Wearable Sensor Technology and Machine Learning for Prediction of Oxygen Uptake and Extraction of Kinetics during Moderate and Heavy Intensity Exercise

by

Eric Thomas Hedge

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Kinesiology

Waterloo, Ontario, Canada 2020

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Chapter 1: Eric T. Hedge drafted the chapter; Eric T. Hedge, Jason S. Au, and Richard L. Hughson edited and revised the chapter; Eric T. Hedge, Jason S. Au, and Richard L. Hughson approved the final version of the chapter.

Chapter 2: Eric T. Hedge and Richard L. Hughson conceived and designed the research; Eric T. Hedge performed experiments; Eric T. Hedge and Richard L. Hughson analyzed the data; Eric T. Hedge and Richard L. Hughson interpreted results; Eric T. Hedge prepared figures; Eric T. Hedge drafted the chapter; Eric T. Hedge and Richard L. Hughson edited and revised the chapter; Eric T. Hedge and Richard L. Hughson approved the final version of the chapter.

Chapter 3: Eric T. Hedge, Robert Amelard and Richard L. Hughson conceived and designed the research; Eric T. Hedge performed experiments; Eric T. Hedge, Robert Amelard and Richard L. Hughson analyzed the data; Eric T. Hedge, Robert Amelard and Richard L. Hughson interpreted results; Eric T. Hedge prepared figures; Eric T. Hedge drafted the chapter; Eric T. Hedge, Robert Amelard and Richard L. Hughson edited and revised the chapter; Eric T. Hedge, Robert Amelard and Richard L. Hughson approved the final version of the chapter.

Chapter 4: Eric T. Hedge drafted the chapter; Eric T. Hedge and Richard L. Hughson edited and revised the chapter; Eric T. Hedge and Richard L. Hughson approved the final version of the chapter.

Abstract

The speed at which pulmonary oxygen uptake ($\dot{V}O_{2p}$) responds to a change in energetic demand provides information about an individual's health and functional capabilities. Advances in wearable sensor technology and machine learning coupled with a novel frequency domain analysis technique (mean normalized gain, MNG) have enabled $\dot{V}O_{2p}$ kinetics to be assessed outside the confines of the laboratory setting. However, the validated $\dot{V}O_{2p}$ prediction algorithm and MNG were only used to predict and extract an indicator of $\dot{V}O_{2p}$ kinetics from low to moderate intensity exercise responses. Therefore, the purpose of this thesis was to investigate if a model based on the previously validated moderate intensity $\dot{V}O_{2p}$ prediction algorithm, and MNG, are appropriate for heavy intensity exercise, where exercise response dynamics are more complex. In order to accomplish this, two studies were conducted. The first study used moderate and heavy pseudorandom binary sequence (PRBS) and constant work rate exercise responses to evaluate the validity of MNG as an indicator of kinetics during heavy intensity exercise. Differences in MNG were detected between moderate and heavy exercise, and strong negative within- and between-participant correlations were observed between MNG and the conventional time domain indicator of kinetics, τ. This indicated that MNG is an accurate indicator of $\dot{V}O_{2p}$ kinetics during heavy exercise. The second study used both moderate and heavy exercise responses to develop a new machine learning algorithm, based on the previously validated moderate exercise algorithm, to predict $\dot{V}O_{2p}$ from easy-to-obtain wearable sensor data. The new model differed from the previously validated algorithm as a result of including the more complex heavy intensity exercise responses in the training dataset. The new model's ability to accurately predict dynamic $\dot{V}O_{2p}$ responses was evaluated during different intensity PRBS exercise. The prediction algorithm was accurate at estimating absolute $\dot{V}O_{2p}$ based on repeated measures correlation and repeated

measures Bland-Altman analysis of the predicted versus measured $\dot{V}O_{2p}$ when all exercise conditions' data were combined. However, the algorithm was biased to predicting slower $\dot{V}O_{2p}$ kinetics, as determined by the MNG method, and only a moderate positive within-participant correlation was observed between the predicted and measured kinetics. This suggests that the previously validated model architecture for $\dot{V}O_{2p}$ prediction during moderate intensity exercise is not easily generalizable to include heavy intensity exercise. Therefore, new machine learning models should be explored in order to improve the prediction of heavy dynamic $\dot{V}O_{2p}$ responses, and to unlock the full potential of frequency domain analysis as a tool for field evaluation of $\dot{V}O_{2p}$ kinetics across a wide range of intensities.

Acknowledgements

I would like to thank all of the people who volunteered to participate in the exercise studies that made this work possible. I truly appreciate each person's willingness to complete the repeated bouts of intermittent and sustained high intensity exercise.

To my mentor, Richard Hughson, thank-you for giving me the opportunity to succeed and grow as a student and researcher. Thank-you for providing the freedom and support which enabled me to explore so many different topics over the course of my master's degree. The knowledge, experiences, and advice that you have shared with me have been invaluable.

To my committee, Paolo Dominelli and Alexander Wong, thank-you for contributing your expertise to this thesis.

To my lab-mates, thank you for always being available to help and support me during my master's program. Thank-you to Kathryn Zuj for showing me the ropes when I first arrived in the lab, and Danielle Greaves for helping me though all the technical difficulties I had experienced. Thank-you to Jason Au, Robert Amelard, and Andrew Robertson for their patience and sense of humor when providing tips on coding and data analysis. A special thank-you to Sean Peterson and Cederick Landry, for giving me the opportunity to learn and work on their engineering projects.

To Kevin Murray and Courtney Patterson, thank-you for all the support and adventures during our master's degrees. Graduate school certainly would not have been as much fun without you.

Lastly, to my family, thank-you for providing me with unwavering support and encouragement throughout this journey.

Table of Contents

List of Figures

List of Tables

List of Equations

List of Abbreviations

$\mathbf{a_{DC}}$	system DC offset amplitude
A	pulmonary oxygen uptake amplitude for a given phase
$\mathbf{A_h}$	cosine amplitude for a given harmonic
$\mathbf{AMP_h}$	total amplitude at a given harmonic
ANOVA	analysis of variance
$\mathbf{B_h}$	sine amplitude for a given harmonic
ECG	electrocardiogram
$\boldsymbol{f_1}$	fundamental frequency
FFT	fast Fourier transformation
$\mathbf{gAMP_h}$	system gain at a given harmonic
LT	lactate threshold
MNG	mean normalized gain
MRT	mean response time
$\mathbf{PO_2}$	partial pressure of oxygen
PRBS	pseudorandom binary sequence
PRTS	pseudorandom ternary sequence
RMSE	root mean square error
$\boldsymbol{r}$	Pearson correlation coefficient
$\boldsymbol{r_{rm}}$	repeated measures correlation coefficient
t	time
TD	time delay for a given phase
$\mathbf{\dot{V}_E}$	minute ventilation
$\mathbf{\dot{V}O_{2p}}$	pulmonary oxygen uptake
$\mathbf{\dot{V}O_{2p}AMP_h}$	pulmonary oxygen uptake amplitude for a given harmonic
$\mathbf{\dot{V}O_{2pBSL}}$	baseline pulmonary oxygen uptake
$\mathbf{\dot{V}O_{2pMax}}$	maximal pulmonary oxygen uptake
VT	ventilatory threshold
WR AMP$_h$	work rate amplitude for a given harmonic
WR	work rate
Δ50%	work rate that corresponds with 50% of the difference between ventilatory threshold and peak pulmonary oxygen uptake
τ	time constant for a given phase
ϕ	phase shift
% HRR	percent heart rate reserve

Chapter 1: General Introductions

Recent technological advancements in wearable sensors and artificial intelligence have enabled the rate of pulmonary oxygen uptake ($\dot{V}O_{2p}$) to be estimated and monitored continuously throughout the day outside of a laboratory environment (14). Coupled with a novel $\dot{V}O_{2p}$ dynamic analysis technique, this advancement has unlocked the possibility of tracking changes in people's fitness over time. This technology could have many applications in the healthcare sector, as it could potentially aid in the diagnosis of disease, promote maintenance of healthy lifestyles, or be used to track a patient's rehabilitation progress. Outside of healthcare, it could be widely used by the general population, and even athletes, who are interested in monitoring their own fitness and activity over time. However, current $\dot{V}O_{2p}$ estimation and analysis methods for continuous monitoring have not been evaluated for their ability to predict and extract dynamics during activities that are beyond low to moderate intensity exercise, where the physiological responses are much more complex. Therefore, this chapter: i) provides an overview of $\dot{V}O_{2p}$ kinetics and how they differ between different intensities of exercise, ii) outlines the common time domain and frequency domain (mean normalized gain; MNG) analysis techniques to extract $\dot{V}O_{2p}$ kinetics, and iii) discusses the application of machine learning algorithms to predict $\dot{V}O_{2p}$. At the end of this chapter, the rationale, research questions and hypotheses for the thesis are stated.

1.1 Overview of Oxygen Uptake Responses to Exercise

When exercising, there is an increased energy demand by the active muscles. This heightened energy requirement is met by an increased rate of oxidative phosphorylation leading to greater oxygen consumption by mitochondria, as oxygen is the final electron acceptor in the

electron transport chain. This increased rate of oxygen usage in the mitochondria of the active muscles is reflected by a higher $\dot{V}O_{2p}$ measured at the mouth, and is governed by the Fick equation, where $\dot{V}O_{2p}$ is the product of systemic oxygen delivery (cardiac output) and total oxygen extraction (arterial-mixed venous oxygen difference) (39). However, oxygen delivery and extraction, as well as the rate of aerobic metabolism, do not instantaneously adjust to fully accommodate a change in energy demand. Therefore, the rate at which $\dot{V}O_{2p}$ adjusts to a change in energy requirement for a given change in work rate (WR), not just the change in $\dot{V}O_{2p}$, is important. This led to the study of oxygen uptake kinetics, which is the quantification of how quickly oxidative metabolism adjusts to a change in energy demand, and the oxygen transport and utilization systems that control the dynamics of the response. Having fast kinetics is crucial, as it reduces the anaerobic energy contribution at the onset of exercise, and it is linked to better aerobic fitness (29, 41, 43, 69). Clinically, faster $\dot{V}O_{2p}$ kinetics are associated with better functional mobility in frail older adults (1), and better prognosis for patients with chronic obstructive pulmonary disease (20) and heart failure (25, 72).

The $\dot{V}O_{2p}$ response to a change in exercise intensity has been predominantly analyzed in the time domain. Conceptually, $\dot{V}O_{2p}$ responses in the time domain are easiest to discuss. Exercise protocols consisting of a single or multiple instantaneous square wave changes in WR to a new constant level facilitate time domain analysis. The on-transient $\dot{V}O_{2p}$ profile for such a transition can be characterized by exponential modeling (refer to Chapter 1.3.1), and this profile can be subdivided into three distinct phases (48, 84) (Fig. 1). Following an abrupt increase in WR during leg exercise, $\dot{V}O_{2p}$ exhibits the brief cardio-dynamic phase (phase I), which reflects an increase in cardiac output that increases pulmonary perfusion (83). The greater pulmonary blood flow allows for increased diffusion of oxygen from the alveoli into the capillaries to increase the measured

$\dot{V}O_{2p}$, but it does not actually reflect an increased oxygen extraction from the blood at the level of the working muscles due to a transit delay from the active muscles to the lungs (83). Despite this transit delay, the oxygen content of the mixed venous blood entering the pulmonary circulation does decrease without a discernable delay at the onset of exercise (27). The reduced oxygen content observed during phase I is due to changes in the in the relative contribution of different vascular beds to venous return (27). The reduced oxygen content of the mixed venous blood results in a larger arterial-mixed venous oxygen difference, which contributes to the abrupt increase in $\dot{V}O_{2p}$ observed during phase I. Next is the primary component (phase II), where $\dot{V}O_{2p}$ appears to exponentially increase towards a new steady-state (84). Last is phase III of the response, where i) $\dot{V}O_{2p}$ achieves a new steady-state, ii) the $\dot{V}O_{2p}$ slow component develops to delay the attainment of steady-state conditions, or iii) the $\dot{V}O_{2p}$ slow component continues to increase until a maximal level ($\dot{V}O_{2pMax}$) (48, 84). The $\dot{V}O_{2p}$ slow component is a delayed onset, secondary increase in $\dot{V}O_{2p}$, and it reflects a fatigue-induced reduction in the efficiency of muscle contractions (90), and increased respiratory muscle work (31).

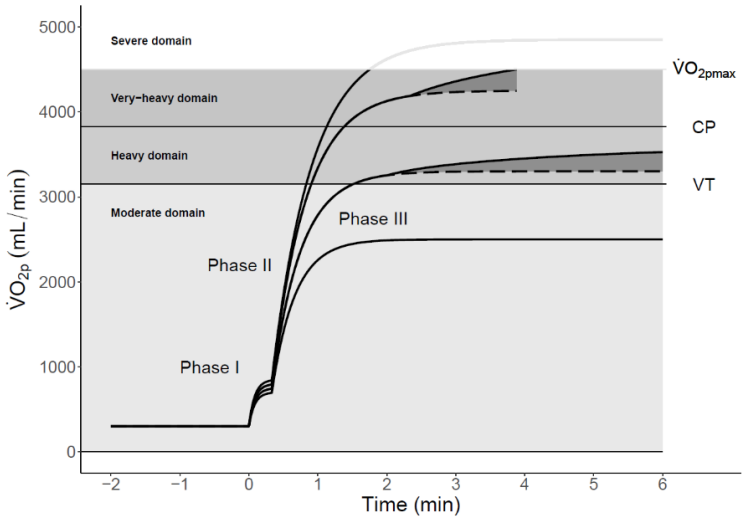

Fig. 1. A representation of the three different phases of the pulmonary oxygen uptake ($\dot{V}O_{2p}$) response to exercise across all intensity domains. $\dot{V}O_{2p}$ response curve characteristics differ depending on the intensity of the exercise being performed. The shaded area between the solid and dashed lines indicate the $\dot{V}O_{2p}$ slow component amplitude. During very-heavy and severe intensity exercise, $\dot{V}O_{2p}$ increases to the maximal level if the work rate is sustained for a long enough duration. Critical power is CP; maximal oxygen uptake is $\dot{V}O_{2pMax}$; ventilatory threshold is VT. Modified from Hughson (48).

The intensity at which a person is exercising dictates the temporal characteristics of these phases of the $\dot{V}O_{2p}$ response. Based on different metabolic thresholds, there are four distinct exercise domains, named moderate, heavy, very-heavy, and severe, that all result in unique dynamic responses (84). These different exercise intensity domains also match closely with peoples' perceived level of effort (73). The lowest intensities are found within the moderate intensity domain, where steady-state $\dot{V}O_{2p}$ is quickly attained (85). The moderate domain contains all exercise intensities below lactate threshold (LT), which is the lowest constant-WR associated with net blood lactate accumulation (84), and this threshold can be non-invasively estimated as the ventilatory threshold (VT) (10). The range of WRs that are above LT but below critical power (CP; the highest constant-WR that can be maintained "indefinitely," without the progressive accumulation of fatigue-inducing metabolites and the depletion of muscle high energy phosphates

(54)) is defined as the heavy domain (84). During heavy domain exercise, the $\dot{V}O_{2p}$ slow component slowly develops resulting in measured $\dot{V}O_{2p}$ being higher than what would be anticipated from extrapolation of the relatively linear moderate intensity $\dot{V}O_{2p}$-WR relationship (70, 91). The range of WRs that are above CP, but do not have a primary phase $\dot{V}O_{2p}$ requirement above $\dot{V}O_{2pMax}$, are defined as the very-heavy intensity domain (84). When performing very-heavy exercise, $\dot{V}O_{2p}$, arterial blood lactate concentration, and arterial blood acidity progressively increase throughout exercise, despite no change in WR, until the limit of tolerance, when $\dot{V}O_{2pMax}$ is achieved and fatigue ensues (48, 84). Lastly, there is severe intensity exercise, in which the $\dot{V}O_{2p}$ requirement is above $\dot{V}O_{2pMax}$ (48, 84). In this domain of exercise, the primary phase of the response projects to a level that exceeds $\dot{V}O_{2pMax}$, and no $\dot{V}O_{2p}$ slow component is evident (84).

1.2 Dynamic Non-Linearity of Oxygen Uptake Responses

The idea of a dynamically linear $\dot{V}O_{2p}$ response to exercise, where the time course of the change and gain ($\Delta\dot{V}O_{2p}/\Delta WR$) are invariant for all changes in WR, has been repeatedly challenged. Overall, there has been an accumulation of indisputable evidence that suggests that $\dot{V}O_{2p}$ is not under the control of a simple linear system (49, 56, 60, 85). However, within the moderate intensity domain, there has been debate over whether the system is, or is not, still considered dynamically linear. In support of a linear system were the findings of Paterson and Whipp (68), and Özyener et al. (67), as they found that below VT, the on- and off-transient kinetics are fairly symmetrical. Despite this on-off symmetry, the principle of superposition is still violated within the moderate domain: from increasing baseline work and metabolic rates, the kinetics of the response become progressively slower (21, 24, 49, 58). Then, once exercising at WRs above

VT, the kinetics become much more complex, which provides further evidence to refute the idea of dynamic linearity.

In exercise that is above VT, attainment of constant $\dot{V}O_{2p}$ conditions are either delayed, or never achieved, as a result of the development of the $\dot{V}O_{2p}$ slow component. This excess oxygen consumption is unable to be predicted by a dynamically linear system (70). Transitions to heavy exercise have been observed to elicit a longer τ than transitions to moderate exercise from the same baseline condition (36, 51, 53, 61, 65, 68), although others have reported no difference (7, 8, 67, 88). Furthermore, phase II kinetics are even slower when initiating heavy exercise transitions from elevated baseline WRs (32, 88). Findings by Özyener et al. (67) directly contrast the idea of slowed kinetics during higher intensity exercise transitions, as they observed similar phase II time constants between moderate, heavy, very-heavy and severe intensity cycling when initiated from the same baseline metabolic and work rates. However, it is important to recognize the limitations of conventional kinetic analysis by exponential curve fitting for exercise intensities above VT, as exponential models assume that the asymptote represents the true end point. Naïve interpretation of phase II τ and gain during exercise above VT may actually suggest faster kinetics and improved exercise efficiency during high compared to moderate intensity exercise, even though that is not actually the case (50) (for further discussion see Chapter 1.3.1). That being said, non-conventional assessment of kinetics has confirmed that $\dot{V}O_{2p}$ does respond with slower kinetics above compared to below VT, as Haouzi et al. (42) observed a larger phase shift and reduced amplitude for $\dot{V}O_{2p}$ during sinusoidal exercise above VT despite oscillations in WR being of the same amplitude as the below VT exercise. Further evidence against a dynamically linear system is demonstrated by its inability to explain the on- and off-transient dynamic asymmetry that is found in above VT exercise (67, 68).

One question that was not resolved until recently, was if the system is nonlinear, then how does it produce a ramp-incremental exercise $\dot{V}O_{2p}$ response that is well described by an exponential function with a single time constant and gain parameter, which suggests dynamic linearity (76, 82)? This apparent dichotomy was clarified by both the works of Keir et al. (56) and Wilcox et al. (87), as they both found with their extended step-incremental exercise protocols that as the pre-transition metabolic rate increases, the kinetics of the $\dot{V}O_{2p}$ response slow down as the gain concomitantly increases. This co-inflation of the time constant and gain over the course of a ramp-incremental exercise test, where WR and metabolic rate are continuously increasing, allows for this pseudo-linear $\dot{V}O_{2p}$ response to occur (56, 87).

The current state of the literature recognizes that $\dot{V}O_{2p}$ responds with non-linear kinetics across exercise intensities; however, within the moderate intensity domain, it is often simplified to a linear response in order to facilitate exercise prescription. Although much is known about $\dot{V}O_{2p}$ response dynamic characteristics of the different exercise domains, the mechanisms underlying slower $\dot{V}O_{2p}$ kinetics with increasing exercise intensity, and each mechanism's relative contribution to the response, are still debated, with results implicating slower adaptation of muscle oxygen delivery (64), hierarchical recruitment of muscle fibers in which aerobic metabolism adjusts more slowly (24), and muscle metabolic status (21). Interestingly, Keir et al. (60) proposed a unifying theory in which the slowing of $\dot{V}O_{2p}$ kinetics is a result of the complex interaction between heterogeneous metabolism and blood flow kinetics within the hierarchy of recruited muscle fibers with higher intensity exercise.

1.3 Analysis Techniques to Extract Oxygen Uptake Kinetics

There are different methods to examine and quantify $\dot{V}O_{2p}$ kinetics; however, the predominant method is time domain analysis. This method is ideal for controlled laboratory setups

in which exercise WRs and duration are able to be tightly controlled. However, recent technological advancements in machine learning and wearable sensors have enabled assessment of $\dot{V}O_{2p}$ outside of the laboratory, in uncontrolled and unsupervised settings (14). In these uncontrolled instances, the $\dot{V}O_{2p}$ data are not well suited for time domain analysis, as WR and $\dot{V}O_{2p}$ are constantly fluctuating. Instead, these data are better suited for less conventional frequency domain analysis to quantify $\dot{V}O_{2p}$ kinetics.

1.3.1 Time Domain Kinetic Analysis

Time domain analysis of the $\dot{V}O_{2p}$ response to an instantaneous step change in WR to a new constant level has been the conventional method used to investigate $\dot{V}O_{2p}$ dynamics. Over the years, there have been different ways that $\dot{V}O_{2p}$ profiles have been fit by exponential models to extract the on-transient kinetics of the response. Nowadays, there are a few common ways to fit $\dot{V}O_{2p}$ data in the time domain. One method fits a mono-exponential function to the entire $\dot{V}O_{2p}$ response (56):

$$\dot{V}O_{2p}(t) = \dot{V}O_{2pBSL} + A \cdot \left(1 - e^{-\left(t/MRT\right)}\right) \quad (1)$$

, where $\dot{V}O_{2p}(t)$ is the value of $\dot{V}O_{2p}$ at any time, *t*, during the transition, $\dot{V}O_{2pBSL}$ is the pre-transition baseline value, *A* is the amplitude from the baseline value, and MRT is the mean response time. This method ignores the discrepancies in the curve shape between the three phases, which results in residual bias, especially when a slow component is present (33). This method has been used for all intensity domains. The other method fits a mono-exponential function to phase II of the response (49, 83):

$$\dot{V}O_{2p}(t) = \dot{V}O_{2pBSL} + A_p \cdot \left(1 - e^{-\left(\left(t-TD_p\right)/\tau_p\right)}\right) \quad (2)$$

, where in addition to previously defined variables, A_p is the phase II amplitude from the baseline value, τ_p is the phase II time constant, and TD_p is the phase II time delay. In the case of heavy and very-heavy exercise, an additional mono-exponential function can be added to equation 2, such that it fits both the primary and slow components of the response (8):

$$\dot{V}O_{2p}(t) = \dot{V}O_{2pBSL} + A_p \cdot \left(1 - e^{-\left(t-TD_p/\tau_p\right)}\right) + A_s \cdot \left(1 - e^{-\left(t-TD_s/\tau_s\right)}\right) \quad (3)$$

, where the previously defined variables above with the subscripts $_p$ and $_s$ represent the primary and slow components of the response, respectively. In equations 2 and 3, it is important to note that the cardio-dynamic phase data have been excluded from the exponential fits. This can be accomplished simply by just excluding the first 20 s of the transition from the exponential fit (15, 16, 66, 67), or more complexly, by starting with the phase II fitting window approximately 30 s into the transition and then progressively moving the fitting window backwards while examining the flatness of the residual profile, and indicators of the quality of fit, such as the 95% confidence interval, to determine the end of the cardio-dynamic phase (56, 57). For multicomponent models that do include an exponential fit of the cardio-dynamic phase, in addition to the primary and slow components, MacDonald et al. (63) developed an alternative method to calculate MRT as the weighted sum of the time delay and time constant for each component:

$$MRT = \left(\frac{A_c}{A_c+A_p+A_s}\right) \cdot (TD_c + \tau_c) + \left(\frac{A_p}{A_c+A_p+A_s}\right) \cdot (TD_p + \tau_p) + \left(\frac{A_s}{A_c+A_p+A_s}\right) \cdot (TD_s + \tau_s) \quad (4)$$

, where the subscripts *c*, *p*, and *s* for the A, TD, and τ represent the cardio-dynamic, primary and slow component phases, respectively. Overall, equations 1, 2, and 3 are all used to assess system dynamics and gain. The MRT and τ are very similar, as they both describe curvature of the response. Mathematically, they both describe the amount of time that it would take to achieve 63% of the change in $\dot{V}O_{2p}$ for the whole curve fit with MRT, or a specific phase fit of the response with τ_p and τ_s. That means after four MRT or τ, greater than 98% of the change in apparent $\dot{V}O_{2p}$

for the specific phase would have occurred. Accordingly, the faster the kinetics, the shorter the MRT or τ, which means the required $\dot{V}O_{2p}$ is achieved more rapidly. The total gain of the whole response fit, and the primary gain of the phase II fit are easily determined by simply dividing the $\dot{V}O_{2p}$ amplitude by the change in WR, and this is also the inverse of delta efficiency. Therefore, these equations allow for quantification of the rate at which the system adapts to an exercise stimulus, while also determining the magnitude of the response with respect to the stimulus magnitude.

As outlined by Hughson (48), these exponential curve fitting models do have uncertainty associated with their parameters. Some of the uncertainty is a result of how sensitively an exponential curve can fit the $\dot{V}O_{2p}$ response with normal breath by breath variability (62). Hughson (48) provided a moderate intensity *in silico* example that demonstrates this limitation. His hypothetical $\dot{V}O_{2p}$ data set was created using an exponential equation with an initial fast and a later slow kinetic portion of the curve, and ±75mL·min^{-1} of random noise was applied to the signal to simulate breath by breath variability, which is well within the range of variability seen in real life data. In the data set, there was no cardio-dynamic phase, and it was only the first 30 s of data that had the faster kinetics (τ = 24 s), and the following 4.5 min of data had 50% slower kinetics (τ = 36 s). Interestingly, modeling these data with equation 2, the mono-exponential fit output was τ = 26.6 s. This τ is clearly much more sensitive to what occurred in the relatively brief time nearest the onset of the exercise transition compared to the rest of the time as $\dot{V}O_{2p}$ approached the new steady-state level. This is an important consideration, as in the early phase of the transient, there is high intracellular partial pressure of oxygen O_2 (PO_2), which should allow oxidative metabolism to increase at a rate that is faster than later on in the transient phase, where intracellular PO_2 is reduced (77). This shows that a single exponential model is unable to parse out some of the

metabolic changes that may affect the transient response. Additionally, in studies that use time domain analysis, multiple repetitions of testing protocols are required to be performed for each participant, the data cleaned to remove aberrant breaths (62), and then like-repetitions ensemble averaged, in order to try to improve the signal-to-noise ratio for a better fit of the data (57). However, even with improved signal-to-noise ratio, the curve fitting uncertainty still applies.

Other curve fitting issues arise in the heavy, very-heavy, and severe exercise domains. One problem with using equation 3 for modeling heavy and very-heavy exercise responses, is that when isolating the phase II and III components, the amplitude that the phase II τ projects toward is less than the predicted $\dot{V}O_{2p}$ for the given WR, while the phase III τ projects to an amplitude above the predicted $\dot{V}O_{2p}$ (48). This is particularly problematic for interpreting $\dot{V}O_{2p}$ dynamics, as during transitions to high intensities, the phase II τ can actually be shorter than the τ measured in the moderate intensity domain (50). Counterintuitively, the reduced phase II τ and gain that can be observed during high intensity exercise compared to moderate intensity exercise actually suggests improved exercise efficiency and faster $\dot{V}O_{2p}$ kinetics; however, this is not the case. Hughson et al. (50) clearly demonstrated that this apparent speeding of $\dot{V}O_{2p}$ kinetics at high intensities (~96 % and 125% of peak$\dot{V}O_{2p}$) is a result of the conventional exponential modeling technique, and in fact, when the data were analyzed using a semi-logarithmic transformation which compared the differences between the measured and required $\dot{V}O_{2p}$ for the given exercise, the high intensity exercise actually revealed slower kinetics. Furthermore, in the very-heavy domain, as the $\dot{V}O_{2p}$ projects towards $\dot{V}O_{2pMax}$, the slow component fit will actually project to an amplitude that exceeds $\dot{V}O_{2pMax}$ (48). There are similar issues in interpreting exponential curve fits of serve intensity exercise, as there is no discernable slow component to fit, and the mathematical amplitude that the fitted mono-exponential function projects to does not correspond with the metabolic demand (48).

Further complications in fitting an exponential function to the severe intensity breath-by-breath $\dot{V}O_{2p}$ response arise due to the short duration that the people can maintain the WR, and the limited number of breaths that occur over that time period. Thus, these issues create further uncertainty around how the model parameters should be interpreted for exercise above VT.

1.3.2 Frequency Domain Kinetic Analysis

Frequency domain analysis is a less commonly used method to investigate $\dot{V}O_{2p}$ kinetics. It has been regarded as a desirable analysis method because it does not try to fit the $\dot{V}O_{2p}$ data to a model with some degree of arbitrariness (37). Not using a specific model to fit the data helps to avoid the uncertainty problems that are common in time domain curve fitting (48), and it enables the analysis of aerobic response dynamics from non-constant WR exercise responses.

Over the years different types of exercise protocols have been used to facilitate frequency domain analysis, such as pseudorandom binary sequences (PRBS) (16, 17, 38, 46, 47, 52), pseudorandom ternary sequences (PRTS) (13–15), sinusoidal fluctuations in WR (28, 42), and even unsupervised and uncontrolled activities of daily living (14). In the pseudorandom protocols, the WR alternates between two (PRBS) or three (PRTS) levels with step-wise increases and decreases in WR, while in sinusoidal protocols, the WR smoothly oscillates between two levels. The use of pseudorandom sequence protocols has become the norm in terms of testing to collect data for frequency domain analysis, as they allow for testing of $\dot{V}O_{2p}$ dynamics across a wide range of frequencies and metabolic rates in a single testing session (37, 52). The input (WR) and output (measured $\dot{V}O_{2p}$) of these types of protocols are collected in the time domain and then must be converted into the frequency domain.

In order to convert the WR and $\dot{V}O_{2p}$ signals from the time domain into the frequency domain, data are transformed using the discrete Fourier transformation. To do this, the following sinusoidal function is solved for an integer number of harmonics, *H*:

$$y(t) = a_{DC} + 2 \cdot \sum_{h=1}^{H} \left(A_h \cdot cos(2\pi \cdot h \cdot f_1 \cdot t) + B_h \cdot sin(2\pi \cdot h \cdot f_1 \cdot t) \right) \quad (5)$$

where *y(t)* is the signal at a given point in the time domain, a_{DC} is the system DC offset amplitude (i.e. the average response), f_1 is the fundamental frequency (1/period of the sinusoid, or the duration of a single cycle of the PRBS or PRTS), *t* is the time, A_h and B_h are the cosine and sine amplitude coefficients for a given harmonic, *h* (an integer), respectively. From A_h and B_h, the total amplitude (AMP_h) and phase shift (ϕ) can be calculated for each harmonic:

$$AMP_h = \sqrt{A_h^2 + B_h^2} \quad (6)$$

$$\phi = tan^{-1}\left(\frac{B_h}{A_h}\right) \quad (7)$$

This conversion from time to frequency domain can also be achieved by using the fast Fourier transformation (FFT). This conversion also decomposes the time-series signals into a series of sinusoids with given amplitudes at frequencies that are multiples of the fundamental frequency (Fig. 2). These amplitudes ($\dot{V}O_{2p}AMP_h$ and WR AMP_h) are then used to generate amplitude spectrums for both the $\dot{V}O_{2p}$ and WR signals.

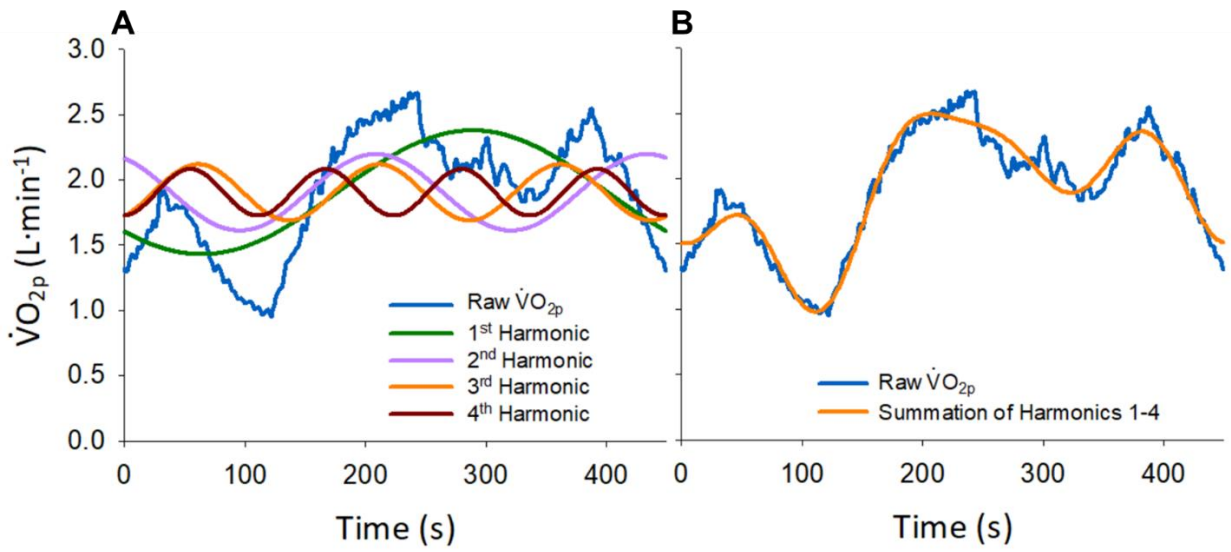

Fig. 2. Illustration of how a PRBS $\dot{V}O_{2p}$ signal in the time domain (blue line) can be decomposed into a series of sinusoids (A). For simplicity, only the first four harmonic frequencies are depicted. Summation of the first four harmonic frequencies displayed in A results in the reconstructed filtered $\dot{V}O_{2p}$ signal (orange line) in the time domain (B).

Once signals are in the frequency domain, they can be used for different types of analysis. One of the earliest, and continued applications of frequency domain analysis of $\dot{V}O_{2p}$ data is to examine the topic of dynamic linearity, and identify nonlinearities within the signal (17, 28, 37, 42, 47). Many studies have used Fourier analysis of $\dot{V}O_{2p}$, in conjunction with other physiological measures, in order to isolate different factors that distort the muscle oxygen uptake parameters that are assumed to be approximated by gas measurement at the mouth (17, 38, 45).

Recently, a new method to quantify $\dot{V}O_{2p}$ dynamics in the frequency domain was developed by Beltrame and Hughson (16), which uses data collected from pseudorandom sequence exercise protocols (13, 15, 16), and even unsupervised activities of daily living (14). Beltrame and Hughson (16) established a new outcome called the mean normalized gain (MNG), which is the ratio between the mean normalized $\dot{V}O_{2p}$ and WR amplitudes. In order to calculate MNG, first the system gain must be determined at each harmonic ($gAMP_h$):

$$gAMP_h = \frac{\dot{V}O_{2p}AMP_h}{WR\ AMP_h} \quad (8)$$

Then the respective gains at each harmonic must be normalized as a percentage of the fundamental harmonic gain ($gAMP_1$), in order to remove the influences of the total gain across the harmonic amplitudes (47). The normalization at each harmonic of interest is very important because it enables MNG, and therefore kinetics, to be compared between individuals. Finally, the MNG is calculated by averaging the normalized gains of second, third and fourth harmonics. These two steps have been combined in the following equation:

$$MNG = \frac{\left(\frac{\sum_2^4 gAMP_h}{3}\right)}{gAMP_1} \cdot 100\% \quad (9)$$

Using the gains from the second, third, and fourth harmonics ensures that only harmonics with sufficient input signal power are used in determining the MNG parameter, as the amplitude of the input signal decreases as frequency increases (52). This also helps to avoid the signal nonlinearities that are common at higher frequencies due to circulatory distortion (47).

In comparison to the standard time domain dynamic parameters, MNG has been found to be strongly negatively correlated with τ and MRT (16), which reinforces MNG's ability to accurately quantify the dynamics of the $\dot{V}O_{2p}$ response. One of the advantages of this frequency domain analysis technique is that when the time domain $\dot{V}O_{2p}$ data are transformed into the frequency domain, the random noise in the signal that is a result of breath by breath variation (62) is discarded as higher order harmonics, thus reducing the impact of inter-breath variability on the determination of $\dot{V}O_{2p}$ dynamics (16). By doing frequency domain analysis, it creates a tighter 95% confidence interval for MNG than in the time domain for τ, as well as the width of the 95% confidence interval for MNG is independent of the number of ensemble averaged repetitions of the exercise protocol (16). The tighter confidence interval observed with MNG compared to τ, regardless of the number of ensemble averaged repeats, is a result of the inherent design of pseudorandom sequence exercise of sufficient length, as the repeated changes in WR are intrinsic

to the protocol, in addition to filtering out the high frequency breath by breath noise. Therefore, this is a distinct advantage of MNG over conventional time domain analysis, where multiple repeats are needed to achieve a sufficient signal-to-noise ratio (57).

The frequency domain analysis technique discussed above is not without its own associated uncertainties, as it may not be applicable to all intensities of exercise. According to Eβfeld et al. (37), "Frequency-domain techniques are based upon the following fact: if a sinusoidal wave, characterized by amplitude A_1 and angular frequency ω, is inputted into a linear system then the steady-state output is also a sinusoidal wave of the same frequency but of generally different amplitude A_0 and phase." This suggests that the analysis and interpretation of the signals are dependent on the idea of dynamic linearity, that there is a linear relationship between the system input signal, WR, and the output signal, $\dot{V}O_{2p}$. Because of this dependency, the WRs used in PRBS (16, 17) and PRTS (14, 15) testing protocols have been constrained to be below the VT, where the system dynamics are relatively linear at the evaluated harmonics (38, 47). As discussed in Chapter 1.2 though, even within the moderate domain, the linearity of the WR-$\dot{V}O_{2p}$ relationship is quite questionable, as there have been many situations where $\dot{V}O_{2p}$ displayed non-linear kinetics below VT (24, 49, 56, 60). However, by using the gains that correspond with only the first four harmonics in the calculation of MNG, it keeps the analysis in a range of harmonics that are quasi-linear, which helps to minimize the contamination of the signal, as often these nonlinearities have a more profound effect on the amplitude of harmonics at higher frequencies (38, 47). An area of concern though is how well, or even if, MNG will work at intensities above VT, where the $\dot{V}O_{2p}$ response is much more complex (67, 85). In exercise above VT, Eβfeld et al.'s (37) principle of linearity for frequency domain analysis will be broken, as $\dot{V}O_{2p}$ kinetics are not linear above VT. In that case, the MNG values could be higher just as a result of higher system complexity, and not due to

faster system adjustment (17). That being said, MNG should be tested for its ability to accurately quantify $\dot{V}O_{2p}$ kinetics to exercise stimuli that are above VT. This is worth investigating, because if the parameter is robust, then the range of activity levels that could be evaluated for kinetics during everyday activities would expand from just moderate exercise and activities of daily living to more demanding tasks and sports, where the metabolic requirements of the activities exceed VT.

1.4 Machine Learning Analysis of Wearable Sensor Data to Predict Oxygen Uptake

Developments in wearable sensor technology have made continuous monitoring of physical activity and vital signs outside of the laboratory setting possible and accessible to the general population. Many companies have developed different smart garments or gadgets of varying complexity that are commercially available. Often these devices monitor one or more of the following outcomes: heart rate, respiration, or physical activity. Nevertheless, each of these outcomes is of limited value individually in their raw forms and can be difficult to interpret (71). However, once refined, the data may be used to estimate energy expenditure, and $\dot{V}O_{2p}$ kinetics (13), which may provide impactful information about health and performance.

In order to instantaneously predict $\dot{V}O_{2p}$ from wearable sensors during activities of daily living and exercise, information is needed about the external work being performed and dynamics of the associated aerobic response. During ambulatory activities, accelerometer data provide an indication of external work and energy expenditure (2), as a strong correlation has been observed between accelerometer derived activity counts and steady-state $\dot{V}O_{2p}$ during walking exercise (86). However, since $\dot{V}O_{2p}$ and energy demand are uncoupled during transitions in activity level, a measurement linked to the $\dot{V}O_{2p}$ response is needed in order to capture the response dynamics.

Heart rate is often the variable used in predictive algorithms to encode aerobic response dynamics (2, 12–14). Estimated minute ventilation ($\dot{V}_E$) and breathing frequency from respiration bands may also provide useful information when predicting $\dot{V}O_{2p}$ during exercise, as breathing is tightly regulated to maintain acid-base balance. How estimated $\dot{V}_E$ and breathing frequency change during exercise may contain important information for predicting $\dot{V}O_{2p}$ dynamics. The rate of carbon dioxide production and occurrence of metabolic acidosis that can influence $\dot{V}_E$ differ based on exercise intensity (79), which creates unique relationships between $\dot{V}_E$ and $\dot{V}O_{2p}$ in different exercise intensity domains.

One non-intrusive device that simultaneously records these physiological signals of interest is the Hexoskin smart shirt (Carré Technologies, Montreal, Canada). It continuously records raw waveforms from integrated electrocardiogram (ECG), abdominal and thoracic respiration bands, and a tri-axial hip accelerometer. From the raw waveforms, heart rate, estimated $\dot{V}_E$, breathing frequency, and tri-axial hip acceleration is extracted using Hexoskin's proprietary data processing algorithms. The accuracy of Hexoskin's data outputs have been previously validated by Villar et al. (78) in comparison to standard laboratory measurement methods.

The application of machine learning techniques to analyze these wearable sensor data sets has enabled useful information to be extracted from complex signals. In 2017, Beltrame et al. (13) published research that pushed the boundaries of how and where $\dot{V}O_{2p}$ can be monitored. Using wearable sensor data obtained from Hexoskin smart shirts in conjunction with a machine learning algorithm, they were able to successfully predict the $\dot{V}O_{2p}$ profile of moderate intensity exercise and activities of daily living. From the predicted $\dot{V}O_{2p}$, they were able to transform the data into the frequency domain and extract an estimate of an individual's $\dot{V}O_{2p}$ dynamics using the MNG method (13). Furthermore, they expanded on this finding, and applied the wearable sensor data

machine learning model to unsupervised activities of daily living, in order to monitor $\dot{V}O_{2p}$ throughout the whole day and extract $\dot{V}O_{2p}$ dynamics from those data (14). This type of research, although very early in development, may allow for early detection of changes in cardiorespiratory fitness and health status by identifying changes in $\dot{V}O_{2p}$ kinetics over time (14).

1.5 Thesis Rationale and Impact

With advancements in wearable sensors and artificial intelligence, it is possible to track physiological responses to physical stressors throughout the day (14). However, the machine learning algorithms and MNG used to predict and analyze $\dot{V}O_{2p}$ responses, respectively, have only been employed during low to moderate intensity exercise and activities of daily living due to concerns about the effects of the gross dynamic non-linearities that are observed during higher intensity exercise. Thus, it is unknown if the $\dot{V}O_{2p}$ prediction framework developed by Beltrame et al. (13) is applicable to estimate instantaneous $\dot{V}O_{2p}$ during heavy intensity exercise when the physiological response inputs are much more complex. Additionally, it is unknown if MNG is sensitive to differences in $\dot{V}O_{2p}$ dynamics above versus below VT. Therefore, the purpose of my thesis is to investigate if a machine learning algorithm based on the validated low to moderate intensity exercise $\dot{V}O_{2p}$ predictor developed by Beltrame et al. is appropriate for estimation of $\dot{V}O_{2p}$ during heavy intensity exercise, and to examine MNG as a quantitative indicator of $\dot{V}O_{2p}$ kinetics above VT.

This research advances the development of an instantaneous $\dot{V}O_{2p}$ predictor that can accurately estimate a wide range of metabolic rates and kinetics, while also validating a model-independent form of kinetic analysis (MNG). The successful pairing of accurate and valid dynamic $\dot{V}O_{2p}$ prediction and kinetic analysis will enable an indicator of $\dot{V}O_{2p}$ kinetics to be extracted from non-constant WR conditions without the use of intrusive equipment. Ultimately, this will unlock the potential of assessing $\dot{V}O_{2p}$ kinetics, as an indicator of health (1, 20, 25, 72), fitness (29, 41, 43, 69), or performance (26), in uncontrolled field conditions across a wide range of exercise intensities.

1.6 Research Questions and Hypotheses

Research Questions:

1. Is MNG sensitive to differences in $\dot{V}O_{2p}$ dynamics during different intensity PRBS exercise?
2. Can machine learning analysis of wearable sensor data obtained from a smart shirt be used to accurately estimate instantaneous $\dot{V}O_{2p}$ during both moderate and heavy intensity exercise?

Hypotheses:

1. MNG will be highest during low to moderate, then low to heavy, and lowest during VT to heavy PRBS exercise, indicating that kinetics are fastest during moderate intensity exercise and slowest during solely heavy intensity exercise.
2. Machine learning analysis of wearable sensor data obtained from a smart shirt will accurately estimate instantaneous $\dot{V}O_{2p}$ during different intensities of exercise.

Two studies were performed to address these research questions and hypotheses. Research question 1 is investigated in Chapter 2, and research question 2 is investigated in Chapter 3.

Chapter 2: Frequency Domain Analysis to Extract Dynamic Response Characteristics for Oxygen Uptake during Transitions to Moderate and Heavy Intensity Exercise

This chapter was under review as:

Hedge, ET, Hughson, RL. Frequency domain analysis to extract dynamic response characteristics for oxygen uptake during transitions to moderate and heavy intensity exercise. Under review at J Appl Physiol since June 16, 2020

2.1 Overview

At the onset of an exercise transition, exponential modeling to calculate a time constant (τ) is the conventional method to analyze pulmonary oxygen uptake ($\dot{V}O_{2p}$) kinetics for moderate and heavy exercise. A new frequency domain analysis technique, mean normalized gain (MNG), has been used to analyze $\dot{V}O_{2p}$ kinetics during moderate exercise, but has not been evaluated for its ability to detect differences in kinetics between moderate and heavy exercise. This study tested the hypothesis that MNG would detect smaller amplitude $\dot{V}O_{2p}$ responses in the heavy compared to moderate exercise domains. Eight young healthy adults (3 female; age: 27 ± 6 yr; peak $\dot{V}O_{2p}$: 43 ± 6 $ml \cdot min^{-1} \cdot kg^{-1}$; mean ± SD) performed three bouts of pseudorandom binary sequence (PRBS) exercise for frequency analysis, with work rate (WR) changing between 25W and 90% ventilatory threshold (VT; low to mod), 25W and 50% of the difference between VT and peak $\dot{V}O_{2p}$ (Δ50%; low to heavy), and VT to Δ50% (VT to heavy). Constant WR exercise tests with equivalent step changes in WR to the PRBS tests were performed to facilitate comparison between MNG and τ. MNG was highest for low to mod (59.1 ± 7.5%), then low to heavy (52.1 ± 6.2%), and lowest for VT to heavy (38.4 ± 6.8%, $F_{(2,14)} = 129.755$, $p < 0.001$) exercise conditions indicating slower kinetics with increasing exercise intensity that correlated strongly in repeated measures with τ from step transitions ($r_{rm} = -0.893$). These results indicate that frequency domain analysis and MNG reliably detect differences in $\dot{V}O_{2p}$ kinetics observed across exercise intensity domains.

2.2 New and Noteworthy

Mean normalized gain is able to detect differences in $\dot{V}O_{2p}$ kinetics between moderate, heavy, and heavy intensity exercise from a raised WR within the same individuals. This new method of kinetic analysis may be advantageous compared to conventional curve fitting as it is less sensitive to breath-by-breath noise, it can provide useful information from a single exercise testing session, and it can be applied to non-constant work rate exercise situations.

2.3 Introduction

Pulmonary oxygen uptake ($\dot{V}O_{2p}$) kinetics, which are used to describe how quickly oxidative phosphorylation adjusts to a change in metabolic demand, can provide prognostic information about endurance exercise performance (26, 43, 61) and human health (1, 20, 25, 72). Analysis of $\dot{V}O_{2p}$ kinetics is typically done in the time domain using exponential modelling to characterize the phase II $\dot{V}O_{2p}$ response and calculate a time constant (τ); however, exponential modelling is restricted to only step-wise changes in work rate (WR) to a new constant level, which impedes the clinical utility of $\dot{V}O_{2p}$ kinetics as an indicator of health. Furthermore, modelling the phase II $\dot{V}O_{2p}$ response as a first order-exponential function requires assumptions to be made about the nature of the response (48), which can incorrectly suggest faster kinetics and improved efficiency when transitioning to high compared to moderate intensities (50). Alternatively, the $\dot{V}O_{2p}$ response to randomly varying WR can be analyzed in the frequency domain by calculating a mean normalized gain (MNG) from the amplitude of $\dot{V}O_{2p}$ relative to WR changes (16). Frequency analysis is an attractive method to analyze $\dot{V}O_{2p}$ kinetics as it does not try to fit the data to an explicit model (37), and it has the benefit of innate breath-by-breath noise filtering (16). Because frequency domain analysis is model-independent, MNG has been used to assess $\dot{V}O_{2p}$ kinetics in non-constant WR conditions (13, 15–17) and activities of daily living (14), but the use of MNG as an indicator of kinetics has been limited to exercise and activities that fall within the moderate intensity domain.

The $\dot{V}O_{2p}$ response to a step increase in WR differs between moderate and heavy intensity exercise (i.e., exercise below versus above the ventilatory threshold (VT)). It has been observed that phase II $\dot{V}O_{2p}$ kinetics are slowed when transitioning to heavy compared to moderate intensity exercise from the same baseline condition (36, 51, 53, 61, 65, 68), although others have reported

no difference (7, 8, 67, 88). Furthermore, slower phase II kinetics and increased gain ($\Delta\dot{V}O_{2p}/\Delta WR$) have been observed when initiating heavy exercise transitions from elevated baseline WRs (32, 88). The slower kinetics observed during heavy compared to moderate intensity exercise have also been confirmed using frequency analysis of the $\dot{V}O_{2p}$ response when WR was varied in a sinusoidal pattern with a period of 4-min (42).

Therefore, the purpose of this study was to investigate whether MNG is sensitive to differences in $\dot{V}O_{2p}$ kinetics between moderate and heavy intensity exercise in the same participants. Three different pseudorandom binary sequence (PRBS) exercise protocols were used to elicit rapid changes in metabolic demand that allowed analysis across a range of sinusoidal-equivalent frequencies. We constrained one protocol to the moderate intensity domain (low to mod), a second protocol had WRs fluctuating between moderate and heavy intensity exercise (low to heavy), and a third protocol had WRs constrained to only the heavy intensity domain (VT to heavy). Three different constant WR step transitions that matched the changes in WR used in the PRBS protocols were also performed to facilitate comparison between frequency and time domain analysis techniques. It was hypothesized that MNG will be highest for the low to mod PRBS indicating fastest kinetics, then the low to heavy PRBS, and lowest for the VT to heavy PRBS indicating the slowest kinetics, and that MNG will be negatively correlated with the conventional time domain indicator of kinetics, τ, with increasing exercise intensity.

2.4 Materials and Methods

Participants

Eight young healthy adults (5 males, 3 females; age: 27 ± 6 yr; height: 1.70 ± 0.08 m; mass: 72 ± 13 kg; peak $\dot{V}O_{2p}$: 43 ± 6 $ml \cdot min^{-1} \cdot kg^{-1}$; mean ± SD) participated in the study. Participants

had no known musculoskeletal, respiratory, cardiovascular, or metabolic conditions, and none were taking any medications that might influence cardiorespiratory or metabolic responses to exercise. Participants were instructed to arrive to the laboratory for testing at least two hours post-prandial, abstain from alcohol consumption and vigorous exercise in the 24 hours preceding each test, and to avoid caffeine consumption on the day of the tests. The study was conducted according to the Declaration of Helsinki, except for registration in a trial database, and all procedures were approved by University of Waterloo's Clinical Research Ethics Committee (PRBS exercise: ORE # 32164 and constant WR exercise: ORE #41551). All participants signed an informed consent before participating in the study, and were made aware of their right to withdraw without prejudice.

Exercise Testing

All participants visited the laboratory on eight separate occasions, with each session separated by at least 48 h, to perform six different types of exercise protocols. All exercise tests were performed in an environmentally controlled laboratory on an electrically braked cycle ergometer (Lode Excalibur Sport, Lode B.V., Groningen, Netherlands). Participants were asked to maintain cadence at 60 revolutions per minute for all the exercise tests.

On the first visit, participants performed a ramp-incremental exercise cycling test to symptom-limited tolerance (25 W baseline for 4 min followed by a 25 $W \cdot min^{-1}$ ramp). The test was terminated when the cadence dropped below 55 revolutions per minute despite strong verbal encouragement. Results from the ramp-incremental test were used to estimate each participant's VT, peak $\dot{V}O_{2p}$, and the WRs for the PRBS and constant WR exercise tests.

In the subsequent seven visits, participants performed one of three different PRBS cycling protocols, and two repetitions of two different constant WR cycling protocols in a random order (Fig. 3). For the PRBS tests, WRs intermittently alternated between 25 W and either the WR that corresponded with 90% of the VT (low to mod), 25 W and the WR that corresponds with 50% of the difference between VT and peak $\dot{V}O_{2p}$ (Δ50%; low to heavy), or the WR that corresponds with VT and Δ50% (VT to heavy). The time series for the changes in WR for PRBS protocols were generated by a digital shift register with an adder module feedback (16–18, 52). A single PRBS was composed of 15 units, each of 30 s in duration, totaling 7.5 min. Each complete PRBS testing session consisted of a 3.5 min warm-up (the last 3.5 min of the 7.5 min PRBS), and then two full repetitions of the PRBS for a total of 18.5 min per session. The two different constant WR cycling protocols that were each repeated twice during separate visits were designed to match the transitions in WR used in the PRBS exercise in order to facilitate comparisons between time and frequency domain analysis techniques. One constant WR exercise protocol consisted of two step transitions from a 5 min 25 W baseline to 90% VT for 5 min (low to mod), and one transition from a 5 min VT baseline to Δ50% for 6 min (VT to heavy). The other constant WR exercise protocol consisted of a single step transition from a 5 min 25 W baseline to Δ50% for 6 min (low to heavy).

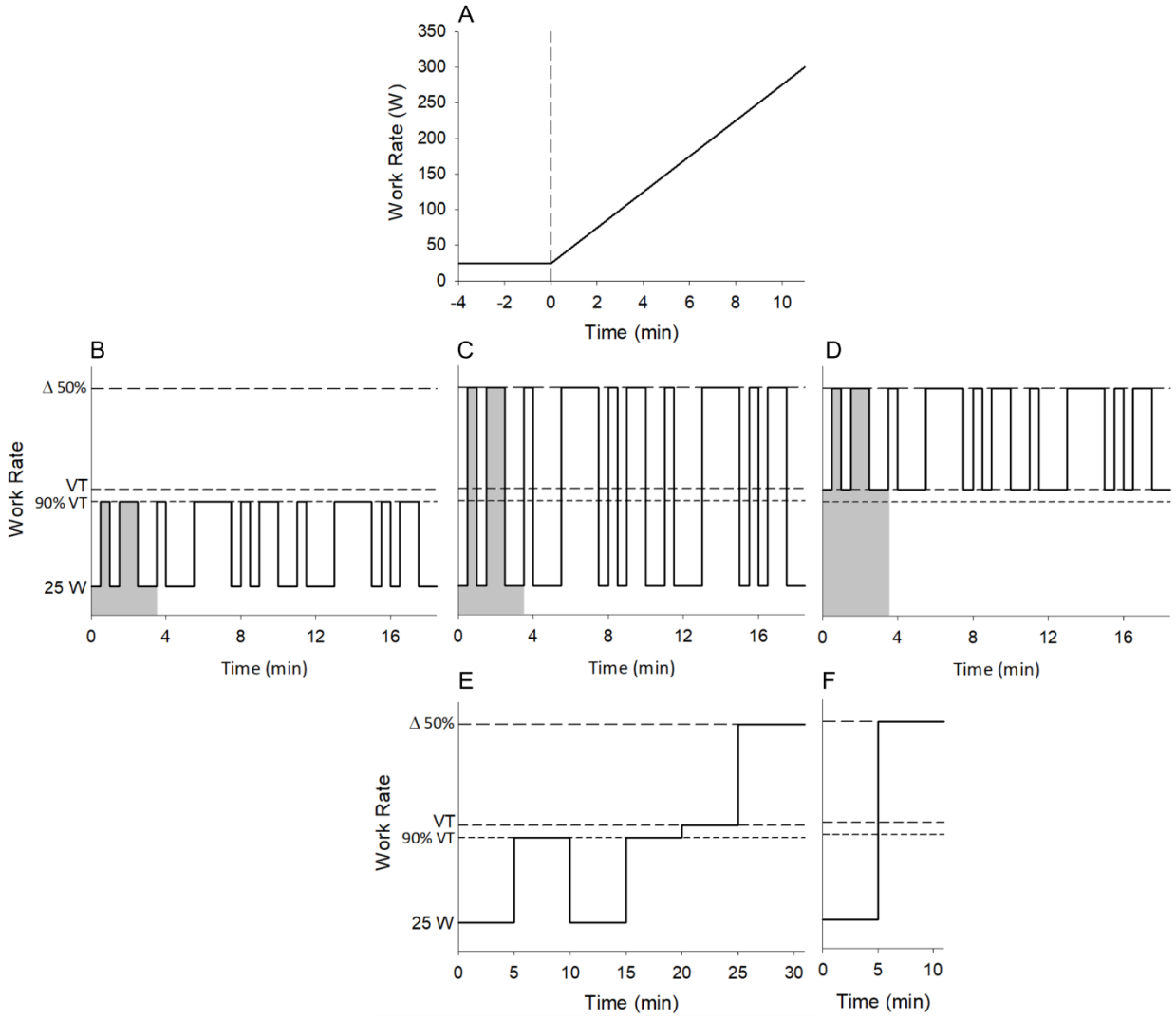

Fig. 3. Schematic showing the six different types of cycling exercise tests. On the first visit, participants performed a 25 W·min^{-1} ramp-incremental exercise test to volitional fatigue (A). In the next seven visits, participants performed three different pseudorandom binary sequence (PRBS) exercise tests (B, C, and D) and two repetitions of two different constant work rate (WR) exercise protocols that contain step transitions that match the changes in WR in the PRBS tests (E and F). Each PRBS test had unique changes in WR, such that WR alternated between 25 W and the WR associated with 90% of the ventilatory threshold (VT) (low to mod; B), between 25 W and the WR associated with 50% of the difference between VT and peak oxygen uptake (Δ50%; low to heavy; C), and between the WR associated with VT and Δ50% (VT to heavy; D). The first 3.5 min of PRBS exercise was considered a warm up (shaded grey), and was followed by two complete repetitions of a 7.5 min pseudorandom WR pattern. One constant WR protocol contained two low to mod, and one VT to heavy exercise step transition (E), while the other constant WR protocol contained a single low to heavy step transition.

Data Collection

For all exercise tests, gas exchange was measured using a portable metabolic system (MetaMax 3B-R2, CORTEX Biophysik, Leipzig, Germany) according to the manufacturer's specifications. Briefly, participants breathed through a mask (7450 SeriesV2™ Mask, Hans Rudolph, Inc., Shawnee, KS, USA), and inspired and expired volumes were measured using a bi-directional volume turbine, which was calibrated before each testing session using a syringe of known volume (3 l). Oxygen and carbon dioxide gas concentrations were continuously sampled at the mouth and were analyzed using a chemical fuel cell and nondispersive infrared sensor, respectively. Gas concentrations were calibrated with precision-analyzed gas mixtures. $\dot{V}O_{2p}$ and carbon dioxide output were calculated using standard breath-by-breath algorithms (80).

Data Analysis

Ramp-incremental exercise response. Peak $\dot{V}O_{2p}$ was defined as the highest $\dot{V}O_{2p}$ computed from a 20 s moving average during the 25 $W \cdot min^{-1}$ ramp test. $\dot{V}O_{2p}$ at VT was estimated by visual inspection using standard ventilatory and gas exchange indices (and their ratios) as previously described (10). WRs at 90% VT, VT, and Δ50% were estimated as the WR associated with the given $\dot{V}O_{2p}$ after left-shifting the $\dot{V}O_{2p}$ response by each individual's response lag time in order to align the $\dot{V}O_{2p}$ and WR profiles. The response lag time was determined using a double-linear model to account for the transit delay time between the exercising legs and lungs, and $\dot{V}O_{2p}$ kinetics at the onset of the ramp. The response lag time was defined as the time at the point of intersection between the forward extrapolation of the average $\dot{V}O_{2p}$ during the last 2 min of 25W

baseline cycling, and the backwards extrapolation of the linear portion of the ramp $\dot{V}O_{2p}$ response that is below VT (59).

Pseudorandom binary sequence exercise response. Breath-by-breath PRBS $\dot{V}O_{2p}$ data were edited on an individual basis to remove aberrant data. The standard aberrant data removal technique used for constant WR exercise, where breaths greater than 3 standard deviations from the local mean are deleted (62), was not possible due to limitations in modelling the frequent changes in $\dot{V}O_{2p}$ that are inherent to PRBS exercise. Instead, a conservative approach was used to edit the data, where breaths that were at least 0.5 l/min away from the local cluster of breaths were identified as suspicious breaths. Tidal volume, breath duration, and the duration of the inspiratory and expiratory phases of the breath were investigated to examine the nature of the suspicious breath. A suspicious breath was only removed if there was not a subsequent compensatory breath. For example, a single breath $\dot{V}O_{2p}$ near zero would be identified, and found to be a result of a very small tidal volume. This breath would be considered erroneous due to aberrant breathing, and consequently, be deleted. Aberrant breaths were only identified and removed in three of the 24 PRBS responses. After editing, breath-by-breath $\dot{V}O_{2p}$ data were linearly interpolated to 1 Hz. For each PRBS tests, the initial 3.5 min of warmup data were excluded *a priori*, and then the two following 7.5 min PRBS repetitions were ensemble averaged, in order to yield a single 7.5 min $\dot{V}O_{2p}$ response for the each of the low to mod, low to heavy, and VT to heavy PRBS exercise conditions, respectively.

Frequency domain analysis was performed on the PRBS data for each individual according to Beltrame and Hughson (16). Briefly, WR and $\dot{V}O_{2p}$ data were transformed into the frequency domain using the discrete Fourier transformation (MATLAB R2016a, TheMathWorks Inc., Natick, MA, USA). The following sinusoidal function for each harmonic was solved:

$$y(t) = a_{DC} + 2 \cdot \sum_{h=1}^{4}\left(A_h \cdot cos(2\pi \cdot h \cdot f_1 \cdot t) + B_h \cdot sin(2\pi \cdot h \cdot f_1 \cdot t)\right) \quad (5)$$

where *y(t)* is the signal at a given point in the time domain, a_{DC} is the system DC offset amplitude (i.e., the average response), f_1 is the fundamental frequency (1/period of the sinusoid, or the duration of a single cycle of the PRBS; 450 s), *t* is the time, A_h and B_h are the cosine and sine amplitude coefficients for a given harmonic, *h* (an integer), respectively. From A_h and B_h, the total amplitude (AMP_h) was calculated for each harmonic:

$$AMP_h = \sqrt{A_h^2 + B_h^2} \quad (6)$$

Subsequently, the system gain was determined at each harmonic ($gAMP_h$):

$$gAMP_h = \frac{\dot{V}O_{2p}AMP_h}{WR\ AMP_h} \quad (8)$$

Then the respective gains at each harmonic were normalized as a percentage of the fundamental harmonic gain ($gAMP_1$), in order to remove the influences of the total gain across the harmonic amplitudes (47). Finally, the MNG was calculated by averaging the normalized gains of second, third and fourth harmonics. These two steps have been combined in the following equation:

$$MNG = \frac{\left(\frac{\sum_2^4 gAMP_h}{3}\right)}{gAMP_1} \cdot 100\% \quad (10)$$

Limiting the calculation of MNG to low frequencies (< 0.01 Hz) ensures that only harmonics with sufficient input signal power are used in determining the MNG, as the amplitude of the input signal decreases as frequency increases (52), and it avoids the $\dot{V}O_{2p}$ response nonlinearities that are common at higher frequencies due to circulatory distortion (47).

Constant work rate exercise response. Breath-by-breath constant WR $\dot{V}O_{2p}$ data were edited on an individual basis to remove aberrant data that were greater than 3 standard deviations from the local mean (62), as previously described (57). After editing, data were linearly interpolated to 1 Hz, and like-repetitions were time aligned such that *t* = 0 s represented the onset

of each step transition in WR. Four low to mod, two low to heavy, and two VT to heavy step transitions were ensemble averaged for each participant to yield a single response for each exercise condition, respectively.

The on-transient $\dot{V}O_{2p}$ response for each step change in WR was fitted with a mono-exponential function (49, 83), using non-linear least squares regression (MATLAB R2016a, TheMathWorks Inc.,):

$$\dot{V}O_{2p}(t) = \dot{V}O_{2pBSL} + A_p \cdot \left(1 - e^{-(t-TD)/\tau}\right) \quad (2)$$

where *t* is time, $\dot{V}O_{2pBSL}$ is the two min pre-transition baseline average $\dot{V}O_{2p}$ at 25 W or VT, A_p is the phase II amplitude for the change in $\dot{V}O_{2p}$, τ is the phase II time constant, and *TD* is the time delay of the exponential function. For all transitions, the cardio-dynamic phase was identified and excluded from analysis by progressively moving the start of the fitting window (from ~35 s after the transition) back towards the transition onset while examining the flatness of the residual profile, τ, and τ 95% confidence interval. For low to mod transitions, the phase II fitting window was limited to approximately 5 τ in order to optimize the fit through the transient data rather than the steady-state data (11). For the heavy intensity transitions, the end of the phase II fitting window was identified as the point immediately preceding a systematic increase in τ and the τ 95% confidence interval, and by examining the flatness of the residual profile. The phase II gain (G_p) was calculated by dividing the A_p by the change in WR. Root mean square error (RMSE) was used to assess modelling error. The amount of noise in in the ensemble averaged $\dot{V}O_{2p}$ signal was calculated for each participant as the standard deviation (in $l \cdot min^{-1}$) of the $\dot{V}O_{2p}$ signal during the last minute of the constant WR exercise. Signal-to-noise ratio of the phase II response for each exercise transition was calculated by dividing the phase II amplitude by the amount of noise.

Statistical Analysis

Statistical analyses were conducted in R (Version 3.5.1). Unless otherwise noted, all analyses used R's standard set of libraries. Shapiro-Wilks normality tests were used to justify the use of parametric statistical testing. One way repeated measures analysis of variance (ANOVA) were used to compare kinetic parameters and variables between the low to mod, low to heavy, and VT to heavy exercise conditions for different PRBS and constant WR protocols, respectively. For all ANOVAs, if Mauchly's test of sphericity was significant, then the Greenhouse-Geisser correction was used. Where significant main effects were found, paired t-tests with Bonferroni corrected *p*-values were used for *post-hoc* analysis. Repeated measures correlation (5) was used to examine the relationship between the MNG and the phase II τ within a participant (addressing the violated assumption of independence), across the three different intensities (R Library: rmcorr). Pearson product-moment correlation was also used to examine the relationship between MNG and phase II τ between participants within each of the three exercise conditions. Statistical significance was set to $p < 0.05$, and data are reported as means ± SD.

2.5 Results

The group mean (n = 8) peak $\dot{V}O_{2p}$ determined from ramp incremental exercise was 3.08 ± 0.57 l·min^{-1}, which corresponded to peak WR values of 273 ± 33 W. The $\dot{V}O_{2p}$ at the estimated VT was 1.68 ± 0.18 l·min^{-1}, which occurred at a WR of 119 ± 6 W after accounting for a ramp incremental $\dot{V}O_{2p}$ response lag time of 47 ± 10 s.

The group mean (n = 8) $\dot{V}O_{2p}$ responses for the low to mod, low to heavy, and VT to heavy constant WR exercise are displayed in Fig. 4A. Group mean step changes in WR for low to mod,

low to heavy, and VT to heavy conditions were 83 ± 7 W, 164 ± 26 W, and 69 ± 21 W, respectively. Table 1 shows the estimated parameters of the phase II exponential models used to describe the $\dot{V}O_{2p}$ response to constant WR exercise. Main effects of condition were observed for $\dot{V}O_{2pBSL}$ ($F_{(2,14)} = 1581.892$, $p < 0.001$), and A_p ($F_{(1.129,7.903)} = 71.423$, $p < 0.001$). A main effect of condition was found for τ ($F_{(2,14)} = 41.623$, $p < 0.001$), such τ was shortest in the low to mod condition, then low to heavy, and longest in the VT to heavy condition (Fig. 5). A main effect of condition was also found for G_p ($F_{(1.117,7.817)} = 10.211$, $p = 0.012$), such that G_p was larger in the VT to heavy condition than both low to mod ($p = 0.046$) and low to heavy ($p = 0.023$), but G_p was not significantly different between the low to mod and low to heavy conditions ($p = 0.335$). Main effects of condition were also observed for phase II fit RMSE ($F_{(2,14)} = 13.591$, $p = 0.001$), and signal-to-noise ratio ($F_{(2,14)} = 14.894$, $p < 0.001$). Phase II fit RMSE was significantly increased for both heavy exercise conditions compared to the low mod condition (low to heavy: $p = 0.003$; VT to heavy: $p = 0.027$), but were not different from each other. However, signal-to-noise ratio was found to be similar between low to mod and low to heavy conditions (p = 0.184), but was significantly lower than both during VT to heavy exercise (low to mod: $p = 0.039$; low to heavy: $p = 0.002$).

The group mean (n = 8) $\dot{V}O_{2p}$ responses to low to mod, low to heavy, and VT to heavy PRBS exercise are displayed in Fig. 4B. A main effect of condition was found for MNG ($F_{(2,14)} = 129.755$, $p < 0.001$), such that MNG was highest in the low to mod condition, then low to heavy, and lowest in the VT to heavy condition (Fig. 6).

Repeated measures correlation analysis, which combines the kinetic data from all three exercise conditions, revealed a strong negative linear correlation between τ and MNG (Fig. 7; $r_{rm} = -0.893$, $p < 0.001$). When examining the relationship between τ and MNG for each exercise

condition individually, significant negative linear correlations were observed in the low to mod (r = -0.873, p = 0.005) and low to heavy (r = -0.716, p = 0.046) conditions, but not in the VT to heavy condition (r = -0.575, p = 0.136).

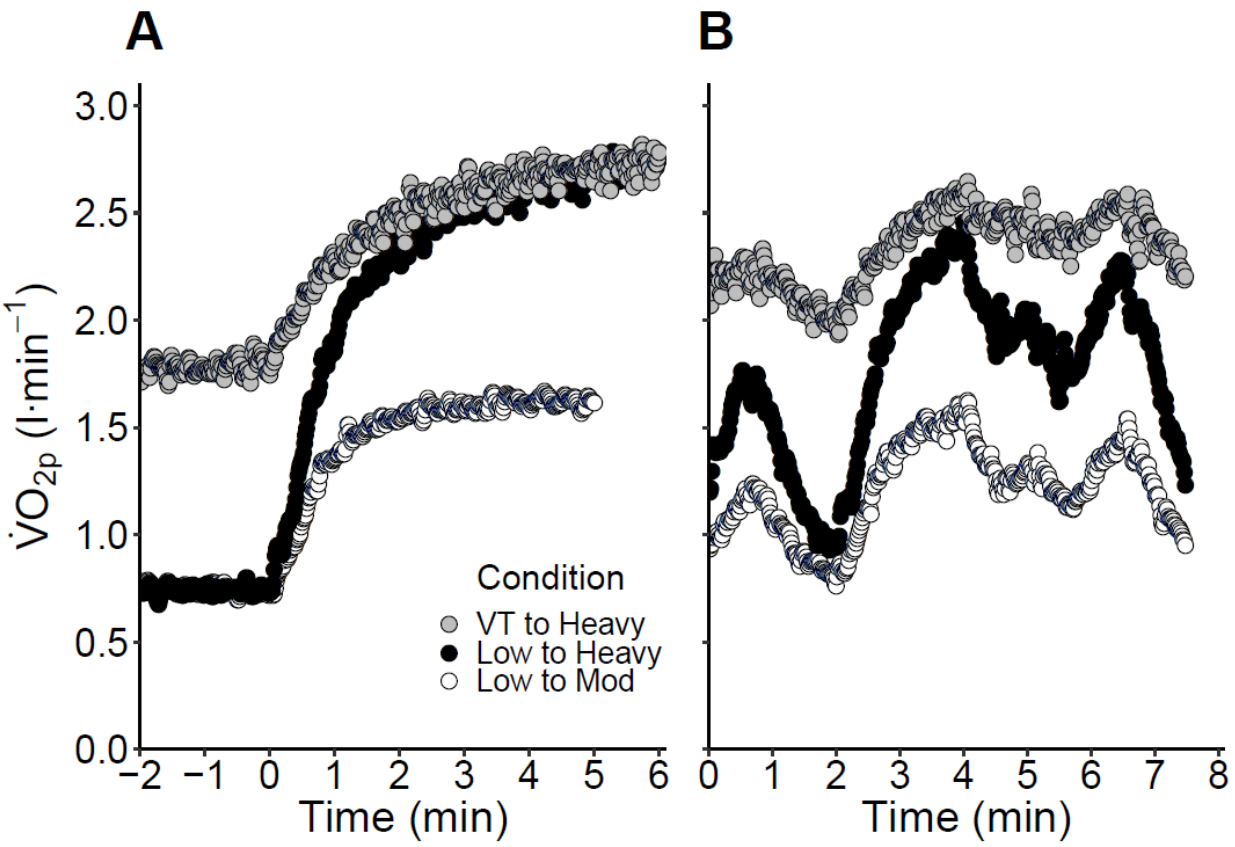

Fig. 4. Group mean (n = 8) pulmonary oxygen uptake ($\dot{V}O_{2p}$) responses to constant work rate (WR) exercise (A), and pseudorandom binary sequence (PRBS) exercise (B). For constant WR exercise, WR increased instantaneously toward a new constant level at time = 0. White, black, and grey circles represent low to moderate, low to heavy, and ventilatory threshold (VT) to heavy exercise $\dot{V}O_{2p}$ responses.

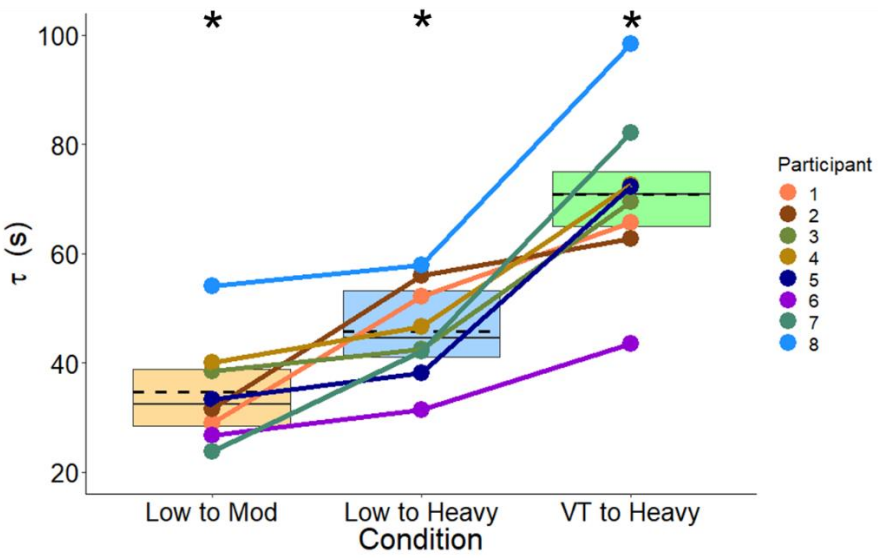

Fig. 5. Comparison of the phase II time constant (τ) between low to moderate, low to heavy, and ventilatory threshold (VT) to heavy exercise conditions. Each color represents a single participant. Lower and upper limits of box plots represent the first and third quartiles, and solid and dashed horizontal lines in the boxes represent the median and mean, respectively. * indicates a significant ($p < 0.05$) difference from all other conditions.

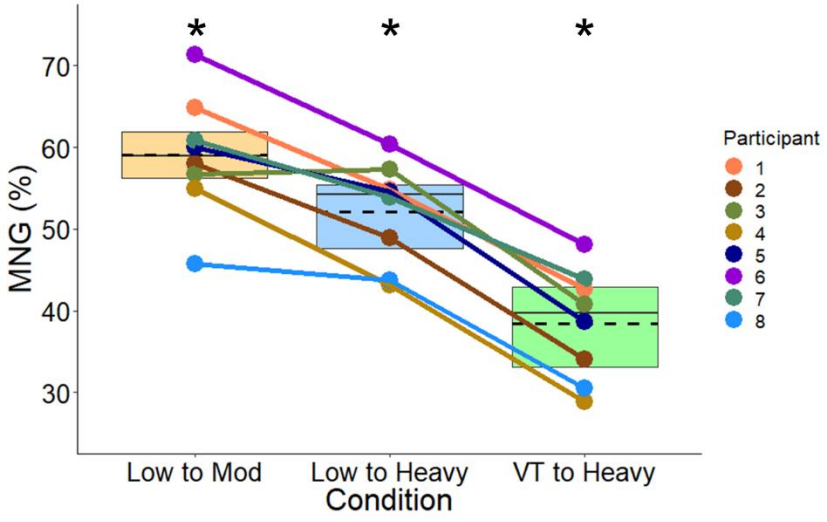

Fig. 6. Comparison of mean normalized gain (MNG) between low to moderate, low to heavy, and ventilatory threshold (VT) to heavy exercise conditions. Each color represents a single participant. Lower and upper limits of box plots represent the first and third quartiles, and solid and dashed horizontal lines in the boxes represent the median and mean, respectively. * indicates a significant ($p < 0.05$) difference from all other conditions.

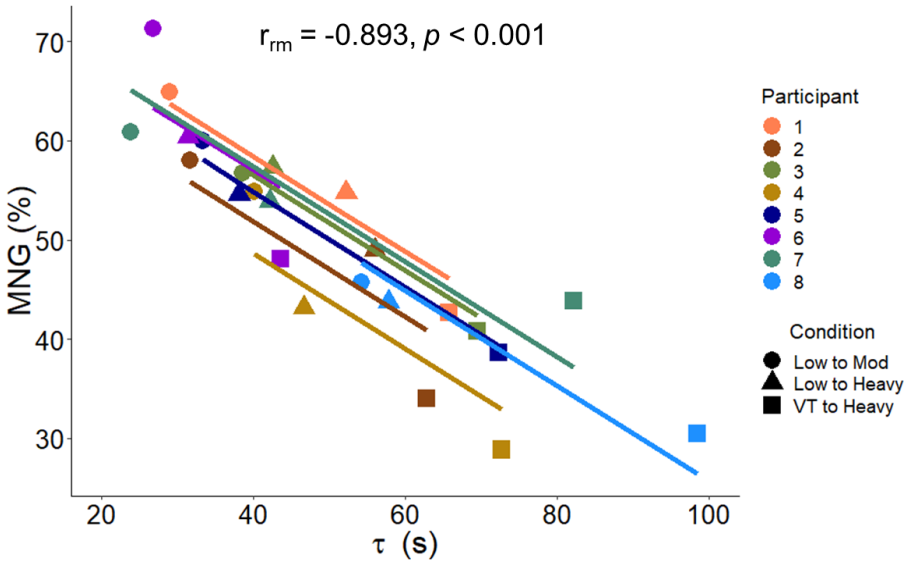

Fig. 7. Relationship between mean normalized gain (MNG) and the phase II time constant (τ) when combining all participants' (n = 8) repeated measures across low to moderate (circles), low to heavy (triangles), and ventilatory threshold (VT) to heavy (squares) exercise conditions. A different color represents each participant. Parallel lines represent the best fit regression line slope across conditions between participants, and the length of each parallel line is constrained to the bounds of each participant's response range.

Table 1. Group mean (n = 8) parameter estimates for pulmonary oxygen uptake kinetics during constant work rate and pseudorandom binary sequence exercise.

	Low to Mod	Low to Heavy	VT to Heavy
$\dot{V}O_{2pBSL}$, $l \cdot min^{-1}$	0.75 ± 0.06 [c]	0.74 ± 0.06 [c]	1.78 ± 0.11 [a,b]
A_p, $l \cdot min^{-1}$	0.85 ± 0.08 [b]	1.77 ± 0.33 [a,c]	0.86 ± 0.30 [b]
TD, s	15 ± 3	10 ± 3	1 ± 13
τ, s	35 ± 10 *	46 ± 9 *	71 ± 16 *
G_p, $ml \cdot min^{-1} \cdot W^{-1}$	10.3 ± 0.4 [c]	10.8 ± 0.4 [c]	12.3 ± 1.5 [a,b]
RMSE, $ml \cdot min^{-1}$	59 ± 20 [b,c]	99 ± 31 [a]	104 ± 47 [a]
Signal-to-noise ratio	16.8 ± 7.7 [c]	22.3 ± 12.2 [c]	10.4 ± 8.0 [a,b]
MNG, %	59.1 ± 7.5 *	52.1 ± 6.2 *	38.4 ± 6.8 *

Values are mean ± SD. A_p, phase II amplitude; G_p, phase II gain; MNG, mean normalized gain; RMSE, phase II fit root mean square error; TD, time delay; τ, phase II time constant; $\dot{V}O_{2pBSL}$, baseline pulmonary oxygen uptake; VT, ventilatory threshold. * significant difference ($p < 0.05$) between all conditions. [a,b,c] significant difference ($p < 0.05$) from low to mod, low to heavy, or VT to heavy exercise condition, respectively.

2.6 Discussion

The purpose of this study was to examine how MNG, as an indicator of $\dot{V}O_{2p}$ kinetics, differs between moderate and heavy intensity exercise. In agreement with our hypotheses, we found that MNG was highest for low to mod exercise, then low to heavy, and lowest for VT to heavy exercise, and that these MNG values were negatively correlated with the standard time domain indicator of kinetics, phase II τ, which was obtained from the same participants. Thus, MNG is sensitive to differences in $\dot{V}O_{2p}$ kinetics between moderate and heavy intensity exercise, as well as differences in kinetics when performing exercise transitions from an elevated WR.

This study is the first to use PRBS exercise to investigate $\dot{V}O_{2p}$ kinetics above VT. Previous investigators (13–16, 37), have limited the use of pseudorandom exercise and the associated frequency domain analysis to moderate intensity exercise in order to assume that the WR-$\dot{V}O_{2p}$ transfer system is linear (i.e., a sinusoidal WR input will create a sinusoidal $\dot{V}O_{2p}$ output of the same frequency, but generally of a different amplitude and phase). However, it is known even within the moderate intensity domain that the $\dot{V}O_{2p}$ response to step incremental exercise is non-linear, as kinetics slow and gain increases with increasing baseline WR (21, 24, 49, 60). Furthermore, the muscle oxygen uptake kinetics that are approximated by gas exchange measured at the mouth are distorted by the circulatory system in transit to the lungs (37). To circumvent these discrepancies from a linear system, the frequencies used in the calculation of MNG are limited to low frequencies (< 0.01 Hz) that have sufficient input signal power, which allows the system to be regarded as quasi-linear (16). One of the major concerns about using MNG as an indicator of kinetics for heavy exercise is that the on- and off-transient kinetics observed during heavy exercise are not symmetrical, whereas they are approximately symmetrical in the moderate intensity domain (67, 68). Based on the results of this study, it appears that the assumption of on- and off-transient symmetry is not crucial for the calculation MNG, as MNG was still sensitive to differences in $\dot{V}O_{2p}$ kinetics, and strongly related to τ across the evaluated exercise intensities despite the differences in on- and off-transient kinetics.

MNG appears to be sensitive to differences in $\dot{V}O_{2p}$ kinetics between low to mod, low to heavy, and VT to heavy exercise. MNG was reduced, indicating slower kinetics, during heavy compared to moderate exercise from the same lower PRBS WR (low to mod versus low to heavy). A further reduction in MNG was observed when performing VT to heavy PRBS exercise. The $\dot{V}O_{2p}$ kinetics extracted from PRBS exercise responses by the MNG method were mirrored by

conventional time domain analysis of the constant WR $\dot{V}O_{2p}$ responses. The longer observed phase II τ in the low to heavy step transition compared to the low to mod condition is consistent with previous reports that found the phase II τ to be longer for step transitions in WR to heavy exercise intensities compared to moderate intensities from the same baseline WR (36, 53, 61, 65, 68), although it does conflict with others (7, 8, 67, 88). A further increase in phase II τ was observed in the heavy intensity step transition from an elevated baseline WR, which is also in agreement with previous works that have found $\dot{V}O_{2p}$ kinetics to be slower when transitioning from an elevated baseline WR (32, 56, 88). Furthermore, the finding that there is a strong negative linear repeated measures correlation between MNG and τ across exercise intensities within the same group of participants demonstrates that MNG is providing an accurate index of $\dot{V}O_{2p}$ kinetics across a wide range of values, regardless of the exercise intensity domain.

Between-participant correlations for each exercise condition revealed strong negative linear correlations between τ and MNG for both the low to mod, and low to heavy exercise conditions, but no significant correlation was found in the VT to heavy condition. The strong negative correlation between τ, derived from constant WR exercise tests, and MNG, derived from PRBS exercise tests, in the low to moderate condition supports the findings of Beltrame and Hughson (16), who also found a similar relationship during PRBS cycling, albeit all their participants performed moderate PRBS exercise of the same absolute WRs (25-100 W), whereas in the present study participants exercised at similar relative intensities, not absolute WRs (25 W to 90% of VT). Additionally, their τ was derived from the two minute on-transient step contained within the PRBS (16). In the present study, we were able to confirm the relationship observed by Beltrame and Hughson (16) using τ derived under ideal time domain analysis conditions (i.e., a single step increase in WR to a constant level). A strong negative linear relationship was also

observed between τ and MNG in the low to heavy condition, as would be expected based on the moderate intensity results.

Interestingly, a significant relationship between τ and MNG was not observed in the VT to heavy condition ($p = 0.136$). The lack of a significant relationship in this condition is likely due to greater breath-by-breath variation and lower phase II amplitude, which is reflected by the signal-to-noise ratio. The signal-to-noise ratio was similar between the low to mod and low to heavy exercise conditions, where significant relationships were found, but was diminished in the VT to heavy condition. Beltrame and Hughson (16) have previously reported that MNG is ~50% less susceptible than τ to the effects of breath-by-breath noise, which is known to create uncertainty in estimating exponential model parameters (62). For this reason, MNG may be considered a more robust form of kinetic analysis than exponential modelling, as the accuracy of MNG is independent of the number of ensemble averaged PRBS repetitions, whereas the accuracy of τ is highly dependent on the number of ensemble averaged step transitions (16). The robustness of MNG to noise is partially a result of its intrinsic filtering characteristics, as the breath-by-breath noise is discarded with the higher order harmonic frequencies and is not included in the calculation of MNG (16). Therefore, the lack of correlation is thought to be a result of inaccuracy in time domain exponential curve fitting due to reduced signal and increased noise, and not because MNG is not accurately reflecting the differences in kinetics. If a greater number of VT to heavy step transitions were performed and ensemble averaged to improve the signal-to-noise ratio, then it is likely that the correlation between τ and MNG within the VT to heavy condition would be significant.

MNG's ability to detect differences in $\dot{V}O_{2p}$ kinetics between exercise conditions within the same participants suggests that MNG may be a promising new analysis method to track changes in aerobic fitness over time due to lifestyle adjustments or interventions. Because MNG

is a model-independent form of analysis, the types of data that MNG can be used to analyze and extract kinetics from are not limited to just pseudorandom exercise performed in a laboratory setting. MNG has previously been used to extract an indicator of $\dot{V}O_{2p}$ kinetics from $\dot{V}O_{2p}$ responses predicted by machine learning analysis of data obtained from wearable sensors integrated into a smart shirt during unsupervised activities of daily living (14). The combination of a non-intrusive, continuous $\dot{V}O_{2p}$ prediction framework (13) and MNG kinetic analysis method allows $\dot{V}O_{2p}$ kinetics, and therefore, aerobic fitness, to be evaluated outside of the laboratory environment. Based on the current study, it appears that the range of exercise intensities that can be assessed outside of the laboratory can now be extended to the heavy intensity domain.

The application of the MNG method to analyze heavy intensity $\dot{V}O_{2p}$ data may be particularly useful in assessing aerobic fitness during "real-world" exercise training and competitions, where exercise intensities often are in the heavy domain. High intensity interval training, which is characterized by interspersed bursts of vigorous activity and low intensity recovery, is a common form of exercise training used to improve aerobic fitness and performance, as two weeks of high intensity interval training has been reported to speed phase II kinetics and improve high intensity exercise tolerance (4). Frequency analysis may be used as a unique method to extract an indicator of aerobic fitness from high intensity interval training sessions, and provide insight into the time course of aerobic system adaptation over a period of training without requiring additional testing. Furthermore, MNG may provide useful information about athlete aerobic fitness and performance during various intermittent high intensity sport competitions (6, 22, 30, 75), if wearable sensors are permitted.

Limitations. Constant WR and PRBS exercise differ in how $\dot{V}O_{2p}$ is able to respond. In constant WR tests, the exercise transitions were performed from a stable baseline $\dot{V}O_{2p}$, whereas

in PRBS tests, each transition to increased WR occurred before the effect of the previous increased WR on $\dot{V}O_{2p}$ had been reduced to baseline. It has been previously demonstrated that prior moderate intensity exercise does not alter $\dot{V}O_{2p}$ kinetics in sequential moderate intensity like-repetitions when $\dot{V}O_{2p}$ is able to return to the baseline level (74). However, it is known that prior heavy intensity exercise does alter the $\dot{V}O_{2p}$ response to subsequent heavy intensity exercise (34, 35, 63). Furthermore, PRBS exercise transitions in WR were initiated from varying metabolic rates within a single condition, as the pre-transition metabolic rate depends on how long the prior WR had been maintained. This is different from the constant WR exercise conditions, where WR transitions were initiated from specific metabolic rates. It has been suggested that a raised metabolic rate is responsible for slowing $\dot{V}O_{2p}$ kinetics independent of WR during moderate intensity exercise (21); although, it has been reported that a raised metabolic rate relative to baseline WR does not affect $\dot{V}O_{2p}$ kinetics during heavy intensity exercise (32). Therefore, it is unknown how much $\dot{V}O_{2p}$ kinetics may vary within a given PRBS test. Consequently, MNG is viewed as a summary metric of many different kinetic responses contained within PRBS exercise, whereas the τ from constant WR exercise is viewed as a specific descriptor of the kinetic response to a single step change in WR performed under very specific conditions.

Conclusion. Previous works have limited the use of MNG to moderate exercise intensities. The present study demonstrated for the first time that MNG is sensitive to differences in kinetics between moderate, heavy, and heavy exercise from an elevated WR, and that MNG across exercise intensities is related to τ. Therefore, the range of exercise intensities that MNG can be used for analysis of $\dot{V}O_{2p}$ kinetics can be extended into the heavy intensity domain. Analysis of exercise responses with MNG also provides the opportunity to assess $\dot{V}O_{2p}$ kinetics in a way that is not as sensitive as exponential modelling to the signal-to-noise ratio, and can be done in a single exercise

session, possibly reducing participant laboratory visits and testing burden. Future research should evaluate if MNG is sensitive to changes in $\dot{V}O_{2p}$ kinetics as a result of exercise training interventions, or sedentary behavior.

Chapter 3: Evaluation of Random Forest Regression Model to Predict Oxygen Uptake Kinetics during Moderate and Heavy Exercise from Wearable Sensor Data

3.1 Overview

Non-intrusive estimation of pulmonary oxygen uptake ($\dot{V}O_{2p}$) has been enabled by advances in wearable sensor technology and artificial intelligence. $\dot{V}O_{2p}$, and its dynamics, have been accurately predicted during moderate exercise using easy-to-obtain sensor inputs. However, $\dot{V}O_{2p}$ prediction algorithms for heavy exercise are still under development. It was hypothesized that a machine learning model based on a previously validated moderate intensity algorithm would be able to accurately estimate $\dot{V}O_{2p}$ during both moderate and heavy exercise, and predict slower $\dot{V}O_{2p}$ kinetics during heavy compared to moderate exercise. Fifteen young healthy adults (7 females; peak $\dot{V}O_{2p}$: 42 ± 5 $ml \cdot min^{-1} \cdot kg^{-1}$) performed a $25W \cdot min^{-1}$ ramp-incremental exercise test to the limit of tolerance, and subsequently, three different pseudorandom binary sequence (PRBS) exercise tests ranging in intensity from low to moderate, low to heavy, and ventilatory threshold to heavy work rates. A random forest regression algorithm was developed to predict instantaneous $\dot{V}O_{2p}$ using the ramp-incremental and PRBS exercise responses, with model features including percent heart rate reserve, estimated minute ventilation, and breathing frequency. Frequency domain analyses were used to extract an indicator of kinetics from the measured and predicted $\dot{V}O_{2p}$ responses. Predicted $\dot{V}O_{2p}$ had low bias ($0.022\ l \cdot min^{-1}$), and was strongly correlated (r_{rm} = 0.911, $p < 0.001$) with the measured $\dot{V}O_{2p}$. Comparison of predicted and measured kinetics revealed a Condition x Data type interaction ($F_{(1.369, 19.167)} = 5.339$, $p = 0.023$), and a bias for slower prediction kinetics. A moderate within-participant correlation between predicted and measured kinetics was observed ($r_{rm} = 0.413$, $p = 0.021$), but no significant between-participant relationships were found. Therefore, the random forest regression algorithm predicts overall changes in $\dot{V}O_{2p}$, but does not precisely predict $\dot{V}O_{2p}$ kinetics when extended to heavy intensity exercise. New models should be explored to improve the characterization of $\dot{V}O_{2p}$ kinetics.

3.2 Introduction

Measuring pulmonary oxygen uptake ($\dot{V}O_{2p}$) enables precise quantification of aerobic energy expenditure, and assessment of how quickly aerobic metabolism is able to adjust to a change in energy demand (i.e., $\dot{V}O_{2p}$ kinetics). $\dot{V}O_{2p}$ kinetics are an indicator of aerobic fitness level (29, 41, 43, 69), which is predictive of endurance exercise performance (26), and health status (1, 20, 25, 72). However, the measurement of $\dot{V}O_{2p}$ outside of the laboratory setting is not widespread as a result of the intrusive equipment and technical expertise that is required to operate the equipment.

Advances in wearable sensors and machine learning techniques have enabled the development of a continuous, minimally-intrusive $\dot{V}O_{2p}$ monitoring system to be used outside of the laboratory environment (14). Wearable sensors are able to acquire large amounts of physiologically relevant data, such as heart rate, breathing frequency and minute ventilation ($\dot{V}_E$), which can then be used to estimate $\dot{V}O_{2p}$. Supervised machine learning techniques can be applied to determine the complex input-output relationship between the wearable sensor data and directly measured $\dot{V}O_{2p}$ in order to develop a $\dot{V}O_{2p}$ predictor (12, 17). Recently, Beltrame et al. (13) developed a dynamic $\dot{V}O_{2p}$ predictor for low to moderate intensity exercise, and activities of daily living. The $\dot{V}O_{2p}$ predictor was found to accurately estimate energy expenditure, and $\dot{V}O_{2p}$ kinetics during paced walking, and simulated activities of daily living (13). However, the $\dot{V}O_{2p}$ predictor was not evaluated for its ability to estimate $\dot{V}O_{2p}$ during heavy intensity exercise. Within the moderate intensity domain, $\dot{V}O_{2p}$ kinetics are relatively linear, although slowing of kinetics is acknowledged with increasing exercise intensity (49, 60), and on- and off-transient kinetics are symmetrical (67, 68). The response is more complex during heavy intensity exercise, as $\dot{V}O_{2p}$ kinetics are much slower (36, 42, 53, 61, 65, 68), on- and off-transient kinetics are asymmetrical

(67, 68), and there is a secondary increase in $\dot{V}O_{2p}$ known as the slow component (81). Furthermore, the relationship between $\dot{V}_E$ and $\dot{V}O_{2p}$ changes as a function of exercise intensity (79). Accordingly, prediction of $\dot{V}O_{2p}$ during both moderate and heavy intensity exercise using a single predictive model is likely more complex than during moderate intensity exercise alone due to non-linear relationships across exercise intensity domains.

Therefore, the purpose of this study was to develop and evaluate a $\dot{V}O_{2p}$ predictor that uses easy-to-obtain wearable sensor data, based on Beltrame et al.'s (13) previously validated moderate intensity algorithm, to estimate both moderate and heavy intensity dynamic exercise responses. The prediction algorithm was evaluated in terms of its accuracy in estimating $\dot{V}O_{2p}$, and $\dot{V}O_{2p}$ kinetics. It was hypothesized that machine learning analysis of wearable sensor data will provide an accurate estimation of $\dot{V}O_{2p}$, and predict slower $\dot{V}O_{2p}$ kinetics during heavy compared to moderate intensity exercise.

3.3 Methods

Participants

Fifteen young healthy adults (8 males, 7 females; age: 27 ± 5 yr; height: 1.70 ± 0.08 m; mass: 71 ± 11 kg; peak $\dot{V}O_{2p}$: 42 ± 5 $ml \cdot min^{-1} \cdot kg^{-1}$) participated in the study, and were used to develop and evaluate the machine learning model. The ramp-incremental $\dot{V}O_{2p}$ responses and measured mean normalized gains (MNGs) of 8 of the 15 participants have been included in analyses reported elsewhere (Chapter 2), where those participants also completed constant work rate (WR) exercise tests during additional testing sessions. A separate group of 7 young healthy adults (5 males, 2 females; age: 25 ± 5 yr; height: 1.73 ± 0.08 m; mass: 69 ± 11 kg; peak $\dot{V}O_{2p}$: 41

± 7 $ml \cdot min^{-1} \cdot kg^{-1}$) were additionally recruited to participate in this study, and were used as an independent validation dataset to optimize the machine learning model. Participants had no known musculoskeletal, respiratory, cardiovascular, or metabolic conditions, and none were taking any medications that might influence cardiorespiratory or metabolic responses to exercise. Participants were instructed to arrive to the laboratory for testing at least two hours post-prandial, abstain from alcohol consumption and vigorous exercise in the 24 hours preceding each test, and to avoid caffeine consumption on the day of the tests. The study was conducted according to the Declaration of Helsinki, except for registration in a trial database, and all procedures were approved by a University of Waterloo Research Ethics committee (ORE # 32164). All participants signed an informed consent before participating in the study.

Exercise Testing

Participants visited the laboratory on four separate occasions to perform a ramp-incremental exercise test, and three different pseudorandom binary sequence (PRBS) exercise tests. Each exercise session was separated by at least 48 h. All exercise tests were performed in an environmentally controlled laboratory on an electronically braked cycle ergometer (Lode Excalibur Sport, Lode B.V., Groningen, Netherlands). Participants were instructed to maintain cadence at 60 revolutions per minute for all exercise tests.

On the first visit, 5 min of seated resting data were collected to determine each participant's resting heart rate. After the resting period, participants performed a ramp-incremental exercise test to symptom-limited tolerance (25 W baseline for 4 min followed by a 25 $W \cdot min^{-1}$ ramp) in order to determine each participant's ventilatory threshold (VT), peak $\dot{V}O_{2p}$, and the WRs for the PRBS

exercise tests. The test was terminated when the cadence dropped below 55 revolutions per minute despite strong verbal encouragement. Peak $\dot{V}O_{2p}$ was defined as the highest $\dot{V}O_{2p}$ computed from a 20 s moving average during the exercise test. $\dot{V}O_{2p}$ at VT was estimated by visual inspection using standard ventilatory and gas exchange indices, and their ratios, as previously described (10). WRs at 90% VT, VT, and the WR that corresponds with 50% of the difference between VT and peak $\dot{V}O_{2p}$ (Δ50%) were estimated by left-shifting the $\dot{V}O_{2p}$ response by each individual's response lag time in order to align the $\dot{V}O_{2p}$ and WR profiles. The response lag time was estimated as the time after the onset of the ramp at the point of intersection between the forward extrapolation of the average $\dot{V}O_{2p}$ during the last 2 min of 25 W baseline cycling, and the backwards extrapolation of the linear portion of the ramp $\dot{V}O_{2p}$ response that is below VT (59).

In visits 2-4, participants performed one of three different PRBS exercise tests in a randomized order. For the three PRBS exercise test, WRs intermittently alternated between 25 W and 90% of the VT (low to mod), 25 W and Δ50% (low to heavy), or VT and Δ50% (VT to heavy) (Fig. 8). The time series for the changes in WR for PRBS protocols were generated by a digital shift register with an adder module feedback (16–18, 52). This process pseudo-randomized the changes in WR, and ensured that there would be sufficient $\dot{V}O_{2p}$ and WR signal amplitude at each of the harmonics of interest for frequency analysis. A single PRBS was composed of 15 units, each of 30 s in duration, totaling 7.5 min. Each complete PRBS testing session consisted of a 3.5 min warm-up (the last 3.5 min of the 7.5 min PRBS), and then two full repetitions of the PRBS for a total of 18.5 min of continuous cycling per session.

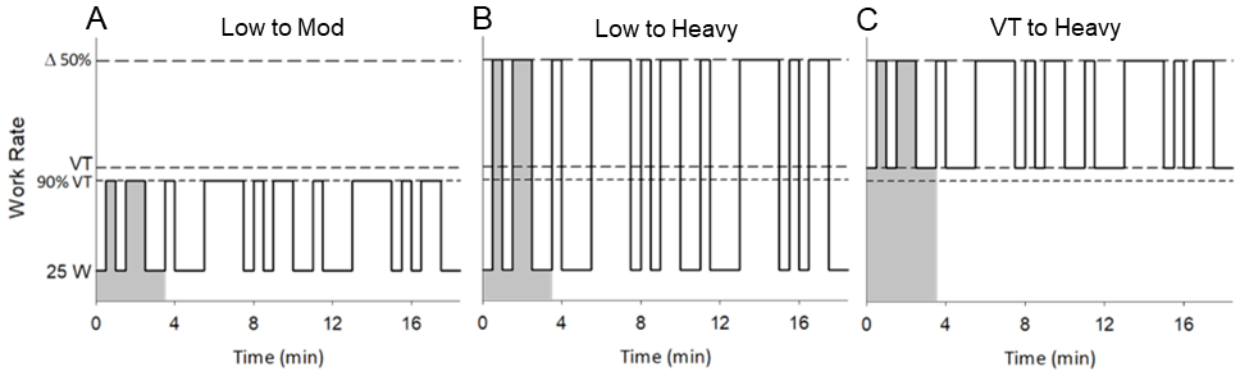

Fig. 8. Work rate (WR) profiles of the three different pseudorandom binary sequence exercise tests. WR alternates between: (A) 25 W and 90% of the ventilatory threshold (VT); (B) 25W and 50% of the difference between VT and peak $\dot{V}O_{2p}$ (Δ50%); (C) VT and Δ50%. The grey shaded region was designed as a warm up, and excluded from subsequent analyses *a priori*.

Data Collection

A portable metabolic system (MetaMax 3B-R2, CORTEX Biophysik, Leipzig, Germany) was used to measure gas exchange during all exercise tests. Participants breathed through a mask (7450 SeriesV2™ Mask, Hans Rudolph, Inc., Shawnee, KS, USA), and inspired and expired volumes were measured using a bi-directional volume turbine. The turbine was calibrated before each testing session using a 3 l syringe. Oxygen and carbon dioxide gas concentrations were continuously sampled at the mouth and were analyzed using a chemical fuel cell and nondispersive infrared sensor, respectively. Precision-analyzed gas mixtures were used to calibrate the oxygen and carbon dioxide gas concentrations. $\dot{V}O_{2p}$ and carbon dioxide output were calculated using standard breath-by-breath algorithms (80).

Participants wore a heart rate monitor (Polar H7, Polar Electro Oy, Kempele, Finland) that wirelessly communicated with the portable metabolic system, such that heart rate data were logged

synchronously with the gas exchange data. Participants were also fitted with a wearable integrated sensors shirt that was sized based on manufacturer guidelines (Hexoskin, Carré Technologies, Montreal, Canada). The shirt contained an electrocardiogram to measure heart rate, and thoracic and abdominal respiration bands in order to obtain estimates of breathing frequency, and minute ventilation ($\dot{V}_E$) via respiratory inductance plethysmography. Estimates of $\dot{V}_E$ provided by the sensor shirt were calibrated to known $\dot{V}_E$ measured using a volume turbine. The estimates of heart rate, $\dot{V}E$, and breathing frequency from the sensor shirt have been previously validated (78). Data recorded by the metabolic system and sensor shirt were time-aligned by cross-correlating the two different heart rate signals. $\dot{V}O_{2p}$ and wearable sensor data were interpolated to 1 Hz.

Machine learning

A $\dot{V}O_{2p}$ prediction algorithm was developed in this study to estimate $\dot{V}O_{2p}$ during both moderate and heavy intensity exercise. The model was based on a validated random forest regression algorithm (23) previously employed by Beltrame et al. to create an ambulatory moderate intensity $\dot{V}O_{2p}$ predictor (13). The candidate model features for this study were motivated by the model features used in Beltrame et al.'s $\dot{V}O_{2p}$ predictor: heart rate, delta heart rate (change in heart rate from previous second), $\dot{V}_E$, breathing frequency, hip acceleration, and walking cadence (13). In the present study, cycling exercise was used as the exercise stimulus instead of ambulatory exercise in order to have tighter control of WR during the heavy intensity exercise bouts. Accordingly, the ambulatory features of Beltrame et al.'s (13) $\dot{V}O_{2p}$ predictor were replaced by the cycling WR. Error and feature importance metrics were calculated on the learned regression trees that used the candidate features. Features with low importance were excluded from the model, resulting in the final set of features used to predict absolute $\dot{V}O_{2p}$: percent heart rate

reserve (%HRR), $\dot{V}_E$, and breathing frequency. %HRR was calculated for each participant by expressing the heart rate at a given point in time as a percentage, where 0% is equal to the resting heart rate obtained during the 5 min of seated rest before the ramp-incremental exercise test and 100% is equal to the maximum heart rate attained during the ramp-incremental test. %HRR was an advantageous model feature to use in comparison to heart rate, as it accounted for inter-individual differences in resting and maximum heart rate.

$\dot{V}O_{2p}$ predictor training was performed in MATLAB R2016a (TheMathWorks Inc., Natick, MA, USA). The random forest $\dot{V}O_{2p}$ predictor was grown as an ensemble of deep bootstrap aggregate regression trees (23) using wearable sensor and $\dot{V}O_{2p}$ data acquired during the ramp-incremental exercise and PRBS exercise tests from the 15 participants with complete datasets. The model was trained in a similar fashion as Beltrame et al. (13). Briefly, each individual regression tree was trained to estimate absolute $\dot{V}O_{2p}$ using a random subset of the training dataset, with one of the three features (%HRR, $\dot{V}_E$, breathing frequency) being randomly selected to be the splitting feature at each node. The splitting value for the feature at a given node was selected as the value that minimized the sum of square residuals of the two child nodes. This process was repeated for each node until each regression tree was completed. The final prediction from the random forest was calculated as the weighted average of each tree's prediction value given the three model inputs. The independent validation dataset composed of PRBS exercise responses from 7 other participants was used to determine the optimal number of regression trees. Independent random forest models with an increasing number of trees (1 to 500) were grown, and root mean square error (RMSE) of the prediction was used as criterion to identify a sufficient number of trees to include while minimizing computational cost. RMSE stabilized at ~0.26 $l \cdot min^{-1}$ by 100 trees (Fig. 9). Therefore, a random forest composed of 100 individual regression trees was used, as evaluating

more trees is computationally more expensive and does not improve prediction accuracy. Feature importance of the $\dot{V}O_{2p}$ predictor was estimated by summing the changes in the mean square error due to the splits on every predictor and dividing by the sum of the number of branch nodes.

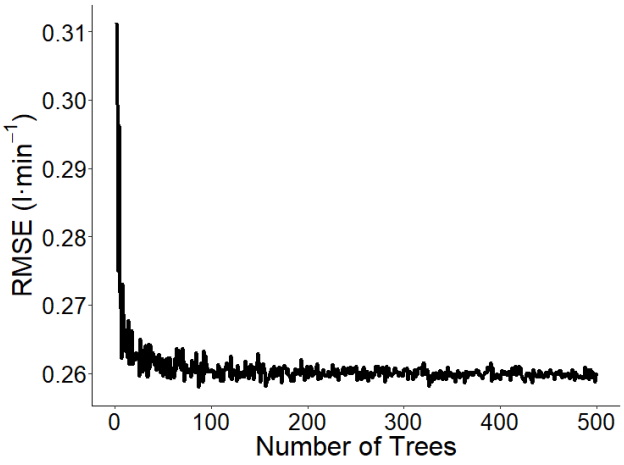

Fig. 9. Influence of number of regression trees included in the random forest model on root mean square error (RMSE) of the predicted $\dot{V}O_{2p}$ using a hold-out validation set.

The $\dot{V}O_{2p}$ predictor's accuracy for the low to mod, low to heavy and VT to heavy PRBSs was evaluated using leave-one-participant-out cross-validation against the directly measured $\dot{V}O_{2p}$ for each of the 15 participants. This method of validation prevented any overlap of participants between the training and testing datasets, which would otherwise influence prediction accuracy. The cross-validated predictions were also used to evaluate the ability of the model to accurately estimate $\dot{V}O_{2p}$ kinetics across different exercise intensities.

Oxygen uptake kinetics

For each measured and predicted PRBS response, the initial 3.5 min of warmup data were excluded *a priori*, and then the two following 7.5 min PRBS repetitions were ensemble averaged,

in order to yield a single 7.5 min $\dot{V}O_{2p}$ response for the each of the low to mod, low to heavy, and VT to heavy PRBS exercise conditions, respectively. The kinetics of measured and predicted $\dot{V}O_{2p}$ for the three different PRBS exercise tests were evaluated by frequency domain analysis, according to Beltrame and Hughson's MNG method (16). This method provided a single metric to summarize the responsiveness of $\dot{V}O_{2p}$ to changes in WR across a range of relevant harmonic frequencies.

MNG was calculated by converting the $\dot{V}O_{2p}$ and WR data into the frequency domain using the discrete Fourier transform. The following sinusoidal function for each harmonic, *h* (an integer), was solved:

$$y(t) = a_{DC} + 2 \cdot \sum_{h=1}^{4} \left(A_h \cdot cos(2\pi \cdot h \cdot f_1 \cdot t) + B_h \cdot sin(2\pi \cdot h \cdot f_1 \cdot t) \right) \quad (5)$$

where $y(t)$ is the signal at a given point t in time, a_{DC} is the system DC offset amplitude (average of the entire signal), f_1 is the fundamental frequency (1/ the duration of a single cycle of the PRBS (450 s)), and A_h and B_h are the cosine and sine amplitude coefficients for a given harmonic, respectively. A_h and B_h were used to calculate the total amplitude for each harmonic (AMP_h):

$$AMP_h = \sqrt{A_h^2 + B_h^2} \quad (6)$$

A system gain at each harmonic ($gAMP_h$) was then calculated using the AMP_h of the $\dot{V}O_{2p}$ and WR signals:

$$gAMP_h = \frac{\dot{V}O_{2p}AMP_h}{WR\ AMP_h} \quad (8)$$

The gain at each harmonic was normalized as a percentage of the fundamental harmonic gain ($gAMP_1$), and MNG was calculated by averaging the normalized gains of second, third and fourth harmonics:

$$MNG = \frac{\left(\frac{\sum_2^4 gAMP_h}{3} \right)}{gAMP_1} \cdot 100\% \quad (9)$$

The normalization of the gain at each harmonic used in the calculation of MNG was important, as it eliminated the effect of total gain across the harmonic amplitudes (47), which enabled MNG to be compared between individuals and exercise conditions. Furthermore, MNG was calculated using only the gains at the first four harmonics in order to avoid $\dot{V}O_{2p}$ nonlinearities that occur at higher frequencies due to circulatory distortion (47).

Statistical Analysis

Statistical analyses were conducted in R (Version 3.5.1). The accuracy of predicted $\dot{V}O_{2p}$ and MNG from the PRBS responses, and their agreement with the directly measured values were assessed using repeated measures correlation (5) (R Library: rmcorr), and repeated measures Bland-Altman analysis, which accounts for the within-participant variance of the repeated measures data (19), respectively. Pearson correlation co-efficient was also used to examine the relationships between predicted and measured MNG between participants within each of the three exercise conditions. Differences between predicted and directly measured MNG were evaluated using two-way repeated measures analysis of variance (ANOVA), with factors of Condition (low to mod, low to heavy, and VT to heavy) and Data type (predicted versus measured). RMSE was used to assess $\dot{V}O_{2p}$ prediction error for each exercise condition. One-way repeated measures ANOVA was then used to detect differences in RMSE between exercise conditions. For the ANOVAs, if Mauchly's test of sphericity was significant, then the Greenhouse-Geisser correction was used. Where significant main effects or an interaction were found, paired t-tests with Holm-Bonferroni corrected *p*-values were used for *post-hoc* analysis. Statistical significance was set to $p < 0.05$, and data are reported as means ± SD.

3.4 Results

Results presented reflect values of the 15 participants used in the leave-one-out cross-validation of the $\dot{V}O_{2p}$ predictor, and do not reflect any of the responses from the independent group of 7 participants which were used to determine the number of trees included in the model.

Ramp-incremental exercise responses. Group mean peak $\dot{V}O_{2p}$ was 2.98 ± 0.50 l·min^{-1}, which corresponded with a mean peak WR was 266 ± 32 W. The $\dot{V}O_{2p}$ at the estimated VT was 1.69 ± 0.17 l·min^{-1}, which occurred at a WR of 120 ± 12 W after accounting for the response lag time (48 ± 10 s). Estimates of WRs corresponding to 90% of VT and Δ50% were 108 ± 10 W and 185 ± 23 W, respectively.

Oxygen uptake prediction. The most important feature (predictor importance) of the $\dot{V}O_{2p}$ predictor was $\dot{V}_E$ (2.89 x 10^{-4} l^2·min^{-2}), followed by %HRR (1.75 x 10^{-4} l^2·min^{-2}), and breathing frequency (1.22 x 10^{-4} l^2·min^{-2}). Group mean predicted and measured $\dot{V}O_{2p}$ responses in low to mod, low to heavy and VT to heavy PRBS exercise are shown in Fig. 10. Fig. 11 shows the agreement results between the predicted and directly measured $\dot{V}O_{2p}$ when considering all three PRBS exercise tests together for each of the 15 participants. Predicted and measured $\dot{V}O_{2p}$ were strongly correlated (r_{rm} = 0.911, p < 0.001), and low systematic bias was observed (0.022 l·min^{-1}) with the distribution of predicted-measured differences centered near zero. A significant effect of Condition was observed on $\dot{V}O_{2p}$ prediction error ($F_{(2,28)}$ = 13.645, p < 0.001). RMSE, expressed relative to body mass in order to facilitate comparison with the American College of Sport Medicine's (ACSM) standard for accurate $\dot{V}O_{2p}$ prediction (RMSE < 5 ml·min^{-1}·kg^{-1}) (40), was lowest in the low to mod condition (2.64 ± 0.76 ml·min^{-1}·kg^{-1}), but was not different between the low to heavy (3.95 ± 0.97 ml·min^{-1}·kg^{-1}) and VT to heavy (3.90 ± 1.26 ml·min^{-1}·kg^{-1}) conditions.

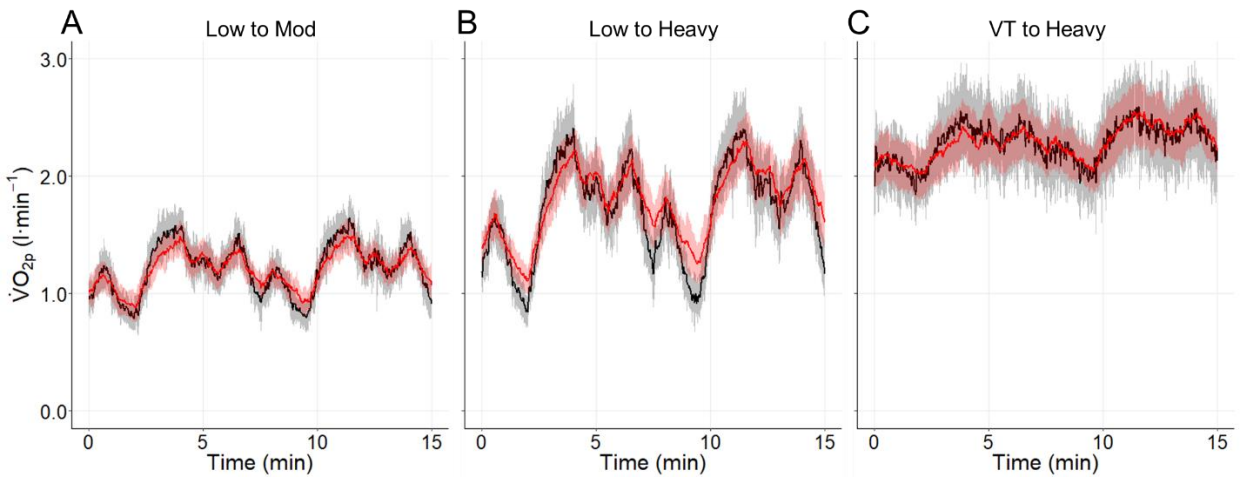

Fig. 10. Group mean ± SD measured (black) and predicted (red) oxygen uptake ($\dot{V}O_{2p}$) responses to pseudorandom binary sequence exercise in low to mod (A), low to heavy (B), and ventilatory threshold (VT) to heavy (C) conditions following 3.5 minutes of warm-up exercise (not shown).

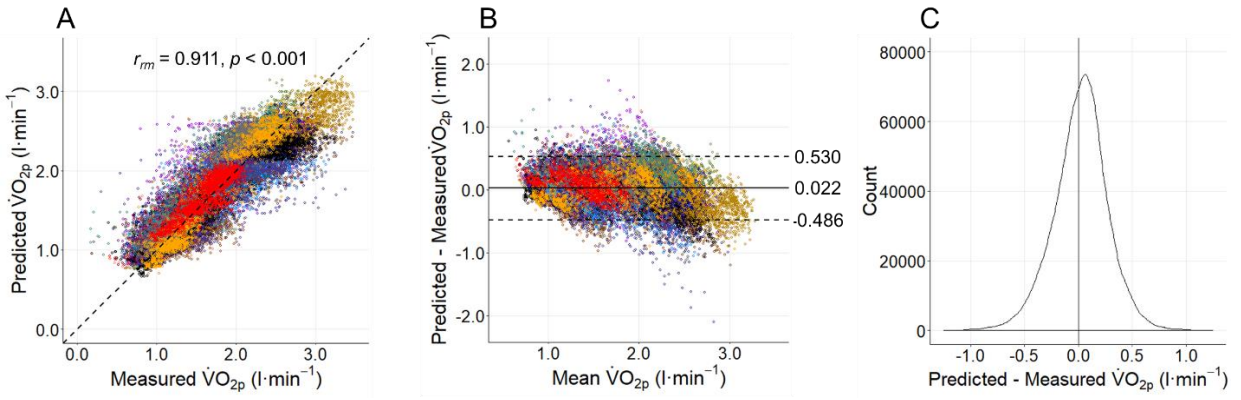

Fig. 11. Accuracy of the proposed oxygen uptake predictor with all three exercise conditions combined. A: repeated measures correlation of predicted and measured oxygen uptake ($\dot{V}O_{2p}$). Parallel regression lines for each participant were omitted for clarity. Dotted line represents the identity line. B: repeated measures Bland-Altman plot. C: distribution of the differences between the predicted and measured $\dot{V}O_{2p}$.

Oxygen uptake kinetics prediction. MNG calculated from predicted $\dot{V}O_{2p}$ differed from that of the measured $\dot{V}O_{2p}$ (Fig. 12). Main effects of Condition ($F_{(1.350, 18.907)}$ = 23.473, p < 0.001) and Data type ($F_{(1,14)}$ = 29.526, p < 0.001) were observed in addition to a Condition x Data type interaction ($F_{(1.369, 19.167)}$ = 5.339, p = 0.023). *Post-hoc* analysis revealed that the measured MNG was highest in the low to mod condition, then low to heavy, and lowest in the VT to heavy

condition. The low to mod predicted MNG was higher than the low to heavy (p = 0.037), but not the VT to heavy (p = 0.225) MNG. The predicted MNG was not different between the low to heavy and VT to heavy conditions. Predicted MNG was lower than the measured MNG in the low to mod (p < 0.001) and low to heavy (p < 0.001) conditions, but not in the VT to heavy condition (p = 0.584). Repeated measures Bland-Altman analysis confirmed that the predicted MNG was systematically lower than the measured MNG (bias: -9.7%; Fig. 13). Despite the negative bias of the predicted MNG, and the observed Condition x Data type interaction, a moderate positive repeated measures correlation was observed between the predicted and measured MNG (r_{rm} = 0.413, p = 0.021; Fig. 14). Negative bias for the predicted MNG is due to significant negative relationships (p < 0.001) between the predicted-measured $\dot{V}O_{2p}$ differences and the mean $\dot{V}O_{2p}$ in the repeated measures Bland-Altman plots when considering each exercise conditions individually (Fig. 15). No significant (p > 0.05) between-participant correlations were observed for MNG when considering each exercise condition independently.

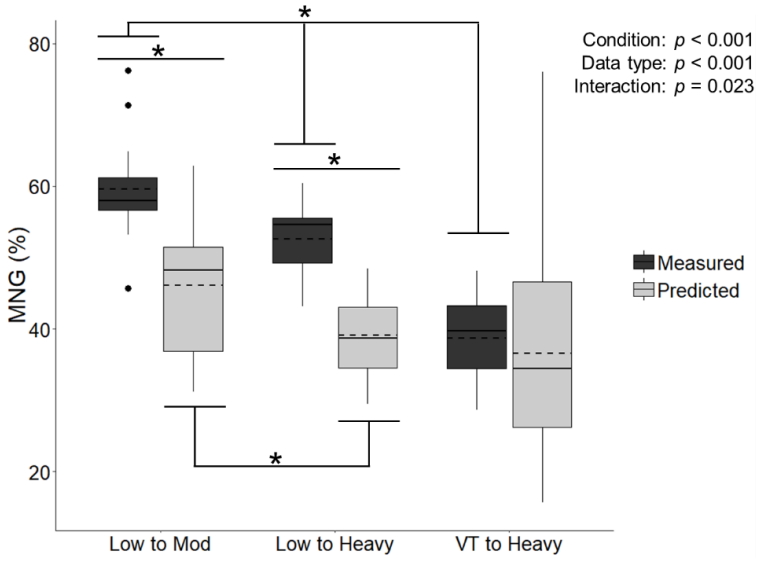

Fig. 12. Comparison of mean normalized gain (MNG) calculated from predicted and measured oxygen uptake data across exercise conditions. * significant (p < 0.05) difference between points. Dashed and solid horizontal lines in box plots are mean and median, respectively.

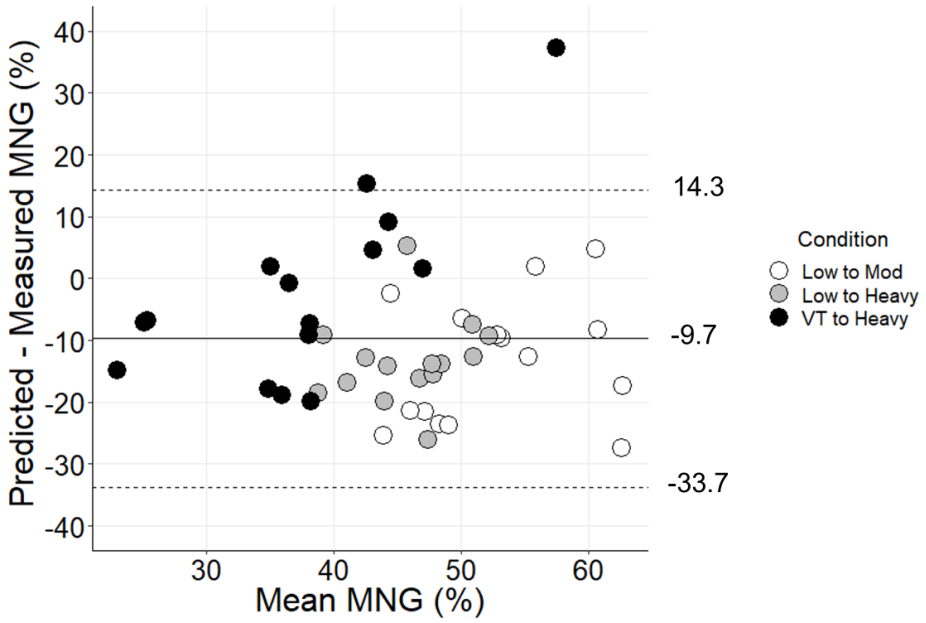

Fig. 13. Repeated measures Bland-Altman plot of the mean normalized gain (MNG) calculated using the predicted and measured oxygen uptake data from the low to moderate (low to mod; white), low to heavy (grey), and ventilatory threshold to heavy (VT to heavy; black) exercise conditions.

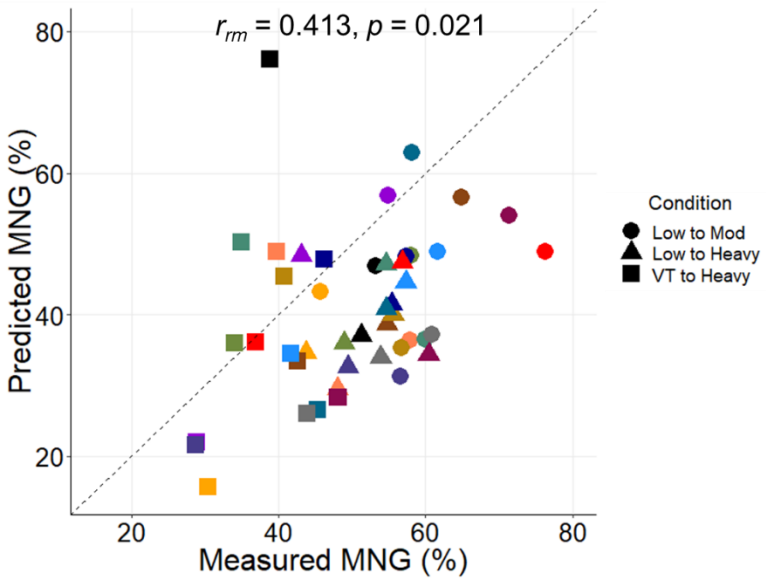

Fig. 14. Repeated measures correlation between the mean normalized gain (MNG) calculated from the predicted oxygen uptake ($\dot{V}O_{2p}$) and the MNG calculated from the directly measured $\dot{V}O_{2p}$. All participants' (n = 15) measurements from low to moderate (circles), low to heavy (triangles), and ventilatory threshold (VT) to heavy (squares) exercise conditions are shown. A different color represents each participant. Parallel regression lines for each participant were omitted for clarity. Dotted line represents the identity line.

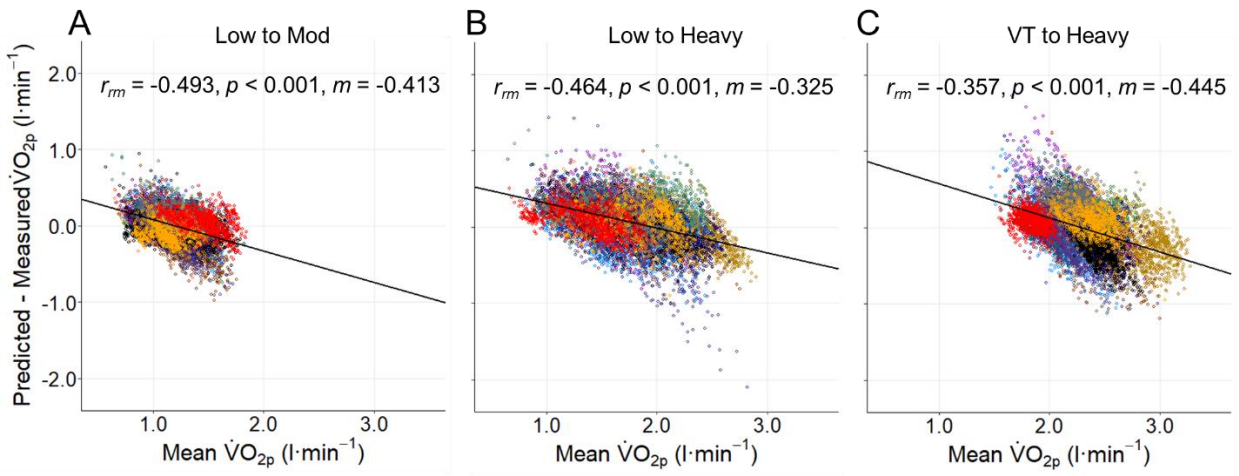

Fig. 15. Repeated measures Bland-Altman plots for oxygen uptake ($\dot{V}O_{2p}$) predictions from low to mod (A), low to heavy (B), and ventilatory threshold (VT) to heavy (C) exercise conditions. A reference line is shown for each condition representing the common regression slope from repeated measures correlation analysis of the data in each plot. Strength of the relationship (r_{rm}) and slope (m) are shown for each individual condition.

3.5 Discussion

The purpose of this study was to develop a $\dot{V}O_{2p}$ predictor that uses data from non-intrusive wearable sensors to accurately estimate dynamic responses to both moderate and heavy intensity exercise. In support of the hypothesis, $\dot{V}O_{2p}$ was able to be predicted across moderate and heavy exercise with low bias, and was strongly correlated with the directly measured $\dot{V}O_{2p}$. The $\dot{V}O_{2p}$ predictor generated responses with slower kinetics during low to heavy compared to low to mod exercise; however the low to mod and low to heavy predicted kinetics were not different from the predicted kinetics in the VT to heavy condition, which was due to large variance in the predicted VT to heavy kinetics. Overall, the predicted $\dot{V}O_{2p}$ kinetics were only moderately correlated with the measured response kinetics within an individual, and were biased to slower kinetics.

Machine learning techniques are powerful tools used to analyze and examine relationships with large amounts of complex data. A random forest regression algorithm that uses wearable

sensor inputs was proposed by Beltrame et al. to be a good candidate for real-time $\dot{V}O_{2p}$ prediction, as it is an ensemble model that generalizes well to different types of data distributions, it is able to generate predictions quickly, and it is memory efficient (13). Beltrame et al. found that this machine learning algorithm was accurate at estimating dynamic changes in $\dot{V}O_{2p}$, but the generalizability of their findings was limited to low to moderate intensities of exercise, as their training and validation data were restricted to below ~ 6 metabolic equivalents (METS; $\dot{V}O_{2p}$ in $ml \cdot min^{-1} \cdot kg^{-1}$ divided by 3.5). Therefore, our aim was to develop and test a model with a similar architecture to Beltrame et al. (13) in order to determine if a random forest model with easy-to-obtain inputs could accurately estimate $\dot{V}O_{2p}$ during both moderate and heavy intensity exercise. In the present study, a random forest algorithm was trained and evaluated via leave-one-participant-out cross-validation using 15 participants' moderate and heavy intensity exercise data. The number of regression trees included in the model was optimized using an independent validation dataset comprised of a separate set of 7 participants' data. The new random forest regression algorithm was able to accurately predict $\dot{V}O_{2p}$ over a much wider range of values (up to ~ 12 METS), as evidenced by the low prediction bias and strong repeated measures correlation between predicted and measured $\dot{V}O_{2p}$. The prediction bias and limits of agreement of the model (bias: 0.419 $ml \cdot min^{-1} \cdot kg^{-1}$, [-6.805, 7.643]) were similar to those reported by Beltrame et al. during activities of daily living (bias: -0.294 $ml \cdot min^{-1} \cdot kg^{-1}$, [-6.460, 5.872]) and walking exercise (bias: 0.259 $ml \cdot min^{-1} \cdot kg^{-1}$, [-3.991, 4.509]) (13). The present model's RMSE was lowest in the low to mod exercise condition, indicating that the predictions were most accurate in the moderate intensity domain. Prediction RMSE in both conditions containing heavy exercise was within ACSM's acceptable range (< 5 $ml \cdot min^{-1} \cdot kg^{-1}$) for obtaining useful estimates of peak $\dot{V}O_{2p}$ (40). Importantly, some participants' measured and predicted $\dot{V}O_{2p}$ approached peak levels during the

VT to heavy exercise condition, which suggests that this random forest algorithm might be useful for accurately and non-intrusively estimating cardiorespiratory fitness from information provided by heart rate and respiratory bands during physical assessments in field settings.

The MNG method was used to assess the measured and predicted $\dot{V}O_{2p}$ kinetics. Group mean measured MNG behaved as expected, such that MNG was highest in the low to mod condition, then low to heavy, and was lowest in the VT to heavy condition, indicating the slowest kinetics. These findings of slowed kinetics with heavy compared to moderate intensity exercise are consistent with the findings from step (32, 36, 51, 53, 61, 65, 68), and sinusoidal exercise (42). Predicted MNG in the low to mod condition was higher than in the low to heavy condition, which reflects faster kinetics during moderate compared to heavy intensity exercise; however, no differences in predicted MNG for either condition were detected in comparison to the VT to heavy condition. The lack of a statistical differences in predicted MNG between these conditions are due to increased variability of the predicted dynamics during the VT to heavy condition. The increased variability may be a result of smaller changes in model input features throughout VT to heavy PRBS exercise compared to the other conditions. This would enable signal noise to have a greater effect on the predicted kinetics.

A moderate positive repeated measures correlation between the predicted and measured MNG was observed, which indicates that there is a significant within-participant relationship. However, no significant between-participant correlations were identified between predicted and measured MNG within each condition individually. Overall, this suggests that the $\dot{V}O_{2p}$ predictor is able to estimate some differences in kinetics between exercise conditions within the same individual, but it is not precise enough to detect between-participant differences within a given exercise condition. The lack of significant between-participant correlations for each exercise

condition individually is not believed to be a result of the form of kinetic analysis performed, as MNG has been previously correlated with the standard time domain indicator of kinetics (τ) between-individuals (16).

Two-way repeated measures ANOVA results in combination with the repeated measures Bland-Altman analysis suggest that the predicted MNG systematically underestimated the measured MNG, which indicates that the predicted $\dot{V}O_{2p}$ is created with slower kinetics than the true kinetic response. This can be visually confirmed in Fig. 3, where the rate of change of the predicted response (red line) is dampened compared to the measured response (black line) in all conditions. The bias observed in this study (~10% lower MNG) was opposite to the bias observed by Beltrame et al. during pseudorandom ternary sequence walking, where the bias was to ~6% higher (13). Beltrame et al. also reported a bias to 16% higher MNG when comparing predicted MNG from activities of daily living to directly measured MNG during pseudorandom ternary sequence walking (14). Differences in bias between the present $\dot{V}O_{2p}$ predictor, and the previous model by Beltrame et al. (13) may be a result of differences in input features, although differential effects of exercise modality (walking versus cycling) cannot be ruled out. Perhaps the cadence and hip acceleration inputs to their model provide the higher frequency stimulus to overestimate MNG, whereas the WR signal was not included as an input into our final $\dot{V}O_{2p}$ predictor due to abrupt square-wave-like changes in predicted $\dot{V}O_{2p}$ that it produced at the onset of transitions in WR.

The bias to estimating slower kinetics in the current model may be due to a limitation of random forest models. Because the random forest algorithm does not account for longitudinal patterns in the relationship between %HRR, $\dot{V}_E$, and breathing frequency, it is not able to adjust predictions based on how the input-output relationship changes as a result of changes in the underlying physiology. For example, the relationship between $\dot{V}O_{2p}$ and heart rate kinetics differ

between on- and off-transients, which is further amplified during the heavy intensity exercise conditions (9). Furthermore, it is known that both $\dot{V}_E$ (49) and heart rate kinetics (9), which are the two most important inputs for this model, are appreciably slower than $\dot{V}O_{2p}$ kinetics when performing repeated exercise transitions. Accordingly, the predicted $\dot{V}O_{2p}$ kinetics may actually more closely reflect the time course of changes in heart rate or $\dot{V}_E$, than $\dot{V}O_{2p}$. Visually, the curve characteristics appear to match more closely between the predicted $\dot{V}O_{2p}$ and the %HRR for a representative participant, than the measured $\dot{V}O_{2p}$ (Fig. 16). It is this dissociation between the true and predicted $\dot{V}O_{2p}$ dynamics that is responsible for the negative slopes in the repeated measures Bland-Altman plots for each respective exercise intensity (Fig. 15). Perhaps the random forest model developed for low to moderate intensity exercise by Beltrame et al. (13) was able to more accurately predict $\dot{V}O_{2p}$ kinetics than the present model because it was only trained and tested in this lower intensity domain, where the input-output relationship between the sensor data and $\dot{V}O_{2p}$ is more linear and tightly coupled. Moreover, the added complexity introduced into the present model by including heavy exercise responses in the training data is likely what compromised our ability to replicate Beltrame et al.'s (13) findings of a strong between-participant correlation for the predicted and measured MNG in the low to mod exercise condition.

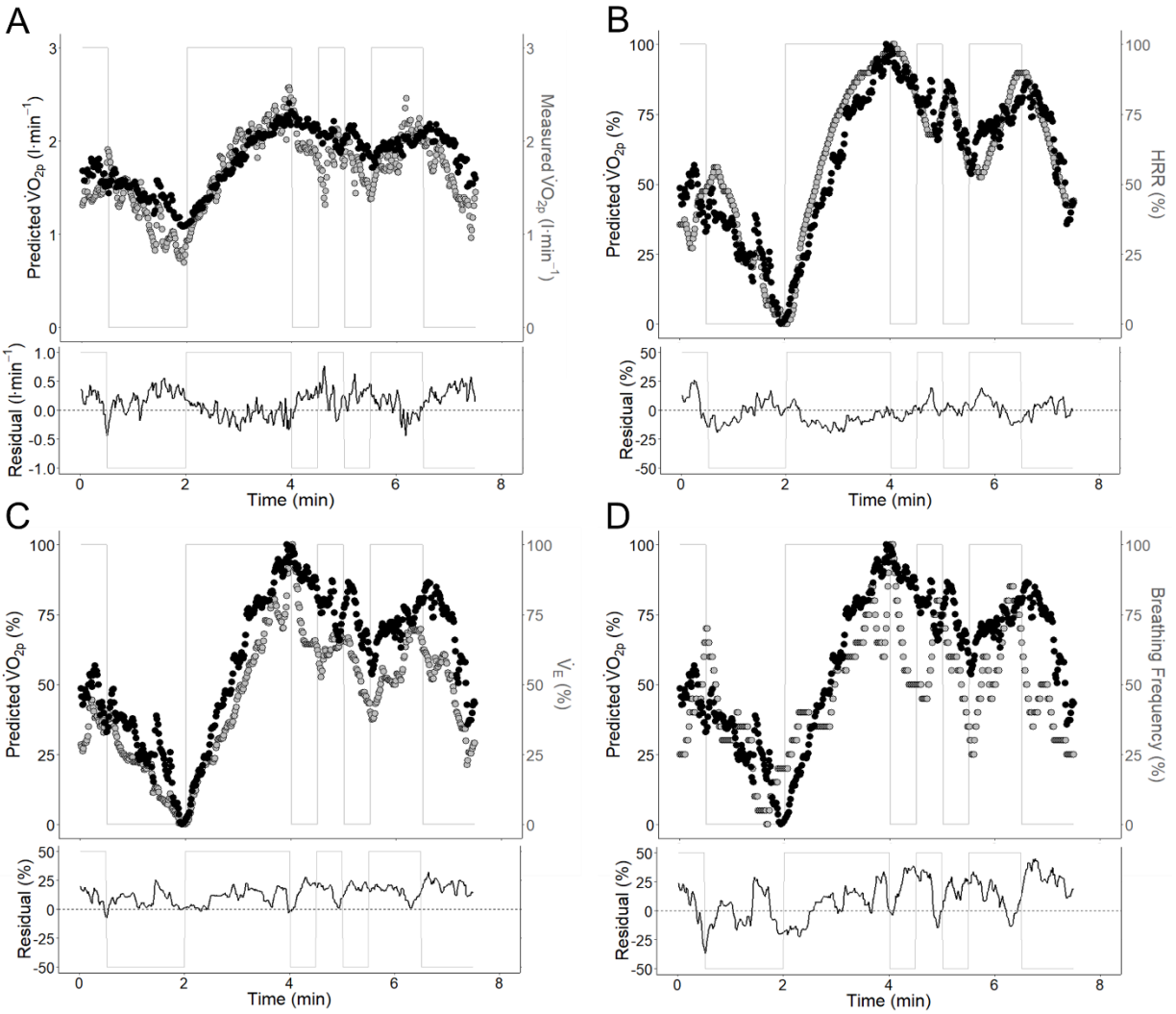

Fig. 16. Time-series comparisons between predicted pulmonary oxygen uptake ($\dot{V}O_{2p}$) (black circle), and measured $\dot{V}O_{2p}$ or model input features (grey circle) during one repetition of a representative participant's low to heavy pseudorandom binary sequence exercise. The time-series for changes in work rate are depicted by the light grey line in each panel. Residuals calculated as the difference between the predicted $\dot{V}O_{2p}$ and measured $\dot{V}O_{2p}$ or each respective model feature are displayed below each graph. Comparison between predicted and measured $\dot{V}O_{2p}$ were performed using absolute values (A). Comparisons between predicted $\dot{V}O_{2p}$ and percent heart rate reserve (HRR; B), estimated minute ventilation ($\dot{V}_E$; C), and breathing frequency (D) were performed using normalized responses as a percentage of the range of observed values during the depicted pseudorandom binary sequence. Dynamics of the predicted $\dot{V}O_{2p}$ appear slower than the measured $\dot{V}O_{2p}$, and match closely with the HRR.

A more complex model is needed to fulfil the goal of being able to accurately characterize $\dot{V}O_{2p}$ response dynamics outside of the laboratory setting across a wide range of exercise

intensities as an indicator of health (1, 20, 25, 72), fitness (29, 41, 43, 69), or performance (26), and not just accurately estimate static gain. The integration of previous predictions into the model, or using longitudinal predictor data inputs, will likely enable the identification of time by exercise intensity interactions. This should improve the characterization of the complex and dynamic relationships between cardiac and respiratory sensor data and $\dot{V}O_{2p}$, leading to a more powerful $\dot{V}O_{2p}$ predictor during heavy exercise. Zignoli et al. (89) recently demonstrated a strong association between measured and predicted $\dot{V}O_{2p}$ values from a long short-term memory recurrent neural network during high intensity exercise; however, they did not evaluate the accuracy of the predicted $\dot{V}O_{2p}$ kinetics, and their prediction generalizability to out-of-sample groups remain unclear since their machine learning model was trained and tested on the same set of participants. Therefore, long short-term memory recurrent neural networks may be a promising model for prediction of heavy intensity $\dot{V}O_{2p}$, but further research is required to evaluate the accuracy of the predicted kinetics.

Limitations. An important consideration for the $\dot{V}O_{2p}$ predictor is the trade-off between being easy to use by the general population, and being accurate. In the present study, percent heart rate reserve was used as a model feature instead of heart rate, and estimated $\dot{V}_E$ from the smart shirt was calibrated to values obtained with a volume turbine. Normalizing the heart rate response and calibrating the $\dot{V}_E$ for each individual helped to improve the accuracy of the $\dot{V}O_{2p}$ prediction, but was at the cost of being a fully self-sufficient $\dot{V}O_{2p}$ prediction system, as a maximum effort test was needed to determine each individual's maximum heart rate, and additional equipment was required to calibrate the estimated $\dot{V}_E$. Normalization of heart rate responses and calibration of estimated $\dot{V}_E$ may affect the feasibility of $\dot{V}O_{2p}$ prediction outside of the laboratory, as individual-specific information and additional equipment is needed.

In conclusion, a random forest regression algorithm that uses easy-to-obtain wearable sensor inputs was developed to predict $\dot{V}O_{2p}$ during moderate and heavy intensity exercise. The model was able to predict $\dot{V}O_{2p}$ responses across exercise intensities with low bias, and was strongly correlated with directly measured $\dot{V}O_{2p}$. However, the indicator of predicted response kinetics was only moderately correlated with the measured value, and was biased to reflect slower kinetics. This suggests that the previously validated random forest regression model that accurately predicted $\dot{V}O_{2p}$ kinetics during low to moderate exercise cannot be easily generalized to include heavy intensity exercise, where response dynamics are more complex. Therefore, new models should be explored to improve the characterization of $\dot{V}O_{2p}$ kinetics for heavy intensity exercise.

Chapter 4: General Discussion

4.1 Summary of Findings

Accurate prediction of $\dot{V}O_{2p}$ and extraction of an indicator of kinetics from non-constant WR exercise will enable assessment of aerobic system responsiveness outside of the confines of a laboratory environment. Previously, Beltrame and Hughson developed a novel frequency domain indicator of kinetics, MNG, for evaluating $\dot{V}O_{2p}$ kinetics during random activities (16). The potential of this frequency analysis technique was realized when Beltrame et al. used it in conjunction with a non-intrusive $\dot{V}O_{2p}$ predictor (13) to evaluate $\dot{V}O_{2p}$ kinetics during unsupervised activities of daily living (14). However, these foundational studies for non-intrusive monitoring of $\dot{V}O_{2p}$ kinetics outside of the laboratory were limited by the fact that all exercise intensities used in the development of the $\dot{V}O_{2p}$ predictor, and validation of MNG, were limited to low to moderate exercise intensities. Therefore, the purpose of this thesis was to investigate if a model based on the aforementioned $\dot{V}O_{2p}$ prediction algorithm (13), and frequency domain kinetic analysis technique (16), are appropriate for heavy intensity exercise, where exercise response dynamics are more complex.

In Chapter 2, Beltrame and Hughson's (16) novel frequency domain indicator of kinetics, MNG, was evaluated for its ability to detect differences in $\dot{V}O_{2p}$ kinetics between pseudorandom moderate and heavy intensity exercise. The main finding of Chapter 2 was that MNG was able to detect the progressive slowing of $\dot{V}O_{2p}$ kinetics from moderate to heavy intensity exercise. Importantly, MNG across the moderate and heavy intensity exercise bouts was strongly correlated with the conventional time domain indicator of kinetics, phase II τ, which reinforces that MNG provides an accurate assessment of $\dot{V}O_{2p}$ kinetics across the moderate and heavy intensity exercise domains. The implications of these findings are that both moderate and heavy $\dot{V}O_{2p}$ kinetics can

be assessed using a frequency domain analysis technique, which is applicable to random changes in intensity that are experienced in real-life situations.

In Chapter 3, a model based on Beltrame et al.'s (13) previously validated moderate intensity $\dot{V}O_{2p}$ predictor was developed to predict $\dot{V}O_{2p}$ during both moderate and heavy intensity exercise. A random forest regression algorithm was trained and evaluated for its ability to accurately predict dynamic $\dot{V}O_{2p}$ responses from easy-to-obtain wearable sensor data. The random forest model was able to accurately estimate absolute $\dot{V}O_{2p}$ across a wide range of intensities, but the kinetics of the predicted response, which were evaluated using the validated MNG method, were systematically slower than the directly measured responses. Furthermore, the predicted and measured MNG were only moderately correlated to each other for within-participant comparisons. This suggests that Beltrame et al.'s (13) validated moderate intensity $\dot{V}O_{2p}$ prediction algorithm cannot be easily generalized to be able to predict both moderate and heavy intensity dynamic exercise responses due to the added complexity of the heavy intensity exercise responses. This study argues for the development of a new single model to predict both moderate and heavy intensity exercise, as it demonstrates that the random forest model is not effective at predicting heavy intensity $\dot{V}O_{2p}$ kinetics, and that other machine learning models that can better characterize time and intensity sensitive relationships between wearable sensor inputs and $\dot{V}O_{2p}$ should be explored to develop a more powerful prediction algorithm.

4.2 Limitations

In this thesis, all studies were conducted using a young healthy sample population of men and women with close to average fitness levels (Chapter 2: 43 ± 6 $ml \cdot min^{-1} \cdot kg^{-1}$, Chapter 3: 42 ± 5 $ml \cdot min^{-1} \cdot kg^{-1}$) for their respective ages (55). All tests for evaluating MNG and the $\dot{V}O_{2p}$

prediction algorithms were PRBS exercise performed on a cycle ergometer. Therefore, studies should be conducted to evaluate MNG as an indicator of heavy intensity kinetics during other modalities of intermittent exercise, such as running, where the activity level will have to be approximated by accelerometers (86). Furthermore, once a $\dot{V}O_{2p}$ predictor is developed to accurately model heavy intensity kinetics using non-intrusive wearable sensor data, the $\dot{V}O_{2p}$ predictor and MNG should be tested for their ability to predict $\dot{V}O_{2p}$ and extract an indicator of kinetics from heavy intensity exercise in the field setting and during intermittent high intensity sports.

4.3 Future Directions

An important next step in this research is the development of a predictive algorithm that more accurately characterizes $\dot{V}O_{2p}$ kinetics. Future work should investigate the applicability of deep learning to this time-series prediction problem, as it can encode and learn temporal dynamics directly from the data. Sequence models, such as recurrent or temporal convolutional neural networks, appear to be strong candidates for $\dot{V}O_{2p}$ prediction as a result of their ability to model dependencies in sequential data. This is an important consideration, as the time course and amplitude of the changes in $\dot{V}O_{2p}$, heart rate, $\dot{V}_E$, and breathing frequency are all highly dependent on the duration and intensity of the current and previous activity. Recurrent and temporal convolutional neural networks both incorporate information about the previous state of the system into their predictions, albeit in different ways. Long short-term memory networks, which are a special class of recurrent neural networks, are able to store the dependency of sequential input data in their memory cells and access this information when needed at later time points (44), whereas temporal convolutional neural networks use longitudinal input data to make

a prediction for a given time point (3). These sequential models will likely improve estimates of dynamic $\dot{V}O_{2p}$ responses by more accurately learning the relationships between the wearable sensor data and $\dot{V}O_{2p}$ as a function of exercise intensity and duration.

4.4 Conclusion

Frequency domain analysis is a viable method to extract an indicator of $\dot{V}O_{2p}$ kinetics from non-constant WR moderate and heavy intensity exercise, but the previously validated random forest architecture for $\dot{V}O_{2p}$ prediction during moderate intensity exercise (13) is not easily generalizable to include heavy intensity exercise, where the physiological response is more complex. Therefore, new machine learning models that account for time- and exercise intensity-dependent interactions between wearable sensor data and $\dot{V}O_{2p}$ should be explored in order to improve the prediction of heavy dynamic $\dot{V}O_{2p}$ responses, and to unlock the full potential of frequency domain analysis as a tool for field evaluation of $\dot{V}O_{2p}$ kinetics across a wide range of intensities.

References

1. **Alexander NB**, **Dengel DR**, **Olson RJ**, **Krajewski KM**. Oxygen-uptake (VO2) kinetics and functional mobility performance in impaired older adults. *J Gerontol* 58: 734–739, 2003.

2. **Altini M**, **Penders J**, **Amft O**. Estimating oxygen uptake during nonsteady-state activities and transitions using wearable sensors. *IEEE J Biomed Heal Informatics* 20: 469–475, 2016.

3. **Bai S**, **Kolter JZ**, **Koltun V**. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv:* 1803.01271, 2018.

4. **Bailey SJ**, **Wilkerson DP**, **Dimenna FJ**, **Jones AM**. Influence of repeated sprint training on pulmonary O2 uptake and muscle deoxygenation kinetics in humans. *J Appl Physiol* 106: 1875–1887, 2009.

5. **Bakdash JZ**, **Marusich LR**. Repeated measures correlation. *Front Psychol* 8: 1–13, 2017.

6. **Barbero-Alvarez JC**, **Soto VM**, **Barbero-Alvarez V**, **Granda-Vera J**. Match analysis and heart rate of futsal players during competition. *J Sports Sci* 26: 63–73, 2008.

7. **Barstow TJ**, **Casaburi R**, **Wasserman K**. O2 uptake kinetics and the O2 deficit as related to exercise intensity and blood lactate. *J Appl Physiol* 75: 755–762, 1993.

8. **Barstow TJ**, **Mole PA**. Linear and nonlinear characteristics of oxygen uptake kinetics during heavy exercise. *J Appl Physiol* 71: 2099–2106, 1991.

9. **Bearden SE**, **Moffatt RJ**. VO2 and heart rate kinetics in cycling: transitions from an elevated baseline. *J Appl Physiol* 90: 2081–2087, 2001.

10. **Beaver WL**, **Wasserman K**, **Whipp BJ**. A new method for detecting threshold by gas exchange anaerobic. *J Appl Physiol* 60: 2020–2027, 1986.

11. **Bell C**, **Paterson DH**, **Kowalchuk JM**, **Padilla J**, **Cunningham DA**. A comparison of modelling techniques used to characterise oxygen uptake kinetics during the on-transient of exercise. *Exp Physiol* 86: 667–676, 2001.

12. **Beltrame T**, **Amelard R**, **Villar R**, **Shafiee MJ**, **Wong A**, **Hughson RL**. Estimating oxygen uptake and energy expenditure during treadmill walking by neural network analysis of easy-to-obtain inputs. *J Appl Physiol* 121: 1226–1233, 2016.

13. **Beltrame T**, **Amelard R**, **Wong A**, **Hughson RL**. Prediction of oxygen uptake dynamics by machine learning analysis of wearable sensors during activities of daily living. *Sci Rep* 7: 45738, 2017.

14. **Beltrame T**, **Amelard R**, **Wong A**, **Hughson RL**. Extracting aerobic system dynamics during unsupervised activities of daily living using wearable sensor machine learning models. *J Appl Physiol* 124: 473–481, 2018.

15. **Beltrame T**, **Hughson RL**. Aerobic system analysis based on oxygen uptake and hip

acceleration during random over-ground walking activities. *Am J Physiol - Regul Integr Comp Physiol* 312: R93–R100, 2017.

16. **Beltrame T**, **Hughson RL**. Mean normalized gain: a new method for the assessment of the aerobic system temporal dynamics during randomly varying exercise in humans. *Front Physiol* 8: 1–13, 2017.

17. **Beltrame T**, **Hughson RL**. Linear and non-linear contributions to oxygen transport and utilization during moderate random exercise in humans. *Exp Physiol* 102: 563–577, 2017.

18. **Bennett FM**, **Reischl P**, **Grodins FS**, **Yamashiro SM**, **Fordyce WE**. Dynamics of ventilatory response to exercise in humans. *J Appl Physiol Respir Environ Exerc Physiol* 51: 194–203, 1981.

19. **Bland JM**, **Altman DG**. Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat* 17: 571–582, 2007.

20. **Borghi-Silva A**, **Beltrame T**, **Reis MS**, **Sampaio LMM**, **Catai AM**, **Arena R**, **Costa D**. Relationship between oxygen consumption kinetics and BODE Index in COPD patients. *Int J COPD* 7: 711–718, 2012.

21. **Bowen TS**, **Murgatroyd SR**, **Cannon DT**, **Cuff TJ**, **Lainey AF**, **Marjerrison AD**, **Spencer MD**, **Benson AP**, **Paterson DH**, **Kowalchuk JM**, **Rossiter HB**. A raised metabolic rate slows pulmonary O2 uptake kinetics on transition to moderate-intensity exercise in humans independently of work rate. *Exp Physiol* 96: 1049–1061, 2011.

22. **Bradley PS**, **Sheldon W**, **Wooster B**, **Olsen P**, **Boanas P**, **Krustrup P**. High-intensity running in English FA Premier League soccer matches. *J Sports Sci* 27: 159–168, 2009.

23. **Breiman L**. Random forests. *Mach Learn* 45: 5–32, 2001.

24. **Brittain CJ**, **Rossiter HB**, **Kowalchuk JM**, **Whipp BJ**. Effect of prior metabolic rate on the kinetics of oxygen uptake during moderate-intensity exercise. *Eur J Appl Physiol* 86: 125–134, 2001.

25. **Brunner-La Rocca HP**, **Weilenmann D**, **Schalcher C**, **Schlumpf M**, **Follath F**, **Candinas R**, **Kiowski W**. Prognostic significance of oxygen uptake kinetics during low level exercise in patients with heart failure. *Am J Cardiol* 84: 741–744, 1999.

26. **Burnley M**, **Jones AM**. Oxygen uptake kinetics as a determinant of sports performance. *Eur J Sport Sci* 7: 63–79, 2007.

27. **Casaburi R**, **Daly J**, **Hansen JE**, **Effros RM**. Abrupt changes in mixed venous blood gas composition after the onset of exercise. *J Appl Physiol* 67: 1106–1112, 1989.

28. **Casaburi R**, **Whipp BJ**, **Wasserman K**, **Beaver WL**, **Koyal SN**. Ventilatory and gas exchange dynamics in response to sinusoidal work. *J Appl Physiol Respir Environ Exerc Physiol* 42: 300–311, 1977.

29. **Chilibeck PD**, **Paterson DH**, **Petrella RJ**, **Cunningham A**. The influence of age and cardiorespiratory fitness on kinetics of oxygen uptake. *Can J Appl Physiol* 21: 185–196, 1996.

30. **Coutts AJ**, **Quinn J**, **Hocking J**, **Castagna C**, **Rampinini E**. Match running performance in elite Australian Rules Football. *J Sci Med Sport* 13: 543–548, 2010.

31. **Cross TJ**, **Winters C**, **Sheel AW**, **Sabapathy S**. Respiratory muscle power and the slow component of O2 uptake. *Med Sci Sports Exerc* 46: 1797–1807, 2014.

32. **Dimenna FJ**, **Bailey SJ**, **Vanhatalo A**, **Chidnok W**, **Jones AM**. Elevated baseline VO2 per se does not slow O2 uptake kinetics during work-to-work exercise transitions. *J Appl Physiol* 109: 1148–1154, 2010.

33. **DiMenna FJ**, **Jones AM**. "Linear" versus "nonlinear" VO2 responses to exercise: reshaping traditional beliefs. *J Exerc Sci Fit* 7: 67–84, 2009.

34. **Dimenna FJ**, **Wilkerson DP**, **Burnley M**, **Bailey SJ**, **Jones AM**. Influence of priming exercise on pulmonary O2 uptake kinetics during transitions to high-intensity exercise at extreme pedal rates. *J Appl Physiol* 106: 432–442, 2009.

35. **Dimenna FJ**, **Wilkerson DP**, **Burnley M**, **Jones AM**. Influence of priming exercise on pulmonary O2 uptake kinetics during transitions to high-intensity exercise from an elevated baseline. *J Appl Physiol* 105: 538–546, 2008.

36. **Engelen M**, **Porszasz J**, **Riley M**, **Wasserman K**, **Maehara K**, **Barstow TJ**. Effects of hypoxic hypoxia on O2 uptake and heart rate kinetics during heavy exercise. *J Appl Physiol* 81: 2500–2508, 1996.

37. **Eßfeld D**, **Hoffmann U**, **Stegemann J**. VO2 kinetics in subjects differing in aerobic capacity: investigation by spectral analysis. *Eur J Appl Physiol* 56: 508–515, 1987.

38. **Eßfeld D**, **Hoffmann U**, **Stegemann J**. A model for studying the distortion of muscle oxygen uptake patterns by circulation parameters. *Eur J Appl Physiol Occup Physiol* 62: 83–90, 1991.

39. **Fick A**. Uber die messung des blutquantums in den herz-ventrikein. *Sitzungsbericht der Phys Gesellschaft Wurzbg* 2: 169–170, 1870.

40. **Gibson A**, **Wagner D**, **Heyward V**. *Advanced Fitness Assessment and Exercise Prescription*. 8th ed. Champaign, Illinois: Human Kinetics, 2018.

41. **Hagberg JM**, **Hickson RC**, **Ehsani AA**, **Holloszy JO**. Faster adjustment to and recovery from submaximal exercise in the trained state. *J Appl Physiol* 48: 218–224, 1980.

42. **Haouzi P**, **Fukuba Y**, **Casaburi R**, **Stringer W**, **Wasserman K**. O2 uptake kinetics above and below the lactic acidosis threshold during sinusoidal exercise. *J Appl Physiol* 75: 1683–1690, 1993.

43. **Hickson RC**, **Bomze HA**, **Holloszy JO**. Faster adjustment of O2 uptake to the energy requirement of exercise in the trained state. *J Appl Physiol Respir Environ Exerc Physiol* 44: 877–881, 1978.

44. **Hochreiter S**, **Schmidhuber J**. Long short-term memory. *Neural Comput* 9: 1735–1780, 1997.

45. **Hoffmann U**, **Drescher U**, **Benson AP**, **Rossiter HB**, **Eßfeld D**. Skeletal muscle VO2

kinetics from cardio-pulmonary measurements: assessing distortions through O2 transport by means of stochastic work-rate signals and circulatory modelling. *Eur J Appl Physiol* 113: 1745–1754, 2013.

46. **Hoffmann U**, **Eßfeld D**, **Leyk D**, **Wunderlich H**, **Stegemann J**. Prediction of individual oxygen uptake on-step transients from frequency responses. *Eur J Appl Physiol Occup Physiol* 69: 93–97, 1994.

47. **Hoffmann U**, **Eßfeld D**, **Wunderlich H-G**, **Stegemann J**. Dynamic linearity of VO2 responses during aerobic exercise. *Eur J Appl Physiol Occup Physiol* 64: 139–144, 1992.

48. **Hughson RL**. Oxygen uptake kinetics: historical perspective and future directions. *Appl Physiol Nutr Metab* 34: 840–850, 2009.

49. **Hughson RL**, **Morrissey M**. Delayed kinetics of respiratory gas exchange in the transition from prior exercise. *J Appl Physiol Respir Environ Exerc Physiol* 52: 921–929, 1982.

50. **Hughson RL**, **O'Leary DD**, **Betik AC**, **Hebestreit H**. Kinetics of oxygen uptake at the onset of exercise near or above peak oxygen uptake. *J Appl Physiol* 88: 1812–1819, 2000.

51. **Hughson RL**, **Tschakovsky ME**, **Houston ME**. Regulation of oxygen consumption at the onset of exercise. *Exerc Sport Sci Rev* 29: 129–133, 2001.

52. **Hughson RL**, **Winter DA**, **Patla AE**, **Swanson GD**, **Cuervo LA**. Investigation of VO2 kinetics in humans with pseudorandom binary sequence work rate change. *J Appl Physiol* 68: 796–801, 1990.

53. **Jones AM**, **Carter H**, **Pringle JSM**, **Campbell IT**. Effect of creatine supplementation on oxygen uptake kinetics during submaximal cycle exercise. *J Appl Physiol* 92: 2571–2577, 2002.

54. **Jones AM**, **Wilkerson DP**, **DiMenna FJ**, **Fulford J**, **Poole DC**. Muscle metabolic responses to exercise above and below the "critical power" assessed using 31 P-MRS. *Am J Physiol - Regul Integr Comp Physiol* 294: R585–R593, 2008.

55. **Kaminsky LA**, **Arena R**, **Myers J**. Reference standards for cardiorespiratory fitness measured with cardiorespiratory exericse testing: data from the fitness registry and the importance of exercise national database. *Mayo Clin Proc* 90: 1515–1523, 2015.

56. **Keir DA**, **Benson AP**, **Love LK**, **Robertson TC**, **Rossiter HB**, **Kowalchuk JM**. Influence of muscle metabolic heterogeneity in determining the VO2p kinetic response to ramp-incremental exercise. *J Appl Physiol* 120: 503–513, 2016.

57. **Keir DA**, **Murias JM**, **Paterson DH**, **Kowalchuk JM**. Breath-by-breath pulmonary O2 uptake kinetics: effect of data processing on confidence in estimating model parameters. *Exp Physiol* 99: 1511–1522, 2014.

58. **Keir DA**, **Nederveen JP**, **Paterson DH**, **Kowalchuk JM**. Pulmonary O2 uptake kinetics during moderate - intensity exercise transitions initiated from low versus elevated metabolic rates: insights from manipulations in cadence. *Eur J Appl Physiol* 114: 2655–2665, 2014.

59. **Keir DA**, **Paterson DH**, **Kowalchuk JM**, **Murias JM**. Using ramp-incremental VO2 responses for constant-intensity exercise selection. *Appl Physiol Nutr Metab* 43: 882–892, 2018.

60. **Keir DA**, **Robertson TC**, **Benson AP**, **Rossiter HB**, **Kowalchuk JM**. The influence of metabolic and circulatory heterogeneity on the expression of pulmonary oxygen uptake kinetics in humans. *Exp Physiol* 101: 176–192, 2016.

61. **Koppo K**, **Bouckaert J**, **Jones AM**. Effects of training status and exercise intensity on phase II VO2 kinetics. *Med Sci Sports Exerc* 36: 225–232, 2004.

62. **Lamarra N**, **Whipp BJ**, **Ward SA**, **Wasserman K**. Effect of interbreath fluctuations on characterizing exercise gas exchange kinetics. *J Appl Physiol* 62: 2003–2012, 1987.

63. **MacDonald M**, **Pedersen PK**, **Hughson RL**. Acceleration of VO2 kinetics in heavy submaximal exercise by hypoxia and prior high-intensity exercise. *J Appl Physiol* 83: 1318–1325, 1997.

64. **MacPhee SL**, **Shoemaker JK**, **Paterson DH**, **Kowalchuk JM**. Kinetics of O2 uptake, leg blood flow, and muscle deoxygenation are slowed in the upper compared with lower region of the moderate-intensity exercise domain. *J Appl Physiol* 99: 1822–1834, 2005.

65. **McNarry MA**, **Kingsley MIC**, **Lewis MJ**. Influence of exercise intensity on pulmonary oxygen uptake kinetics in young and late middle-aged adults. *Am J Physiol - Regul Integr Comp Physiol* 303: R791–R798, 2012.

66. **Murias JM**, **Kowalchuk JM**, **Paterson DH**. Speeding of VO2 kinetics in response to endurance-training in older and young women. *Eur J Appl Physiol* 111: 235–243, 2011.

67. **Özyener F**, **Rossiter HB**, **Ward SA**, **Whipp BJ**. Influence of exercise intensity on the on- and off- transient kinetics of pulmonary oxygen uptake in humans. *J Physiol* 533: 891–902, 2001.

68. **Paterson DH**, **Whipp BJ**. Asymmetries of oxygen uptake transients at the on- and offset of heavy exercise in humans. *J Physiol* 443: 575–586, 1991.

69. **Powers SK**, **Dodd S**, **Beadle RE**. Oxygen uptake kinetics in trained athletes differing in VO2max. *Eur J Appl Physiol Occup Physiol* 54: 306–308, 1985.

70. **Roston WL**, **Whipp BJ**, **Davis JA**, **Cunningham DA**, **Effros RM**, **Wasserman K**. Oxygen uptake kinetics and lactate concentration during exercise in humans. *Am Rev Respir Dis* 135: 1080–1084, 1987.

71. **Sawka MN**, **Friedl KE**. Emerging wearable physiological monitoring technologies and decision aids for health and performance. *J Appl Physiol* 124: 430–431, 2018.

72. **Schalcher C**, **Rickli H**, **Brehm M**, **Weilenmann D**, **Oechslin E**, **Kiowski W**, **Brunner-La Rocca HP**. Prolonged oxygen uptake kinetics during low-intensity exercise are related to poor prognosis in patients with mild-to-moderate congestive heart failure. *Chest* 124: 580–586, 2003.

73. **Scherr J**, **Wolfarth B**, **Christle JW**, **Pressler A**, **Wagenpfeil S**, **Halle M**. Associations

between Borg's rating of perceived exertion and physiological measures of exercise intensity. *Eur J Appl Physiol* 113: 147–155, 2013.

74. **Spencer MD**, **Murias JM**, **Lamb HP**, **Kowalchuk JM**, **Paterson DH**. Are the parameters of VO2, heart rate and muscle deoxygenation kinetics affected by serial moderate-intensity exercise transitions in a single day? *Eur J Appl Physiol* 111: 591–600, 2011.

75. **Stojanovic E**, **Stojiljkovic N**, **Scanlan AT**, **Dalbo VJ**, **Berkelmans DM**, **Milanovic Z**. The activity demands and physiological responses encountered during basketball match-play: a systematic review. *Sport Med* 48: 111–135, 2018.

76. **Swanson GD**, **Hughson RL**. On the modeling and interpretation of oxygen uptake kinetics from ramp work rate tests. *J Appl Physiol* 65: 2453–2458, 1988.

77. **Tschakovsky ME**, **Hughson RL**. Interaction of factors determining oxygen uptake a the onset of exercise. *J Appl Physiol* 86: 1101–1113, 1999.

78. **Villar R**, **Beltrame T**, **Hughson RL**. Validation of the Hexoskin wearable vest during lying, sitting, standing, and walking activities. *Appl Physiol Nutr Metab* 40: 1019–1024, 2015.

79. **Ward SA**. Exercise physiology: exercise hyperpnea. *Curr Opin Psychol* 10: 166–172, 2019.

80. **Wasserman K**, **Hansen JE**, **Sue DY**, **Casaburi R**, **Whipp BJ**. *Principles of exercise testing and interpretation: including pathophysiology and clinical applications*. 3rd ed. Philadelphia: Lippincott Williams & Wilkins, 1999.

81. **Whipp BJ**. The slow component of O2 uptake kinetics during heavy exercise. *Med Sci Sports Exerc* 26: 1319–1326, 1994.

82. **Whipp BJ**, **Davis JA**, **Torres F**, **Wasserman K**. A test to determine parameters of aerobic function during exercise. *J Appl Physiol* 50: 217–221, 1981.

83. **Whipp BJ**, **Ward SA**, **Lamarra N**, **Davis JA**, **Wasserman K**. Parameters of ventilatory and gas exchange dynamics during exercise. *J Appl Physiol Respir Environ Exerc Physiol* 52: 1506–1513, 1982.

84. **Whipp BJ**, **Ward SA**, **Rossiter HB**. Pulmonary O2 uptake during exercise: conflating muscular and cardiovascular responses. *Med Sci Sports Exerc* 37: 1574–1585, 2005.

85. **Whipp BJ**, **Wasserman K**. Oxygen uptake kinetics work for various intensities of constant-load work. *J Appl Physiol* 33: 351–356, 1972.

86. **Whitcher L**, **Papadopoulos C**. Accelerometer derived activity counts and oxygen consumption between young and older individuals. *J Aging Res* 2014: 1–8, 2014.

87. **Wilcox SL**, **Broxterman RM**, **Barstow TJ**. Constructing quasi-linear VO2 responses from nonlinear parameters. *J Appl Physiol* 120: 121–129, 2016.

88. **Wilkerson DP**, **Jones AM**. Effects of baseline metabolic rate on pulmonary O2 uptake on-kinetics during heavy-intensity exercise in humans. *Respir Physiol Neurobiol* 156:

203–211, 2007.

89. **Zignoli A**, **Fornasiero A**, **Ragni M**, **Pellegrini B**, **Schena F**, **Biral F**, **Laursen PB**. Estimating an individual's oxygen uptake during cycling exercise with a recurrent neural network trained from easy-to-obtain inputs: A pilot study. *PLoS One* 15: e0229466, 2020.

90. **Zoladz JA**, **Gladden LB**, **Hogan MC**, **Nieckarz Z**, **Grassi B**. Progressive recruitment of muscle fibers is not necessary for the slow component of VO2 kinetics. *J Appl Physiol* 105: 575–580, 2008.

91. **Zoladz JA**, **Rademaker ACHJ**, **Sargeant AJ**. Non-linear relationship between O2 uptake and power output at high intensities of exercise in humans. *J Physiol* 488: 211–217, 1995.

Appendix: Alternative Model Results

The results of a random forest regression model that used model features of work rate, percent heart rate reserve, delta heart rate (change in heart rate from the previous second), estimated minute ventilation, and breathing frequency to predict absolute oxygen uptake ($\dot{V}O_{2p}$). This model most closely reflects the features used in the validated moderate intensity exercise $\dot{V}O_{2p}$ predictor developed by Beltrame et al. (13).

Group mean (n = 15) time-series comparisons between measured and predicted $\dot{V}O_{2p}$ are presented in Fig. 17. $\dot{V}O_{2p}$ prediction accuracy was similar between the proposed model in Chapter 3 (Fig. 11), and the alternative model presented in the appendix (Fig. 18). However, using work rate as a model feature resulted in abrupt changes in predicted $\dot{V}O_{2p}$ at the time points that corresponded with changes in work rate (Fig. 19).

The alternative predictive model appears to be a poorer predictor of $\dot{V}O_{2p}$ kinetics than the model used in Chapter 3. Significant main effects of Condition ($F_{(1.402, 19.634)} = 42.274$, $p < 0.001$) and Data type ($F_{(1, 14)} = 33.269$, $p < 0.001$), as well as a Condition x Data type interaction ($F_{(2, 28)} = 14.491$, $p < 0.001$) were observed on mean normalized gain (MNG). *Post-hoc* analysis revealed a similar pattern of pairwise differences as in Chapter 3 (Fig. 20). The alternative model was also biased to predicting slower kinetics compared to the measured responses (Fig. 21). Interestingly, the measured and predicted MNG by the alternative model was not correlated between-participants for individual exercise conditions ($p > 0.05$), nor within-participants across all exercise conditions ($r_{rm} = 0.204$, $p = 0.271$; Fig. 22).

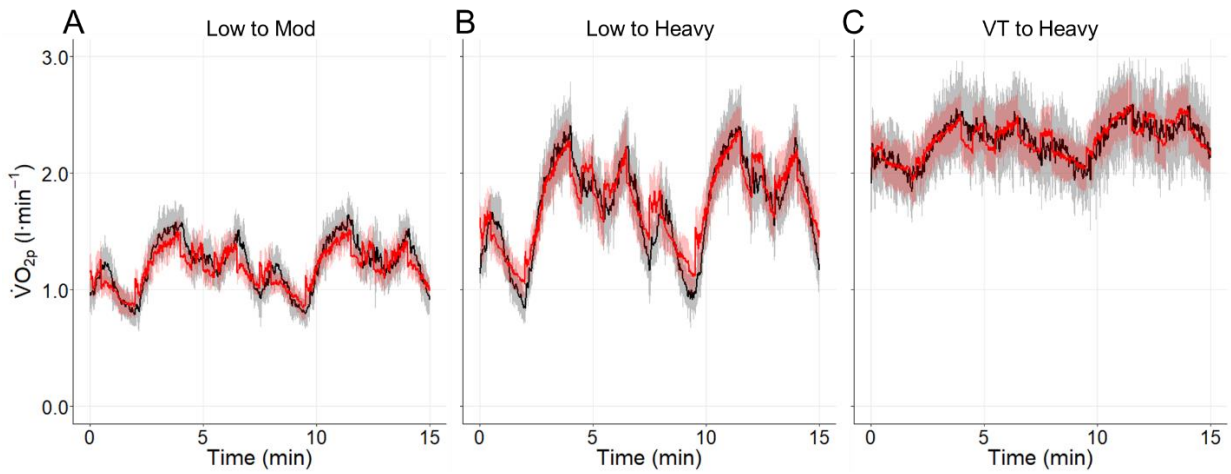

Fig. 17. Group mean ± SD measured (black) and predicted (red) oxygen uptake ($\dot{V}O_{2p}$) responses to pseudorandom binary sequence exercise in low to mod (A), low to heavy (B), and ventilatory threshold (VT) to heavy (C) conditions following 3.5 minutes of warm-up exercise (not shown).

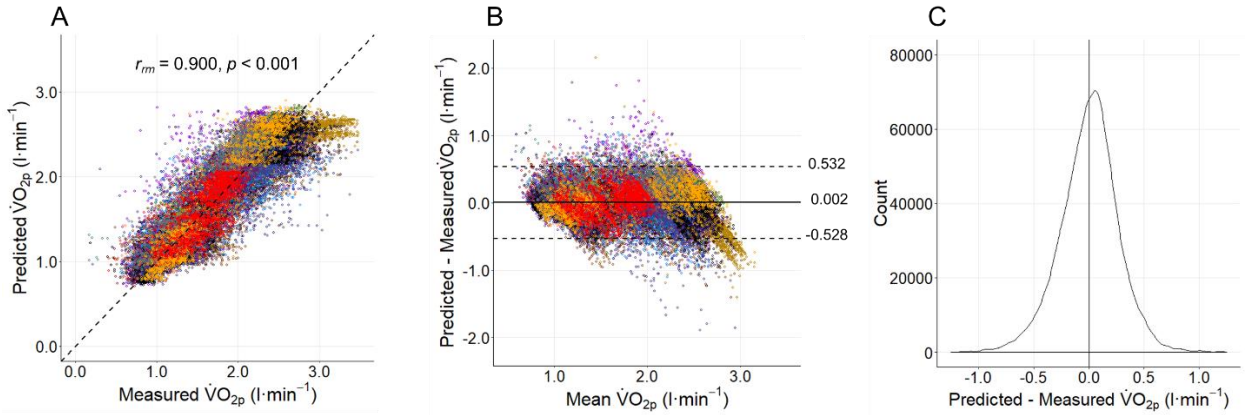

Fig. 18. Accuracy of the alternative oxygen uptake predictor ($\dot{V}O_{2p}$) with all three exercise conditions combined. A: repeated measures correlation of predicted and measured $\dot{V}O_{2p}$. Parallel regression lines for each participant were omitted for clarity. Dotted line represents the identity line. B: repeated measures Bland-Altman plot. C: distribution of the differences between the predicted and measured $\dot{V}O_{2p}$.

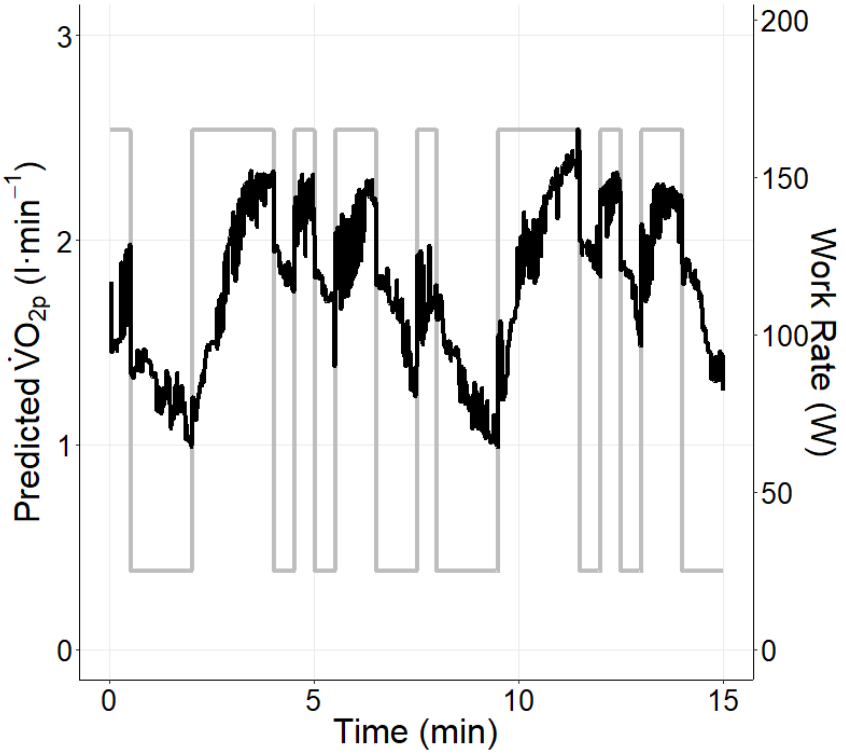

Fig. 19. Predicted oxygen uptake ($\dot{V}O_{2p}$; black) and work rate (grey) profiles of a representative participant during low to heavy pseudorandom binary sequence exercise. Abrupt changes in predicted $\dot{V}O_{2p}$ occur concomitantly with changes in work rate.

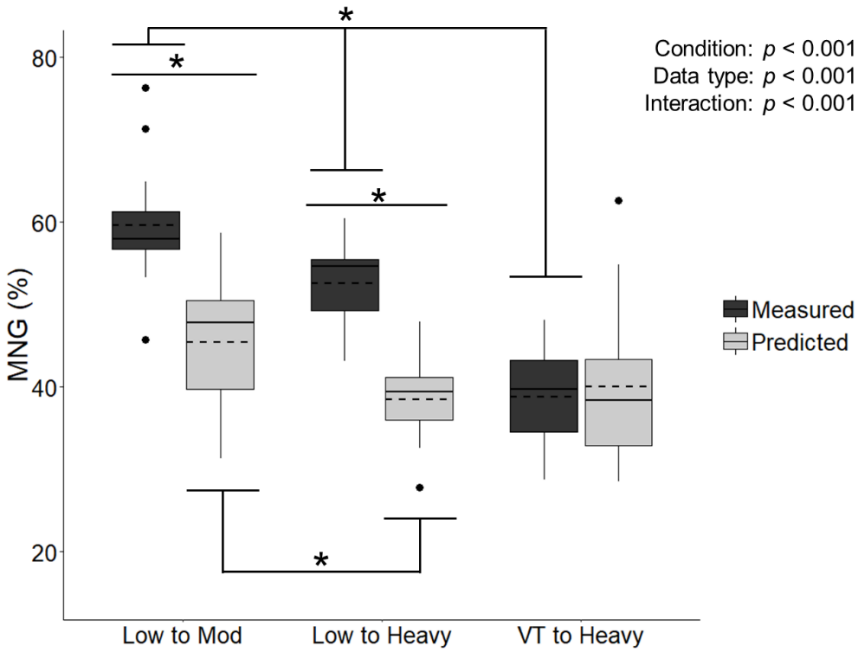

Fig. 20. Comparison of mean normalized gain (MNG) calculated from predicted and measured oxygen uptake data across exercise conditions. * significant ($p < 0.05$) difference between points. Dashed and solid horizontal lines in box plots are mean and median, respectively.

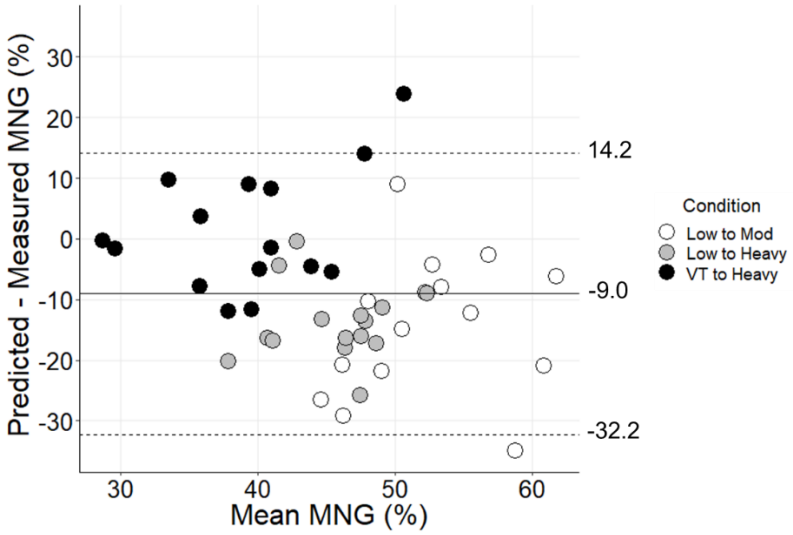

Fig. 21. Repeated measures Bland-Altman plot of the mean normalized gain (MNG) calculated using the predicted and measured oxygen uptake data from the low to moderate (low to mod; white), low to heavy (grey), and ventilatory threshold to heavy (VT to heavy; black) exercise conditions.

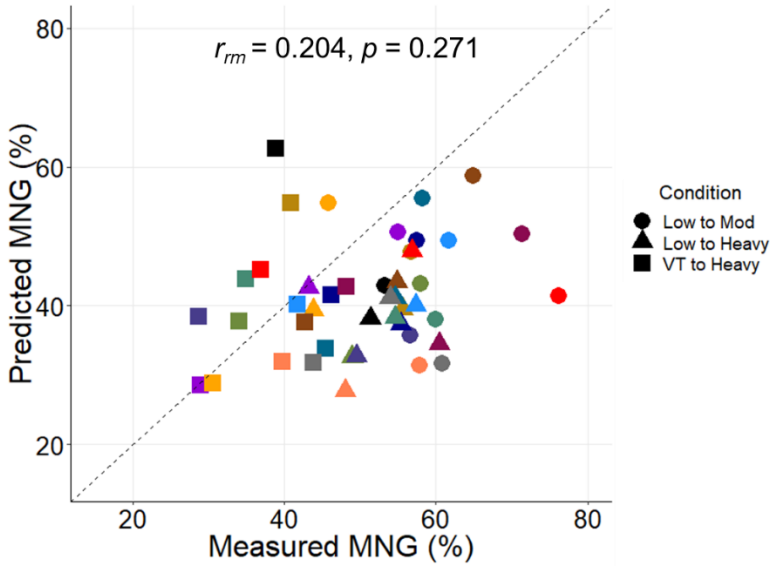

Fig. 22. Repeated measures correlation between the mean normalized gain (MNG) calculated from the predicted oxygen uptake ($\dot{V}O_{2p}$) and the MNG calculated from the directly measured $\dot{V}O_{2p}$. All participants' (n = 15) measurements from low to moderate (circles), low to heavy (triangles), and ventilatory threshold (VT) to heavy (squares) exercise conditions are shown. A different color represents each participant. Parallel regression lines for each participant were omitted for clarity. Dotted line represents the identity line.