | 0.20 (0.12 to 0.27) |
| Adolescents | 326 | -4.9 (-5.5 to -3.5) | 5.5 (5.2 to 5.8) | 0.70 (0.62 to 0.78) | 36 | -0.63 (-1.4 to 0.22) | 5.6 (5.0 to 6.1) | 0.29 (0.24 to 0.44) |
| Adult | 1389 | -2.3 (-2.9 to -1.7) | 4.7 (4.0 to 5.5) | 0.52 (0.45 to 0.59) | 307 | 0.06 (-0.69 to 0.75) | 5.6 (4.5 to 7.8) | 0.29 (0.22 to 0.37) |
| Aerobic | 1196 | -2.6 (-3.2 to -1.7) | 4.9 (4.2 to 5.7) | 0.54 (0.48 to 0.61) | 359 | -0.02 (-0.70 to 0.68) | 5.6 (4.5 to 7.6) | 0.29 (0.22 to 0.37) |
| Resistance | 140 | -2.4 (-3.0 to -1.8) | 4.7 (4.1 to 5.5) | 0.47 (0.42 to 0.57) | 359 | -0.02 (-0.70 to 0.67) | 5.6 (4.5 to 7.8) | 0.29 (0.22 to 0.37) |
| Combined | 139 | -2.9 (-3.6 to -2.3) | 5.1 (4.2 to 5.8) | 0.61 (0.53 to 0.68) | 359 | -0.03 (-0.72 to 0.65) | 5.6 (4.4 to 5.6) | 0.29 (0.23 to 0.38) |

**Supplemental Table S3.** Estimated means, standard deviations (SD), and proportions exceeding minimum clinically important difference (MCID) for changes in body mass by participant characteristic subgroups.

| | Exercise (90% Credible intervals) | | | | Control (90% Credible intervals) | | | |
|---|---|---|---|---|---|---|---|---|
| **Subgroup** | **N** | **Mean (cm)** | **SD (cm)** | **Proportion ≥ MCID** | **N** | **Mean (cm)** | **SD (cm)** | **Proportion ≥ MCID** |
| Males | 489 | -1.7 (-2.9 to -0.60) | 5.2 (4.0 to 6.6) | 0.45 (0.36 to 0.56) | 99 | 0.62 (-0.66 to 1.8) | 4.6 (3.5 to 6.2) | 0.23 (0.13 to 0.37) |
| Females | 1046 | -1.4 (-2.0 to -0.89) | 3.7 (3.3 to 4.4) | 0.41 (0.34 to 0.49) | 276 | -0.22 (-0.87 to 0.35) | 4.4 (3.6 to 6.7) | 0.27 (0.20 to 0.35) |
| Non-diabetic | 1123 | -1.8 (-2.6 to -1.1) | 3.7 (3.2 to 4.2) | 0.44 (0.35 to 0.54) | 287 | -0.48 (-1.3 to 0.23) | 4.1 (3.3 to 5.9) | 0.28 (0.19 to 0.39) |
| Diabetic or pre-diabetic | 326 | -1.0 (-1.5 to -0.62) | 3.4 (3.1 to 3.8) | 0.35 (0.31 to 0.42) | 36 | 0.18 (-1.7 to 1.1) | 4.5 (4.0 to 5.2) | 0.30 (0.25 to 0.48) |
| Adolescents | 86 | -1.0 (-1.8 to -0.31) | 5.6 (4.0 to 7.0) | 0.44 (0.33 to 0.48) | 52 | 1.6 (-0.05 to 2.4) | 6.2 (5.1 to 7.3) | 0.19 (0.15 to 0.23) |
| Adult | 1449 | -1.6 (-2.4 to -0.99) | 3.8 (3.4 to 4.3) | 0.42 (0.35 to 0.52) | 323 | -0.25 (-1.1 to 0.41) | 4.3 (3.6 to 5.8) | 0.26 (0.18 to 0.36) |
| Aerobic | 1249 | -1.4 (-2.2 to -0.71) | 3.9 (3.4 to 4.6) | 0.41 (0.33 to 0.50) | 375 | 0.03 (-0.77 to 0.73) | 4.4 (3.6 to 5.8) | 0.25 (0.18 to 0.34) |
| Resistance | 141 | -1.2 (-2.0 to -0.54) | 4.0 (3.4 to 4.8) | 0.38 (0.30 to 0.47) | 375 | 0.04 (-0.77 to 0.73) | 4.4 (3.6 to 6.0) | 0.25 (0.18 to 0.35) |
| Combined | 145 | -1.6 (-2.4 to -0.83) | 4.1 (3.5 to 5.0) | 0.48 (0.40 to 0.57) | 375 | 0.03 (-0.77 to 0.73) | 4.4 (3.6 to 5.8) | 0.25 (0.18 to 0.34) |