European Journal of Applied Physiology (2019) 119:889–900
https://doi.org/10.1007/s00421-019-04078-w

ORIGINAL ARTICLE

The application of repeated testing and monoexponential regressions to classify individual cardiorespiratory fitness responses to exercise training

Jacob T. Bonafiglia[1] · Robert Ross[1] · Brendon J. Gurd[1]

Received: 9 October 2018 / Accepted: 11 January 2019 / Published online: 21 January 2019

Abstract

Purpose We tested the hypothesis that monoexponential regressions will increase the certainty in response estimates and confidence in classification of cardiorespiratory fitness (CRF) responses compared to a recently proposed linear regression approach.

Methods We used data from a previously published RCT that involved 24 weeks of training at high amount–high intensity (HAHI; $N=28$), high amount–low intensity (HALI; $N=48$), or low amount–low intensity (LALI; $N=33$). CRF was measured at 0, 4, 8, 16, and 24 weeks. We fit the repeated CRF measures with monoexponential and linear regressions, and calculated individual response estimates, the error in these estimates ($TE_{MONOEXP}$ and TE_{SLOPE}, respectively), and 95% confidence intervals (CIs). Individuals were classified as responders, uncertain, or non-responders based on where their CI lay relative to a minimum clinically important difference. Additionally, responses were classified using observed pre–post-changes and the typical error of measurement.

Results Comparing the error in response estimates revealed that monoexponential regressions were a better fit than linear regressions for the majority of individual responses ($N=81/109$) and mean CRF data (mean $TE_{MONOEXP}$:TE_{SLOPE}; HAHI = 2.00:2.58, HALI = 1.91:2.46, LALI = 1.63:2.18; all $p<0.01$). Fewer individuals were confidently classified as responders with linear regressions ($N=29/109$) compared to monoexponential ($N=55/109$). Additionally, response estimates were highly correlated across all three approaches (all $r>0.92$).

Conclusions Future studies should determine the type of regression that best fits their data prior to classifying responses. The similarity in response estimates and classification from regressions and observed pre–post-changes questions the purported benefit of using repeated measures to characterize CRF responses to training.

Keywords Cardiorespiratory fitness · Repeated measures · Individual response · Individual regressions · Typical error · Non-responder

Abbreviations

ANOVA	Analysis of variance
BMI	Body mass index
CI	Confidence interval
CRF	Cardiorespiratory fitness
HAHI	High amount and high intensity
HALI	High amount and low intensity
LALI	Low amount and low intensity
MCID	Minimum clinically important difference
MET	Metabolic equivalent task
RCT	Randomized controlled trial
SEM	Standard error of measurement
TE	Typical error
$TE_{MONOEXP}$	Error in monoexponential regression
TE_{SLOPE}	Error in linear regression

Communicated by Keith Phillip George.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00421-019-04078-w) contains supplementary material, which is available to authorized users.

✉ Brendon J. Gurd
gurdb@queensu.ca

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

Introduction

Since Bouchard and colleagues' (1999) demonstration of interindividual heterogeneity in cardiorespiratory fitness (CRF) responses to standardized exercise training, there has been a growing interest in classifying individuals as CRF "responders" or "non-responders" (Sisson et al. 2009; Scharhag-Rosenberger et al. 2012; Astorino and Schubert 2014; Ross et al. 2015a; Bonafiglia et al. 2016; Gurd et al. 2016; Raleigh et al. 2016; Montero and Lundby 2017). Although the majority of the CRF non-response literature has used thresholds based on the typical error of measurement to dichotomously classify individuals as responders or non-responders (Scharhag-Rosenberger et al. 2012; Astorino and Schubert 2014; Ross et al. 2015a; Bonafiglia et al. 2016; Gurd et al. 2016; Raleigh et al. 2016; Montero and Lundby 2017), these response classification approaches have received criticism given their risk in misclassifying individual response (Williamson et al. 2017; Bonafiglia et al. 2018; Hecksteden et al. 2018). Further, dichotomous classification approaches involve calculating individual responses using only two time points (i.e., pre- and post-training) and may, therefore, be associated with large error in individual response estimates (Hecksteden et al. 2018). In light of these criticisms, recent reports have proposed alternative statistical approaches for classifying individual responses (Hecksteden et al. 2015, 2018; Swinton et al. 2018).

Hecksteden et al. (2018) recently described an individual response classification method that involves taking repeated measures over the course of a training intervention. Specifically, an individual's response was estimated as the slope of a linear regression of their observed measures over time (Hecksteden et al. 2018). The certainty in an individual's response estimate [i.e., the typical error (TE) of the slope] was used to calculate individual confidence intervals (CIs), and individuals were classified as "responders", "uncertain", or "non-responders" if their CI lay above, crossed, or fell below the smallest worthwhile change [herein referred to as the minimum clinically important difference (MCID)], respectively. Unlike individual response classification approaches that derive response estimates from pre–post-training changes (Hopkins 2000b; Swinton et al. 2018), Hecksteden et al. (2018) argued that classifying response based on repeated measures provides a more accurate estimate of an individual's true response. Additionally, although a repeated cross-over study design (participants are exposed to both the experimental and control condition more than once) can theoretically provide the most accurate estimate of an individual's response to an intervention (Senn et al. 2010), adopting this study design for an exercise training trial would be costly, time consuming, and risks introducing carryover and/or seasonal effects that may bias response estimates (Atkinson and Batterham 2015; Hecksteden et al. 2015, 2018). Accordingly, collecting repeated measures over the course of an intervention represents a practical alternative for classifying individual responses to training (Hecksteden et al. 2015, 2018).

Importantly, Hecksteden et al. (2018) performed linear regressions on CRF responses to a 1-year training program despite CRF adaptations to prolonged exercise training appearing to be non-linear (Scharhag-Rosenberger et al. 2009). Specifically, CRF adaptations to training appear to follow a monoexponential (non-linear) shape (Hickson et al. 1981; Govindasamy et al. 1992) as the rate of increase in CRF decreases over the course of prolonged exercise training (Scharhag-Rosenberger et al. 2009). Considering the non-linear shape of CRF adaptations to training, Hecksteden et al. (2018) suggest that using non-linear modelling may decrease the error in response estimates and thus improve the confidence in classification of individual responses. Therefore, the primary purpose of the present study was to test the hypothesis that, compared to individual linear regressions, non-linear regressions will decrease the error in CRF response estimates and consequently improve the ability to confidently classify individuals as CRF responders or non-responders. A secondary purpose of this study was to test the hypothesis that compared to the commonly adopted (Vollaard et al. 2009; Sisson et al. 2009; Bouchard et al. 2012; Scharhag-Rosenberger et al. 2012; Astorino and Schubert 2014; Ross et al. 2015a; Bonafiglia et al. 2016; Gurd et al. 2016; Raleigh et al. 2016, 2018; Montero and Lundby 2017) approach of deriving response estimates from observed pre–post-training differences, estimating responses from regressions of repeated measures will reduce the error in response estimates and thus increase the number of participants that can be confidently classified. To test these hypotheses, we utilized CRF data from a rigorously controlled 24-week randomized controlled trial (Ross et al. 2013, 2015a, b). Given the growing interest in the application of personalized exercise-based medicine (Buford et al. 2013), this study aims to better inform exercise scientists and clinicians about the strengths and weaknesses of several approaches for deriving individual CRF response estimates and classifying CRF responses to exercise training.

Methods

Trial design

Details of the trial design and methods (Ross et al. 2013), and findings from the primary (Ross et al. 2015b) and secondary analysis (Ross et al. 2015a) have been published elsewhere. Briefly, 300 sedentary and abdominally obese adults were randomized to a control group ($n = 75$) or to one of three exercise training groups: low amount (kcals per session: 180 for females and 300 for males) and low intensity (50% of maximal oxygen consumption; LALI; $n = 73$), high amount (kcals per session: 360 for females and 600 for males) and low intensity (50% of maximal oxygen consumption; HALI; $n = 76$), or high amount (kcals per session: 360 for females and 600 for males) and high intensity (75% of maximal oxygen consumption; HAHI; $n = 76$). All training involved supervised walking/jogging on a treadmill five times per week. All 300 participants did not complete the entire intervention and dropout information (number and reasons for dropout) is presented in the primary analysis (Ross et al. 2015b). Sample sizes for the present analysis are presented in Table 1.

Participants in the control group did not complete repeated CRF testing over the course of the intervention (Ross et al. 2015a) and are, therefore, excluded from the repeated measures analysis. However, baseline (pre)- and 24-week (post)-CRF data from the control group were used to calculate the typical error of measurement (described below). Participants in the exercise groups completed 24 weeks of supervised training that consisted of five treadmill exercise sessions per week. Training attendance and adherence data are reported in the primary analysis (Ross et al. 2015b). CRF was measured at baseline and after 4, 8, 16, and 24 weeks of training and only participants with CRF data at each time point were included in the current analysis (baseline characteristics are presented in Table 1).

CRF measurement

CRF was measured during a graded exercise test using standard open-circuit spirometry techniques (SensorMedics) as previously described (Ross et al. 2015a). Briefly, participants walked on a treadmill at a self-selected speed at zero elevation for 3 min, after which the incline was increased by five percent for 2 min, then by two percent every subsequent 2 min until volitional fatigue.

Repeated measures analysis overview

Although previous attempts to define the shape of CRF adaptations to exercise training are limited with small sample sizes, short durations (i.e., ~ 10 weeks), and/or prescribed an inconsistent exercise dose over the course of training (Hickson et al. 1981; Govindasamy et al. 1992; Morris et al. 2002; Gass et al. 2004; Murias et al. 2010; Astorino et al. 2013), these studies have generally indicated that CRF adaptations are best described with exponential or linear fits (Hickson et al. 1981; Govindasamy et al. 1992; Morris et al. 2002; Gass et al. 2004). Therefore, to test the hypothesis that non-linear modelling improves the accuracy in individual response classification, we compared exponential (non-linear) and linear regressions. All regressions were performed on GraphPad Prism (v. 5.01; GraphPad Software, Inc., La Jolla CA, USA), and exponential regressions were run according to Prism's one-phase association (i.e., monoexponential) equation:

$$Y_i = Y_0 + \left(\text{Plateau} - Y_0\right) \times \left(1 - e^{(-K \times x_i)}\right), \quad (1)$$

where Y_i is the observed CRF value at a given time, Y_0 is the predicted CRF value at baseline (y-intercept), plateau is the highest predicted CRF value at infinite time, K is the time constant value (reciprocal of tau), and x_i is time. For the purpose of our analysis, all the time points were expressed relative to 24 weeks; 0, 4, 8, 16, and 24 weeks were assigned numerical values of 0.0, 0.167, 0.333, 0.667, and 1.0, respectively.

We first performed monoexponential and linear regressions on mean CRF data for each exercise group (HAHI, HALI, and LALI). Based on the results from this group analysis (see details below) we then performed individual monoexponential and linear regressions and calculated individual CIs. Results from the individual regression analysis were subsequently utilized to classify individuals as "responders", "uncertain", or "non-responders" according to the methods outlined by Hecksteden et al. (2018) (described below).

Table 1 Baseline characteristics

	HAHI ($N = 28$)	HALI ($N = 48$)	LALI ($N = 33$)	Control ($N = 42$)
Age (years)	53 ± 8	52 ± 8	55 ± 6.6	51 ± 8.3
CRF (L/min)	2.74 ± 0.76	2.69 ± 0.63	2.60 ± 0.65	2.74 ± 0.82
CRF (mL/kg/min)	28.3 ± 5.4	28.9 ± 5.0	28.0 ± 5.5	29.7 ± 5.9
BMI (kg/m^2)	37.6 ± 7.0*†	36.0 ± 5.5†	33.0 ± 6.9	32.2 ± 4.3

All data mean ± standard deviation

CRF cardiorespiratory fitness, *BMI* body mass index

*Significantly different than LALI ($p < 0.05$)

†Significantly different than control ($p < 0.05$)

Group-level regressions

We ran both monoexponential and linear regressions between time (independent variable) and mean observed CRF values (dependent variable) separately for each exercise group. The goodness of the monoexponential and linear fits was compared by assessing the error associated with each fit (described in more detail below).

Individual linear regressions

Following the methods outlined by Hecksteden et al. (2018), we ran linear regressions for each participant in the exercise groups that was included in our analysis. The slope of the linear regressions represented each individual's response estimate and the TE of the slope was calculated using the following equation:

$$TE_{SLOPE} = \frac{\sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{3}}}{\sqrt{\sum (x_i - \bar{x})^2}}, \tag{2}$$

where $\sum (y_i - \hat{y}_i)^2$ represents the sum of squared differences between the observed (actual) CRF value (y_i) and the predicted CRF value ($\hat{y}_i$) at each time point (i), $\sum (x_i - \bar{x})^2$ represents the sum of squared differences between each numerical time value (x_i) and the mean time value ($\bar{x}$), and 3 represents the degrees of freedom with five timepoints ($df = n - 2$). Each individual's TE_{SLOPE} was used to calculate individual 95% CIs using the following equation (Swinton et al. 2018):

$$\text{Individual 95\% CI} = \text{Response estimate} \pm (1.96 \times \text{Error}), \tag{3}$$

where the response estimate and error were the slope and TE_{SLOPE} of an individual's linear regression, respectively (see Fig. 2 for example). Individuals were then classified as "responders", "uncertain", or "non-responders" if their 95% CI lay above, crossed, or fell below the MCID (see Fig. 3 for examples). We chose a MCID of 1 MET (3.5 mL/kg/min) because a 1 MET increase in CRF confers a ~ 12–20% decrease in risk of all-cause mortality (Ross et al. 2016).

Individual monoexponential regressions

We extended the methods outlined by Hecksteden et al. (2018) to run individual monoexponential regressions. We constrained the time constant (K) value for each individual to match the K value associated with their respective group-level monoexponential regression. Constraining the K value assumed that the predicted rate of CRF adaptation (i.e., the shape of the monoexponential curves) was the same for each individual within a given exercise group (K values: HAHI = 2.583, HALI = 3.155, LALI = 4.023). We feel this assumption is justified because failure to constrain the K value during individual regression analysis resulted in physiologically impossible predicted peak CRF values for several participants (Supplemental Fig. 1).

Similar to the approach using linear regressions, individual response estimates ($\Delta\hat{Y}$)were calculated as the predicted CRF at 24 weeks minus the predicted CRF at baseline. Specifically, the y-intercept (Y_0) of the monoexponential regression was used as the predicted CRF value at baseline and the predicted CRF value at 24 weeks was calculated using Eq. 1 with $x = 1.0$ (i.e., the numerical value for 24 weeks, see above). Similar to TE_{SLOPE}, the TE of each individual's monoexponential regression ($TE_{MONOEXP}$) was calculated using Eq. 2. Individual 95% CIs were then calculated using Eq. 3 with $\Delta\hat{Y}$ and $TE_{MONOEXP}$ inputted as the response estimate and error, respectively. These individual 95% CIs were then used to classify response relative to the MCID following the same procedures outlined above (see Fig. 2 for example).

Observed pre–post-changes

Following recent recommendations (Williamson et al. 2017), the typical error of measurement was calculated using the pre–post-CRF data in the control group with the following equation (Hopkins 2000a):

$$\text{Typical error of measurement} = \frac{SD_{diff}}{\sqrt{2}}, \tag{4}$$

where SD_{diff} is the standard deviation of the difference scores (i.e., post/24-week CRF minus pre/baseline CRF) for control group participants. Unlike the analysis using repeated measures, we derived response estimates for this analysis as the observed change in CRF from pre-training (0 weeks) to post-training (24 weeks) for participants in the three exercise groups. Individuals were then classified following the same procedures described above but with observed changes and the typical error of measurement inputted as the response estimate and error in Eq. 3, respectively. Because this approach does not produce an individualized error estimate, the CIs were the same width for all participants. We compared the width of this CI to the average CI width from individual linear and monoexponential regressions.

Additional statistical analysis

The goodness of fit for the monoexponential and linear models of mean CRF data was compared using the TE of each fit (Eq. 2; described above). The average of the individual $TE_{MONOEXPS}$ and TE_{SLOPES} was calculated for each group and these values were compared within each group (i.e., average $TE_{MONOEXP}$ vs. average TE_{SLOPE}) using paired *t* tests (separate *t* test for each group). Paired *t* tests were also used to compare the average width of individual CIs when using monoexponential and linear regressions within each group. Three separate bivariate correlations were run to assess the similarity of response estimates derived from monoexponential regressions, linear regressions, and observed pre–post-changes. These bivariate correlations were run separately for each exercise group and also on pooled data across all three groups. Additionally, Chi-square analyses were performed to compare the number of participants that could not be confidently classified (uncertain) vs. the number of participants that were confidently classified (responder or non-responder) between the linear and monoexponential regression approaches. Separate 2 × 2 Chi-square tests (independent variable: linear or monoexponential regression; dependent variable: uncertain or not uncertain) were conducted for total (pooling all participants together), HAHI, HALI, and LALI (four Chi-square tests in total). Lastly, one-way analysis of variance (ANOVA) and Bonferroni post hoc analyses were used to compare baseline characteristics across groups. All data presented in the text of "Results" are reported as mean ± standard deviation.

Results

Baseline characteristics and mean regressions

With the exception of body mass index (BMI), which significantly ($p < 0.05$) differed between control vs. HAHI, control vs. HALI, and LALI vs. HALI (Table 1), there were no other significant differences in baseline characteristics between groups (Table 1). Figure 1 presents the mean monoexponential and linear regressions for HAHI, HALI, and LALI. For each exercise group, the TE_{SLOPE} exceeded the $TE_{MONOEXP}$, suggesting that the group mean CRF data were best fit with a monoexponential regression (Fig. 1).

Individual monoexponential regressions vs. individual linear regressions

Similar to the group-level regressions, TE_{SLOPE} exceeded the $TE_{MONOEXP}$ for most individuals in HAHI (21 of 28), HALI (38 of 48), and LALI (22 of 33). Paired *t* tests revealed that the mean TE_{SLOPE} (average of all individual TE_{SLOPES}) was significantly higher than the mean $TE_{MONOEXP}$ (average of all individual $TE_{MONOEXPS}$) for HAHI ($TE_{SLOPE} = 2.58 \pm 0.97$; $TE_{MONOEXP} = 2.00 \pm 0.79$; $p < 0.01$), HALI ($TE_{SLOPE} = 2.46 \pm 0.95$; $TE_{MONOEXP} = 1.91 \pm 0.81$; $p < 0.01$), and LALI ($TE_{SLOPE} = 2.18 \pm 0.84$; $TE_{MONOEXP} = 1.63 \pm 0.72$; $p < 0.01$). The average individual CI width was also significantly ($p < 0.05$) smaller when using individual monoexponential regressions compared to linear regressions for all three groups (monoexponential CI width vs. linear CI width: HAHI = 7.8 ± 3.1 vs. 10.1 ± 3.8, HALI = 7.5 ± 3.2 vs. 9.7 ± 3.7, LALI = 6.4 ± 2.8 vs. 8.5 ± 3.3). The significantly smaller TE and CI widths associated with monoexponential regressions resulted in more individuals that were classified as responders when using monoexponential rather than linear regressions (Table 2). Chi-square analyses revealed a significant ($p < 0.01$) difference in the number of participants classified as uncertain between monoexponential and linear regressions for HALI and total (pooling all participants together) but not HAHI ($p = 0.10$) or LALI ($p = 0.12$).

Figure 2 presents a representative monoexponential regression, linear regression, and individual CI for one participant to demonstrate the impact of TE_{SLOPE} exceeding $TE_{MONOEXP}$ on response classification. As seen in Fig. 2, despite this participant being classified as a responder with a monoexponential regression, they were classified as uncertain with a linear regression because their comparatively larger TE_{SLOPE} resulted in a larger individual CI that crossed the MCID. Similar to the representative participant presented in Fig. 2, several of the participants that were classified as a responder using a monoexponential regression were classified as uncertain with a linear regression (total = 28(uncertain with linear fit)/55(responders with monoexponential fit): HAHI = 6/18; HALI = 15/26; LALI = 7/11). Conversely, only two participants (one in HALI and one in LALI) were classified as uncertain with a monoexponential regression and a responder with a linear regression. Very few individuals were classified as non-responders ($n = 4$ of 109 total participants) and these participants were consistently classified as non-responders with both regression approaches (Table 2).

Individual monoexponential regressions

To illustrate our proposed extension of Hecksteden et al.'s (2018) approach, Fig. 3 presents individual monoexponential regressions, individual CIs, and response classification for six representative participants. These participants were not included in Fig. 3 to represent the observed mean change in CRF following exercise training or based on any baseline characteristic (e.g., age, sex, baseline CRF, etc.). Instead, these six participants are highlighted because they represent the six possible combinations of response estimates/$TE_{MONOEXP}$ and response classification that generally characterized most participants in our analysis. Specifically, participants were classified as responders if they had a large response estimate (e.g., > 8 mL/kg/min; Fig. 3a) or if they had a response estimate that barely exceeded the MCID but also had a low $TE_{MONOEXP}$ (Fig. 3b). Participants were classified as uncertain if they had a large $TE_{MONOEXP}$ (Fig. 3c) or a response estimate that approximated the MCID (Fig. 3d).

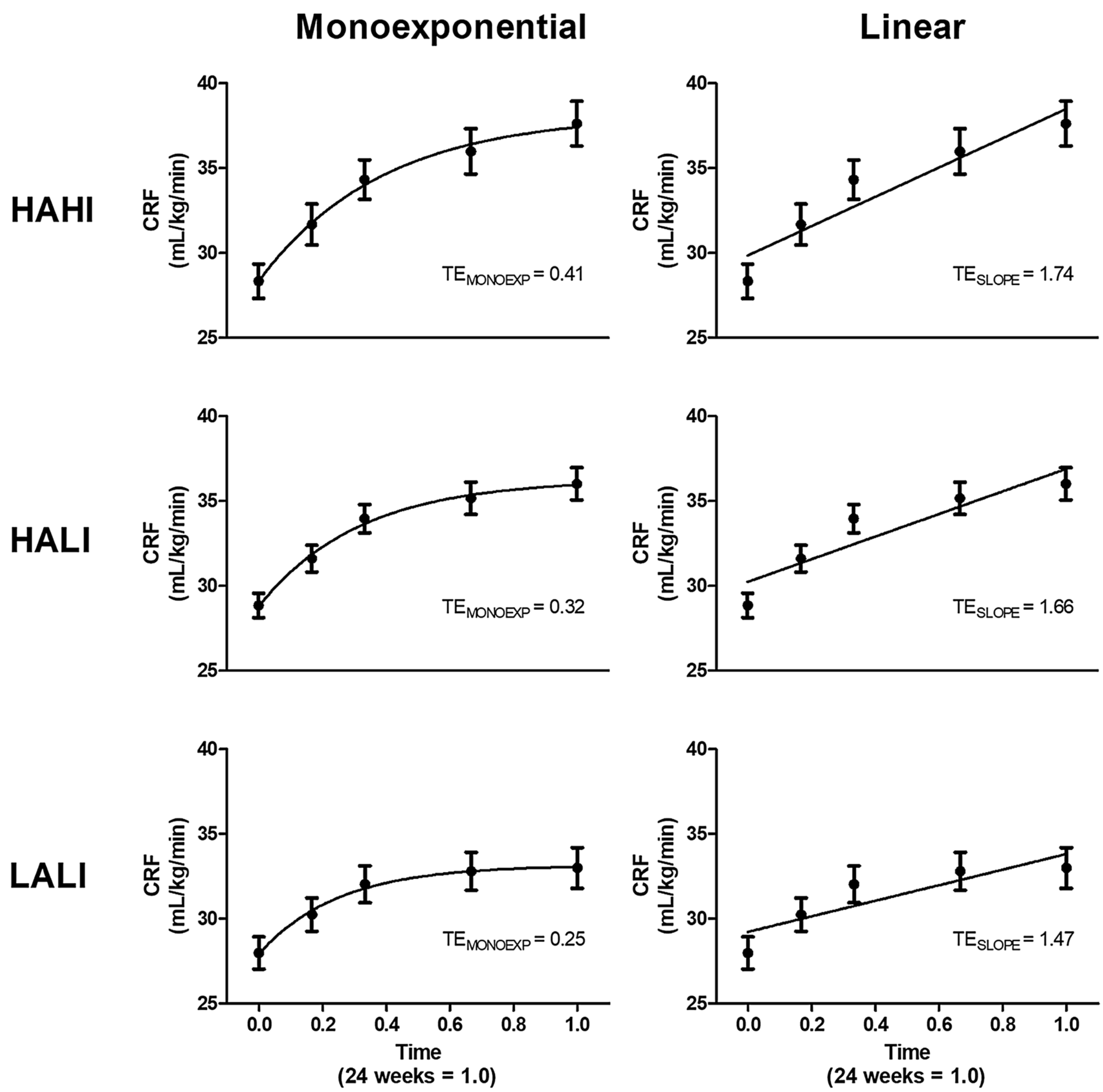

Fig. 1 Monoexponential and linear regressions of mean cardiorespiratory fitness (CRF) data for HAHI (high amount and high intensity), HALI (high amount and low intensity), and LALI (low amount and low intensity). $TE_{MONOEXP}$ and TE_{SLOPE} values were calculated using Eq. 2 and are expressed in CRF units (mL/kg/min). Data are plotted as mean ± SEM

Table 2 Number and percentage of individuals classified as responders, uncertain, or non-responders using individual linear or monoexponential regressions

	HAHI (N = 28)	HALI (N = 48)	LALI (N = 33)	Total (N = 109)
Responders				
Linear	12 (43%)	12 (25%)	5 (15%)	29 (27%)
Monoexponential	18 (64%)	26 (54%)	11 (33%)	55 (50%)
Uncertain				
Linear	16 (57%)	36 (75%)[a]	24 (73%)	76 (69%)[a]
Monoexponential	10 (36%)	22 (46%)	18 (55%)	50 (46%)
Non-responders				
Linear	0 (0%)	0 (0%)	4 (12%)	4 (4%)
Monoexponential	0 (0%)	0 (0%)	4 (12%)	4 (4%)

[a]Significantly greater than monoexponential ($p < 0.01$)

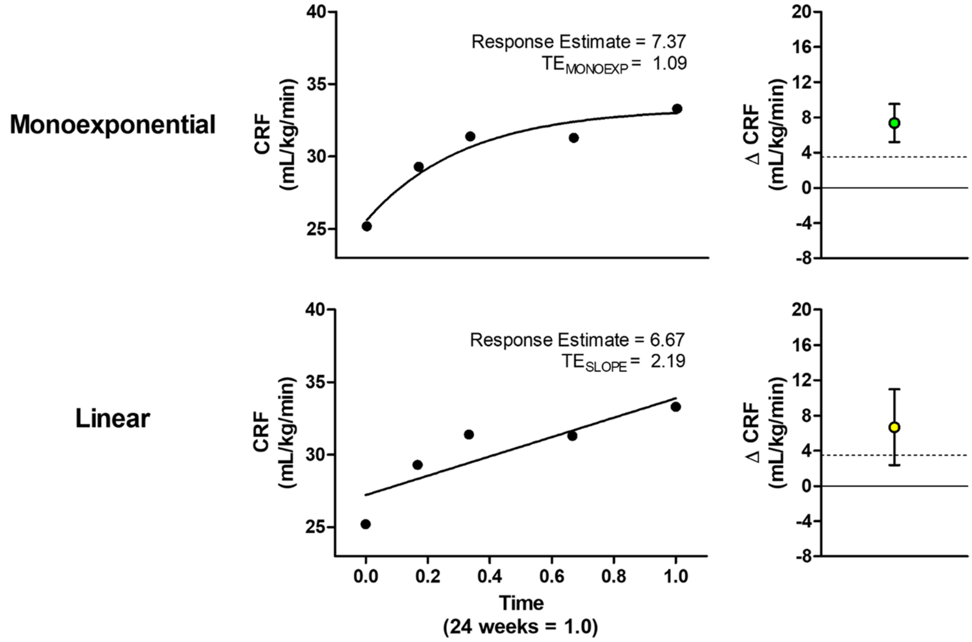

Fig. 2 Individual monoexponential (upper left panel) and linear (bottom left panel) regressions, and the corresponding individual CIs and response classification (right panels) for one representative participant. The green circle represents a classification of "responder" whereas the yellow circle represents a classification of "uncertain". The dashed line represents the MCID (1 MET). Refer to the text for information regarding the calculation of response estimates and their error ($TE_{MONOEXP}$/TE_{SLOPE}), and for the rationale of illustrating this participant's data

Finally, participants were classified as non-responders if they had a response estimate that fell below (Fig. 3e) or approximated zero (Fig. 3f).

Individual regressions vs. observed pre–post-changes and TE

The analysis performed for our secondary purpose revealed that using the typical error of measurement (derived from observed pre–post-changes in the control group) resulted in a wider CI (8.32 mL/kg/min) than the average CI width (7.24 mL/kg/min) from individual monoexponential regressions but not linear regressions (9.44 mL/kg/min). The larger CI width with the typical error of measurement resulted in slightly more participants being classified as uncertain (51%, 56/109 of total participants) compared to our analysis using monoexponential regressions (46%, 50/109 of total participants; Table 2) but not linear regressions (69%, 76/109 of total participants; Table 2). Individual response estimates derived from monoexponential regressions, linear regressions, and observed pre–post-changes were highly correlated for all groups and for the pooled data across groups (all bivariate correlations had r values > 0.92). Figure 4 presents individual response estimates and corresponding CIs derived from all three of these approaches (contains HAHI participant responses as representative data). Data in Fig. 4 are participant-matched and are presented in ascending order based on response estimates derived from monoexponential regressions (i.e., participant 1 had the smallest response estimate). The strong correlations and the individual data presented in Fig. 4 demonstrate that despite using three distinct approaches, a given individual's response estimates appear to be quite similar regardless of the approach used. The similarity in individual response estimates can also be seen in the trend that response estimates generally increased from participant 1–28 in each panel of Fig. 4 suggesting that all three approaches produced similar individual response estimates. Although Fig. 4 only contains data from HAHI, we found similar results with HALI and LALI (data not shown).

Discussion

The current study used CRF data from a large RCT to test the hypothesis that using individual monoexponential regressions rather than linear regressions improves the ability to confidently classify individuals as CRF responders or non-responders. The novel findings are (1) the smaller error in the group-level monoexponential regressions suggested that monoexponential curves were a better fit than linear regressions for the mean CRF data, (2) the average error and CI width of the individual monoexponential regressions were significantly smaller than the average error and CI width of the linear regressions, (3) very few participants were classified as non-responders and these participants were consistently classified as non-responders regardless of the regression used to derive their response estimate, and (4) compared to individual linear regressions, the smaller error in monoexponential regressions resulted in narrower CIs and an increased number of individuals who were more confidently classified as CRF responders in all three groups. Taken together, these findings suggest that the non-linear shape of CRF adaptations in our dataset was best evaluated with monoexponential regressions.

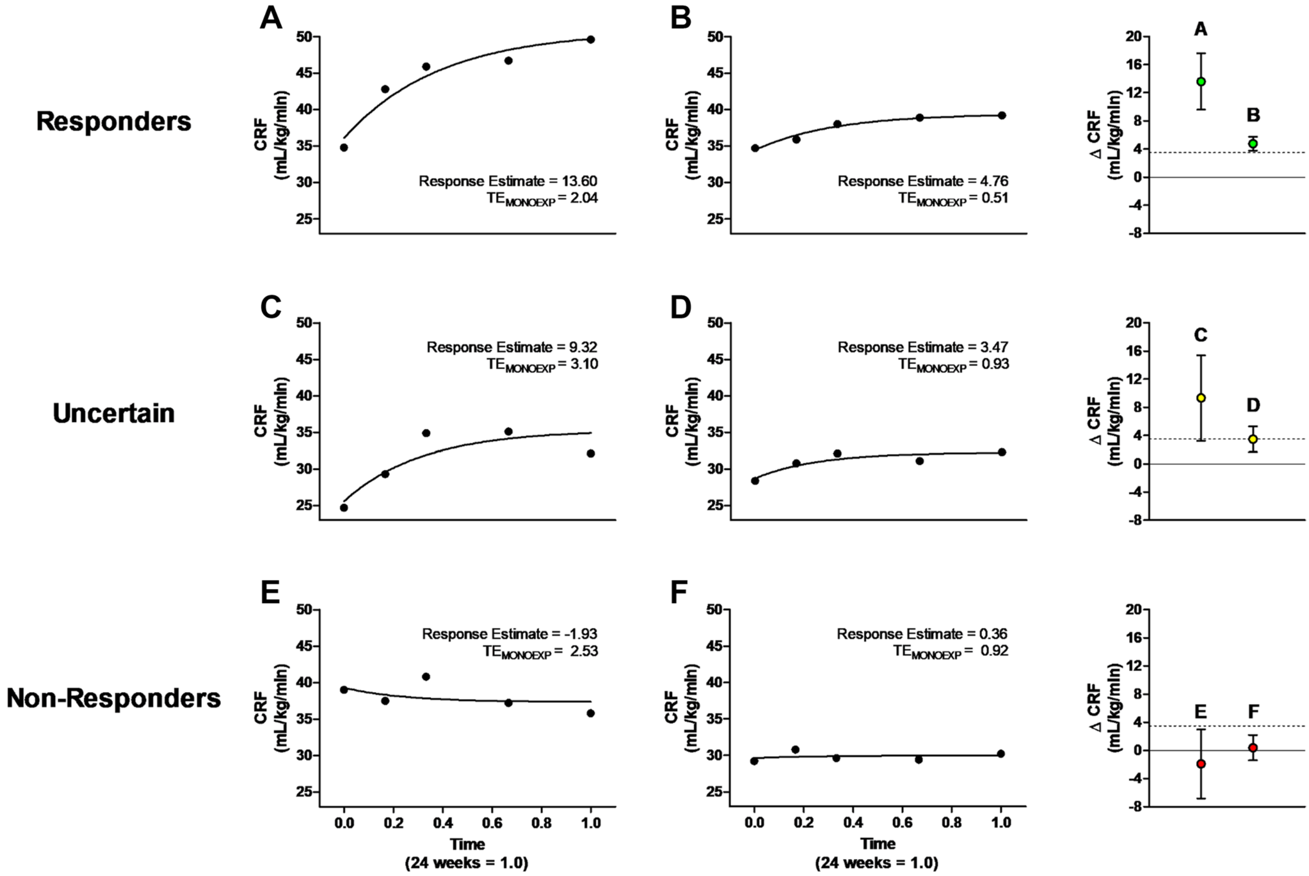

Fig. 3 Individual monoexponential regressions (A–F), and individual CIs and response classification (right panels) for six representative participants. Green circles represent responders, yellow circles represent uncertain, and red circles represent non-responders (right panels). The letters in the right panel are matched to the letters of the individual monoexponential regressions panels. The dashed line represents the MCID (1 MET). Refer to the text for information regarding the calculation of response estimates and their error ($TE_{MONOEXP}$). These participants are highlighted because they represent the six possible combinations of response estimates/$TE_{MONOEXP}$ and response classification that generally characterized most participants in our analysis. The participant in **a** completed HAHI, **b** and **c** completed HALI, and **d**–**f** completed LALI

As we move toward the application of personalized exercise-based medicine (Buford et al. 2013), an individual's response classification may be used to guide exercise prescription decision-making. For example, a recent study by Montero and Lundby (2017) increased the prescribed training frequency for participants classified as maximal power output "non-responders" in an attempt to increase these participants' observed responses. Because the use of monoexponential regressions decreased the number of individuals who could not be classified with confidence (i.e., "uncertain"), this approach may be superior for making exercise prescription decisions based on the classification of CRF responders or non-responders. However, future work using data simulations where true responses and measurement error are known is needed to compare the performance of different response classification approaches.

Additionally, our analysis for our secondary purpose revealed that using observed pre–post-changes and the typical error of measurement resulted in only slightly more individuals ($n = 6$) being classified as uncertain compared to our monoexponential regression approach. CRF response estimates were similar (Fig. 4) and highly correlated across approaches suggesting that using pre–post-changes results in similar response estimates as regressions of repeated measures. These data contradict the suggestion that using repeated measures will provide more accurate estimates of individual CRF responses (Hecksteden et al. 2018) and questions the benefit of using repeated measures when classifying CRF response to exercise training compared to the observed pre–post-change approach. Importantly, this suggests that the additional cost, time, and participant burden associated with repeated measures may not be warranted in future exercise training studies seeking to characterize individual CRF response. Further, given the possibility that short-term training studies (e.g., 2–6 weeks) are limited in the number of repeated measures that can be collected

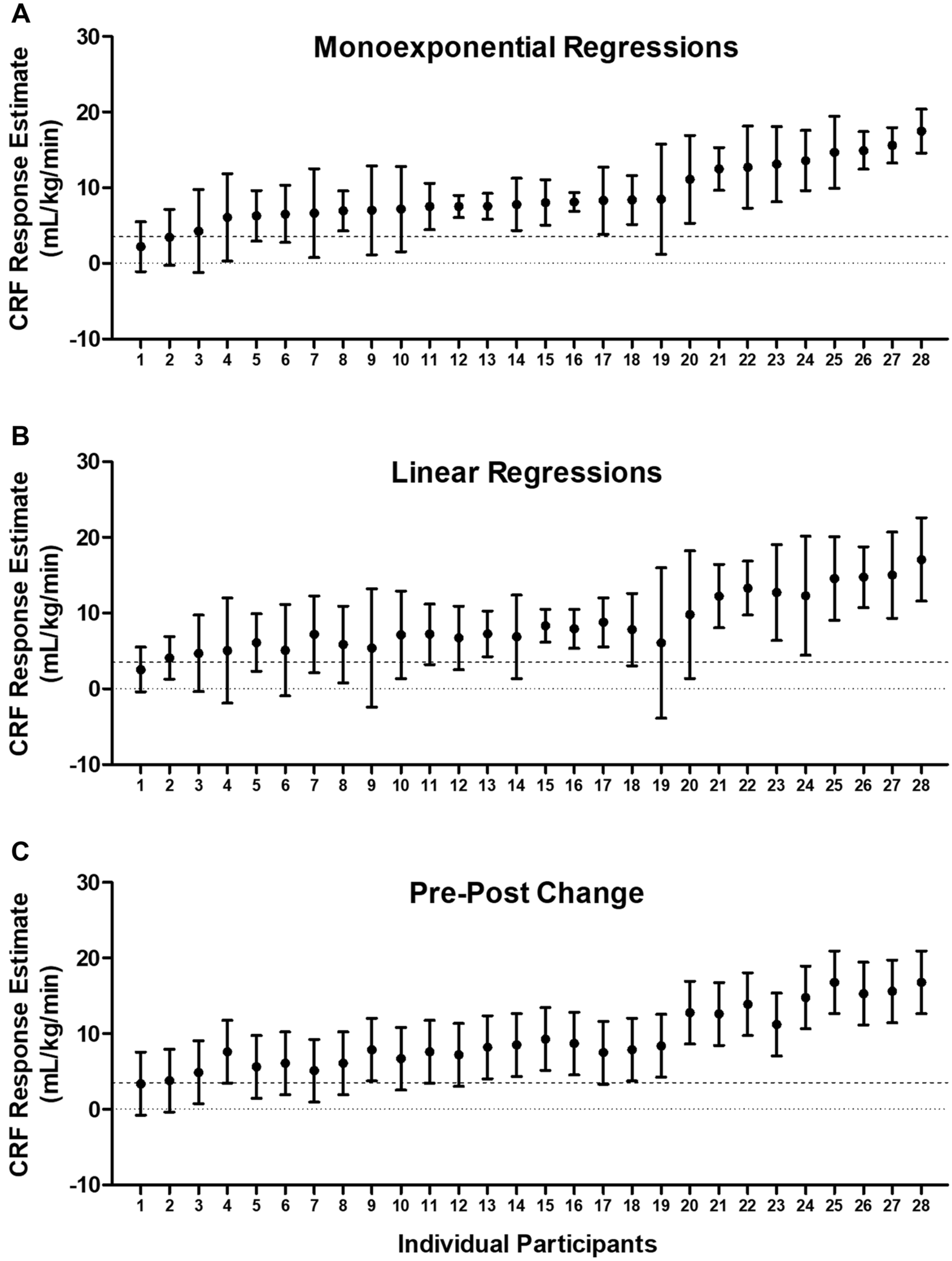

Fig. 4 Individual response estimates and CIs derived from monoexponential regressions (**a**), linear regressions (**b**), and observed pre–post-changes (**c**) for participants in HAHI ($n=28$). Participant numbers are matched across panels and are presented in ascending order based on response estimates from monoexponential regressions (**a**). Black dashed line represents the MCID of 1 MET (3.5 mL/kg/min)

throughout the intervention, it may be more suitable to classify CRF responses to short-term exercise training using the observed pre–post-change approach. However, future work is needed to examine the appropriateness of collecting repeated measures and subsequently classifying CRF responses for training interventions of different durations.

It is important to emphasize that the generalizability of our findings is currently unclear as our results are specific to our dataset. Future work is needed to confirm that monoexponential regressions best characterize CRF adaptations to exercise training in other data sets across a range of training types, durations and within different populations. Further, the trial used in the present study (Ross et al. 2013) adjusted exercise intensity after every CRF measurement (i.e., at 4, 8, and 16 weeks of training) and our findings may not be generalizable to studies that progress/adjust exercise intensity or amount differently throughout an exercise training intervention. It is possible that the specific protocol to adjust exercise intensity used in the trial by Ross and colleagues (2013) contributed to the monoexponential shape of changes in CRF to exercise training. Generalizability aside, our findings suggest that future studies should determine the type of regression that best fits the repeated measures data prior to classifying individual responses. This recommendation is supported by our finding that monoexponential regressions resulted in better fits than the linear regressions for all three exercise groups. Interestingly, this finding may suggest that the time course of CRF adaptations to exercise training follows a monoexponential shape regardless of exercise intensity and/or amount; however, future work is needed to confirm this contention.

Our findings may not be generalizable to exercise training interventions using different outcomes and/or intervention durations. For instance, certain outcomes including body weight (Sopko et al. 1985; Keim et al. 1990), fat mass (Kirk et al. 2003), and markers of cardiac function (Arbab-Zadeh et al. 2014) appear to adapt linearly or remain unchanged in the first several months of exercise training, which suggests that monoexponential regressions are not the best fit for these outcomes. Additionally, previous studies have indicated that linear models best describe CRF adaptations to short-term (i.e., 10 weeks) training interventions (Morris et al. 2002; Gass et al. 2004). Therefore, future work using different outcomes and/or intervention durations should determine the model that best fits their data prior to classifying individual responses. Our analysis demonstrates that choosing a model that best fits a given dataset can decrease the error in individual response estimates and consequently increase the confidence with which individuals are classified.

The importance of specificity of the modeling/fitting approach is also apparent when comparing the error and shape of the monoexponential curves for the mean CRF response across the three exercise groups (Fig. 1). For instance, the larger K value (reciprocal of tau) in LALI is consistent with a delayed plateau in the mean CRF response compared to HALI and HAHI. These between-group differences resulted in different mean CRF response estimates whereby LALI had the smallest and HAHI had the largest estimated mean CRF response (Fig. 1). Importantly, despite differences in the characteristics of the monoexponential curves between groups, the finding that a monoexponential regression was a better fit than a linear regression was consistent across all three groups.

It is important to note that our approach for deriving individual response estimates assumed that the time course of CRF adaptation followed the same monoexponential curve (i.e., contained the same K value; supplemental Fig. 1) across all participants within a given group. Despite completing the same exercise training program, it is likely that individuals experienced different rates of CRF adaptations (e.g. individuals within a group may reach a 'plateau' in CRF improvements at different timepoints). Therefore, although our approach to constrain the K value eliminated physiologically impossible CRF response estimates, this approach also prevented potential individuality in the shape of monoexponential regressions. The potential individuality in the rate/shape of CRF adaptations to training supports the need for future studies to characterize the shape of group/mean responses to determine the most appropriate approach to derive individual response estimates.

Although the current study involved a single CRF measure at each time point, the error in response estimates can be reduced by performing repeated measures at each time point (Hopkins 2004; Hecksteden et al. 2015). Therefore, in an attempt to improve the accuracy in individual response classification, future studies would benefit from including repeated measures at each time point.

Conclusion

Our findings demonstrate that individual monoexponential regressions were a superior fit compared to linear regressions for CRF responses to prolonged exercise training. Accordingly, using monoexponential regressions resulted in a larger number of individuals who could be confidently classified. Importantly, the similarity in response estimates and classification when using observed pre–post-changes and the typical error of measurement questions the need to include repeated measures when analyzing individual CRF responses to exercise training. Nevertheless, future work that wishes to use repeated measures should determine the response classification approach that provides the lowest error/greatest certainty in individual response estimates prior to classifying individual responses.

Author contributions Conceptualization: JB, RR, BG. Data curation: JB, RR, BG. Formal analysis: JB, RR, BG. Funding acquisition: RR. Methodology: JB, RR, BG. Writing—original draft: JB, BG. Writing—review and editing: JB, RR, BG.

Funding This work was supported by the Canadian Institutes of Health Research [Grant OHN-63277; http://www.cihr-irsc.gc.ca]. RR received this funding. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Research involving human participants All procedures performed in studies involving human participants were in accordance with ethical standards of the Health Sciences Human Research Ethics Board at Queen's University, verbal and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study. Verbal and written explanation of the experimental protocol and associated risks were provided to all participants prior to obtaining written informed consent.

References

Arbab-Zadeh A, Perhonen M, Howden E et al (2014) Cardiac remodeling in response to 1 year of intensive endurance training. Circulation 130:2152–2161. https://doi.org/10.1161/CIRCULATIONAHA.114.010775

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

Astorino TA, Schubert MM, Palumbo E et al (2013) Magnitude and time course of changes in maximal oxygen uptake in response to distinct regimens of chronic interval training in sedentary women. Eur J Appl Physiol 113:2361–2369. https://doi.org/10.1007/s00421-013-2672-1

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

Bonafiglia JT, Rotundo MP, Whittall JP et al (2016) Inter-individual variability in the adaptive responses to endurance and sprint interval training: a randomized crossover study. PLoS One 11:e0167790. https://doi.org/10.1371/journal.pone.0167790

Bonafiglia JT, Nelms MW, Preobrazenski N et al (2018) Moving beyond threshold-based dichotomous classification to improve the accuracy in classifying non-responders. Physiol Rep 6:e13928. https://doi.org/10.3390/rs6109653

Bouchard C, An P, Rice T et al (1999) Familial aggregation of VO2 max response to exercise training: results from the HERITAGE Family Study. J Appl Physiol 87:1003–1008

Bouchard C, Blair SN, Church TS et al (2012) Adverse metabolic response to regular exercise: Is it a rare or common occurrence? PLoS One 7:e37887. https://doi.org/10.1371/journal.pone.0037887

Buford TW, Roberts MD, Church TS (2013) Toward exercise as personalized medicine. Sport Med 43:157–165. https://doi.org/10.1007/s40279-013-0018-0

Gass G, Gass E, Wicks J et al (2004) Rate and amplitude of adaptation to two intensities of exercise in men aged 65–75 year. Med Sci Sports Exerc 36:1811–1818. https://doi.org/10.1249/01.MSS.0000142405.51319.FB

Govindasamy D, Paterson DH, Poulin MJ, Cunningham D (1992) Cardiorespiratory adaptation with short term training in older men. Eur J Appl Physiol Occup Physiol 65:203–208

Gurd BJ, Giles MD, Bonafiglia JT et al (2016) Incidence of nonresponse and individual patterns of response following sprint interval training. Appl Physiol Nutr Metab 41:229–234. https://doi.org/10.1139/apnm-2015-0449

Hecksteden A, Kraushaar J, Scharhag-Rosenberger F et al (2015) Individual response to exercise training: a statistical perspective. J Appl Physiol 118:1450–1459. https://doi.org/10.1152/japplphysiol.00714.2014

Hecksteden A, Pitsch W, Rosenberger F, Meyer T (2018) Repeated testing for the assessment of individual response to exercise training. J Appl Physiol 124:1567–1579. https://doi.org/10.1152/japplphysiol.00896.2017

Hickson RC, Hagberg JM, Ehsani AA, Holloszy JO (1981) Time course of the adaptive responses of aerobic power and heart rate to training. Med Sci Sport Exerc 13:17–20

Hopkins WG (2000a) Measures of reliability in sports medicine and science. Sports Med 30:1–15. https://doi.org/10.2165/00007256-200030050-00006

Hopkins WG (2000b) Precision of the estimate of a subject's true value (Excel spreadsheet). In: A new view of statistics. Internet society for sport science. http://www.sportsci.org/resource/stats/xprecisionsubject.xls

Hopkins WG (2004) How to interpret changes in an athletic performance test. Sportscience 8:1–7. https://doi.org/10.1097/00005768-199804000-00026

Keim NL, Barbieri TF, Van Loan MD, Anderson BL (1990) Energy expenditure and physical performance in overweight women: response to training with and without caloric restriction. Metabolism 39:651–658. https://doi.org/10.1016/0026-0495(90)90035-B

Kirk EP, Jacobsen DJ, Gibson C et al (2003) Time course for changes in aerobic capacity and body composition in overweight men and women in response to long-term exercise: the Midwest Exercise Trial (MET). Int J Obes 27:912–919. https://doi.org/10.1038/sj.ijo.0802317

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

Morris N, Gass G, Thompson M et al (2002) Rate and amplitude of adaptation to intermittent and continuous exercise in older men. Med Sci Sports Exerc 34:471–477. https://doi.org/10.1097/00005768-200203000-00014

Murias JM, Kowalchuk JM, Paterson DH (2010) Time course and mechanisms of adaptations in cardiorespiratory fitness with endurance training in older and young men. J Appl Physiol 108:621–627. https://doi.org/10.1152/japplphysiol.01152.2009

Raleigh JP, Giles MD, Scribbans TD et al (2016) The impact of work-mathced interval training on VO_{2peak} and VO_2 kinetics: diminishing returns with increasing intensity. Appl Physiol Nutr Metab 41:706–713. https://doi.org/10.1139/apnm-2015-0614

Raleigh JP, Giles MD, Islam H et al (2018) Contribution of central and peripheral adaptations to changes in VO_{2max} following four weeks of sprint interval training. Appl Physiol Nutr Metab 43:1059–1068. https://doi.org/10.1139/apnm-2017-0864

Ross R, Hudson R, Day AG, Lam M (2013) Dose-response effects of exercise on abdominal obesity and risk factors for cardiovascular disease in adults: Study rationale, design and methods. Contemp Clin Trials 34:155–160. https://doi.org/10.1016/j.cct.2012.10.010

Ross R, De Lannoy L, Stotz PJ (2015a) Separate effects of intensity and amount of exercise on interindividual cardiorespiratory fitness response. Mayo Clin Proc 90:1506–1514. https://doi.org/10.1016/j.mayocp.2015.07.024

Ross R, Hudson R, Stotz PJ, Lam M (2015b) Effects of exercise amount and intensity on abdominal obesity and glucose tolerance in obese adults. Ann Intern Med 162:325. https://doi.org/10.7326/M14-1189

Ross R, Blair SN, Arena R et al (2016) Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the american heart association. Circulation 134:e653–e699. https://doi.org/10.1161/CIR.0000000000000461

Scharhag-Rosenberger F, Meyer T, Walitzek S, Kindermann W (2009) Time course of changes in endurance capacity: a 1-yr training study. Med Sci Sports Exerc 41:1130–1137. https://doi.org/10.1249/MSS.0b013e3181935a11

Scharhag-Rosenberger F, Walitzek S, Kindermann W, Meyer T (2012) Differences in adaptations to 1 year of aerobic endurance training: Individual patterns of nonresponse. Scand J Med Sci Sport 22:113–118. https://doi.org/10.1111/j.1600-0838.2010.01139.x

Senn S, Rolfe K, Julious SA (2010) Investigating variability in patient response to treatment—a case study from a replicate cross-over study. Stat Methods Med Res 0:1–11. https://doi.org/10.1177/0962280210379174

Sisson SB, Katzmarzyk PT, Earnest CP et al (2009) Volume of exercise and fitness nonresponse in sedentary, postmenopausal women. Med Sci Sports Exerc 41:539–545. https://doi.org/10.1249/MSS.0b013e3181896c4e

Sopko G, Leon a S, Jacobs DR et al (1985) The effects of exercise and weight loss on plasma lipids in young obese men. Metabolism 34:227–236

Swinton PA, Hemingway BS, Saunders B et al (2018) A Statistical framework to interpret individual response to intervention: paving the way for personalised nutrition and exercise prescription. Front Nutr 5:41. https://doi.org/10.3389/FNUT.2018.00041

Vollaard NBJ, Constantin-Teodosiu D, Fredriksson K et al (2009) Systematic analysis of adaptations in aerobic capacity and submaximal energy metabolism provides a unique insight into determinants of human aerobic performance. J Appl Physiol 106:1479–1486. https://doi.org/10.1152/japplphysiol.91453.2008

Williamson PJ, Atkinson G, Batterham AM (2017) Inter-individual responses of maximal oxygen uptake to exercise training: a critical review. Sport Med 47:1501–1513. https://doi.org/10.1007/s40279-017-0680-8

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.