Measuring the Sympathetic Response to Intense Exercise in a Practical Setting

Shiva Kaul SHIVA.KAUL@GMAIL.COM
55 Prospect St.
Morristown, NJ 07960

Dr. Anthony Falco, PT, DPT ANTHONY@FALCOPHYSIO.COM
Performance Physical Therapy & Sports Conditioning
2025 Hamburg Turnpike
Wayne, NJ 07470

Karianne Anthes KARIANNE@CFMOTOWN.COM
CrossFit Morristown
38 Dumont Pl
Morristown, NJ 07960

Abstract

Brief, intense exercise can improve health due to its acute effect on the autonomic nervous system, particularly the sympathetic nervous system. Salivary amylase is a marker of sympathetic activity during exercise, but it requires specialized equipment to measure. We investigate the feasibility of estimating the amylase response from heartbeat data recorded by commodity sensors. We collect heartbeat and amylase data for $n = 71$ sessions of intense exercise performed in a commercial setting. Our machine learning model exploits structure in the heartbeat signal: by identifying and removing the contribution of the parasympathetic nervous system, we obtain a residual with sympathetic information, to which we apply a convolutional neural network. This model has better accuracy than existing measures of exercise response, such as maximum heart rate, even though it doesn't use metadata such as age and gender. This suggests sympathetic activity may be (weakly) discerned from heartbeat data. With a larger dataset, a practical measure of sympathetic response to exercise could potentially be developed. Our quantification of parasympathetic activity is more powerful than existing approaches and may have independent value.

1. Introduction

Intense exercise elicits a much different physiological response than moderate exercise. At low and moderate intensities, energy demands are satisfied by the oxidative (aerobic) pathway, and the initial increase in heart rate is accompanied by withdrawal of the parasympathetic nervous system. At high intensities, the glycolytic (anaerobic) energy pathway predominates; sweating, lipolysis, gluconeogenesis, and other characteristic responses are driven mostly by activation of the sympathetic nervous system rather than further parasympathetic withdrawal (White and Raven, 2014; Michael et al., 2017a; Koistinen and Laitinen, 2004). Intense training has been shown to improve insulin sensitivity, blood pressure, aerobic capacity (VO_2max), and body composition (Batacan et al., 2017; Jelleyman et al.,

2015; Jelleyman, 2018; Kessler et al., 2012; Milanović et al., 2015). Even a single bout of intense exercise can have acute health benefits, such as enhancing glucose control (Marliss and Vranic, 2002; Jelleyman, 2018) or inhibiting the growth of colon cancer (Devin et al., 2019). Intense exercise interventions are safe (Wewege et al., 2018; de Jong et al., 2003; Carl et al., 2016) and have the potential to improve outcomes for numerous patient populations (Weston et al., 2014; Jelleyman et al., 2015; Elliott et al., 2015; Hannan et al., 2018).

It is hard to measure if the desired response to intense exercise was actually achieved. Subjective measures, such as RPE (rate of perceived exertion) or RIR (repetitions in reserve), are common in athletic training. However, these measures are often unreliable in patient populations unaccustomed or indisposed to exercise (Unick et al., 2014; Aamot et al., 2014; Strzelczyk et al., 2001). As previously mentioned, the activity of the sympathetic nervous system is of fundamental importance. There are many methods of measuring sympathetic activity, but none is considered a gold standard (Grassi and Esler, 1999). The most common measure of sympathetic tone is the concentration of plasma epinephrine (a.k.a. adrenaline) or norepinephrine. Unfortunately, this requires invasive, confounding blood draws and complex laboratory analysis. The cardiac preejection period has been proposed as a valid, noninvasive measure of sympathetic activity (Michael et al., 2017a). Unfortunately, it is recorded by bioimpedance cardiography, which is sensitive to the postural changes and heavy breathing that occur during exercise. Measurements can also be made at the periphery. Microneurography involves electrodes inserted directly into muscle or skin nerves, which prohibits large movements during exercise (Vallbo et al., 2004). Electrodermal activity (EDA) of sweat glands correlates with sympathetic activity during exercise (Boettger et al., 2010; Posada-Quintero et al., 2018). However, peripheral measurements are not uniform across the body (Shoemaker et al., 2018,?), the most accurate measurement locations are not convenient (van Dooren et al., 2012), and are affected by local phenomena, such as vasoconstriction (Edelberg, 1964).

Salivary α-amylase (briefly, "amylase") has been identified as a marker of sympathetic tone, especially during exercise (Chicharro et al., 1998; Koibuchi and Suzuki, 2014; Nater and Rohleder, 2009); (nor)epinephrine activates β_1-adrenergic receptors in the salivary glands, which causes granules of this enzyme to be released. Below a minimum threshold of intensity – which seems to coincide with the accumulation of lactate in blood (Calvo et al., 1997; Bocanegra et al., 2012; Akizuki et al., 2014) – the change in amylase is negligible. Above that threshold, it rises proportionally with intensity (Li and Gleeson, 2004; De Oliveira et al., 2010). Amylase is typically measured by immunoassay of saliva sampled by passive drool. It can also be immediately measured by point-of-care devices (Shetty et al., 2011). However, both methods incur nonneglible marginal cost due to nonreusable materials. Because salivary flow rate changes during exercise (Rohleder and Nater, 2009; Bosch et al., 2011), the point-of-care devices are prone to substantial error, if used without careful adherence to protocol (Peng et al., 2016).

Practitioners have resorted to metrics based on heart rate, because it is cheap and practical to measure with commodity sensors. These metrics include average and maximum heart rate, the decrease in heart rate 60 seconds after exercise (HRR60), and the rate at which heart rate returns to baseline (HRRτ). With sensors capable of recording the times between individual heartbeats, an assortment of heart rate variability (HRV) metrics, such as SDNN, RMSSD, PNN50, and LF, can be calculated during or after exercise (Shaffer and

Ginsberg, 2017). Though these metrics may be useful for monitoring recovery, assessing fitness, or related tasks, their utility as measures of the sympathetic response to intense exercise is limited. LF — the weight placed on low-frequency ($0.04 - 0.15$ Hz) components of the Fourier decomposition of the heartbeat signal — was previously thought to reflect slower-responding sympathetic tone, but is now understood to be parasympathetically driven (Reyes del Paso et al., 2013; Thomas et al., 2019; Moak et al., 2007). During exercise, HRV reaches a near-minimum at a relatively low intensity, analogous to parasympathetic tone (Michael et al., 2017a; Boettger et al., 2010). Following exercise, HRV recovery, unlike sympathetic withdrawal, is substantially delayed by duration (Michael et al., 2017c) and moderate intensity (Michael et al., 2017b). Similarly, HRR60 primarily, though not entirely, reflects parasympathetic reactivation (Kannankeril et al., 2004). HRRτ seems to have a stronger dependence on sympathetic tone (Buchheit et al., 2007); in skeletal muscle, accumulated metabolites stimulate metaboreceptors, which in turn maintain sympathetic tone (Fisher et al., 2013). HRRτ, along with the other recovery metrics, require monitoring from 10 minutes to hours after the cessation of exercise, during which upright posture and further activity may confound results.

Since intense exercise is hard to monitor, it is hard to prescribe. Exercise intensity is typically specified as a percentage of some unknown, estimated quantity, such as maximum perceived exertion, maximum heart rate, or VO_2max. The error in this estimate is amplified when prescribing intense exercise at 95% VO_2max rather than moderate exercise at 60% VO_2max. Such imprecision stokes lingering concerns of overexertion during intense exercise. Mann et al. (2013) review the large variation of physiological responses to "poorly standardized" exercise protocols. Imprecise dosing of exercise raises concerns of overexertion and results in worse health outcomes. The SMARTEX heart study (Ellingsen et al., 2017a) found intense exercise was not better than moderate exercise for rehabilitating cardiac-failure patients, because some patients did not exercise at the correct intensity. In their view, "tight control of prescribed exercise intensity and intended load increase was somehow lost in the translation from a small proof-of-principle study to a larger multicenter trial of the efficacy under conditions closer to standard clinical practice" (Ellingsen et al., 2017b).

1.1. Our contribution

We investigate the feasibility of measuring sympathetic response without specialized equipment. Using machine learning, we develop an algorithm which estimates the change in amylase enzyme activity (from before exercise to after, denoted as ΔAmylase) from heartbeat data readily collected during exercise. This is a supervised regression problem: the input is an arbitrary-length sequence of interbeat intervals, and the label is ΔAmylase, a real value with units U/mL. We also consider the associated comparison problem: determining if one workout resulted in greater ΔAmylase than another workout.

To train and evaluate the algorithm, we collect a realistic, albeit noisy, dataset. Previous studies were conducted in laboratories or elite athletic settings, involved a narrow population of participants, exerted tight control over their diet and schedule, and specified a small number of exercises (primarily cycling). Our dataset is collected in commercial group training sessions, involves a diverse group of participants, observes them in their day-to-day

exercise routine, and involves a wide variety of exercises. However, there may be substantial noise in both the heartbeat data (due to sensors slipping in full-body exercises) as well as the amylase measurements (due to changes in salivary flow and inadvertent misuse of equipment). Whereas laboratory settings have a tendency to produce optimistic results, our study is designed to produce pessimistic ones.

Our machine learning model is informed by the physiology of the autonomic nervous system. The parasympathetic system contributes a substantial, high-frequency component to the heartbeat signal. Heart rate variability (HRV) is a fairly reliable indicator of parasympathetic activity. By removing an HRV-derived component from the heartbeat signal, we obtain a residual with information about sympathetic activity. Multiple HRV metrics may be computed from second moments of the interbeat intervals. To quantify the parasympathetic contribution, rather than using an existing HRV metric, we allow more general, parametrized metrics of the same underlying moments. The generalization is mild enough to still consider them variability metrics rather than arbitrary statistics. The parameters of this metric are (pre)trained upon the data. Subsequently predicting ΔAmylase from the residual is a straightforward application of convolutional neural networks.

Clinical Relevance. A practical measure of sympathetic response could greatly improve the clinical practice of intense exercise. For doctors and researchers, a measure could ease the development and specification of intense exercise regimens. For patients and athletes, it could help ensure that improvements to health and fitness are actually being made. We initiate the study of this problem in the hope of motivating further research. On our dataset, as a measure of sympathetic response, heart rate (the predominant metric) is no better than random guessing. Our model achieves a modest but discernible improvement. However, our dataset is too small ($n = 71$) to train a practically usable model.

Technical Significance. Our work restores some optimism that heartbeat data contains information about sympathetic activity. Our novel approach of eliminating parasympathetic influence from the heartbeat signal proves useful; a naive application of CNNs is not as accurate. Our generalization of HRV metrics may have some independent utility for quantifying parasympathetic activity.

1.2. Outline

In Section 2, we describe the cohort and the data collection process. In Section 3, we present basic statistics of our dataset. These suggest our estimation problem is challenging, and motivate the use of machine learning. In Section 4, we present our machine learning algorithm, and compare it to previous work. In Section 5, we present the results of fitting our algorithm to the data. In Section 6, we review our findings and offer guidance for future research.

2. Study Design

Our goal is to collect a dataset in which the participants and their activities are realistically observed rather than tightly controlled. This makes the data noisier, and in turn makes the estimation problem harder. However, a model trained upon this data has a better chance of transferring to practical use. Importantly, the goal of our study is just to estimate amylase, not to bolster its physiologic validity as a measure of sympathetic response.

2.1. Cohort Selection

The participants in this study are members of a commercial fitness facility in New Jersey. This membership is evenly split between genders and has an average age of 45, not including minors who are not eligible for study. Approximately 80% of the members are white and 20% of the members are of another race. No competitive or professional athletes are included.

Each exercise session is performed as part of an hour-long, instructor-led class. The workout is relatively brief, typically between 10 and 15 minutes. It is preceded by a warmup involving dynamic exercises, static stretching, and weightlifting practice. The workouts involve a wide variety of exercises, including bodyweight movements, rowing, cycling, running, gymnastics, dumbbell exercises, and barbell lifts. Most of the workouts are driven by one of two goals: to complete the workload as quickly as possible ("time priority"), or to complete as many rounds as possible within a given time ("task priority"). The workout prescriptions are generated semirandomly, to avoid repetition of exercises and to keep participants engaged. The workout prescriptions are just guidelines; participants make modifications, such as reducing the weight, so they can complete the workout.

Most of the workouts are intended to be intense, though longer workouts are necessarily less intense than the shorter ones. A small fraction of the workouts focus on "accessory" exercises which are performed attentively with large rest periods and lower overall difficulty. Relatively low ΔAmylase is expected for these workouts. They may be thought of as informal controls, with relatively low changes in amylase. Another portion of the workouts are performed competitively while being judged; relatively high ΔAmylase is expected for these.

2.2. Data collection

The study began in February 2019 and lasted 4 weeks. Participants who consented to the study were instructed how to use the necessary equipment. Participants arrived at class as they usually would; we did not control their food intake, time of day, or any other factors. However, they were instructed to arrive 10-15 minutes prior to class in order to initiate data collection. Participants performed their own measurements, having been individually instructed on proper protocol. (Since they exercised throughout the day, it was not feasible to supervise them at all times.) Roughly 5 minutes before the class started, they measured their amylase by inserting a saliva measurement stick (pictured in Figure 1) under their tongue for 30 seconds. They took special care to saturate the stick with saliva, to avoid problems with salivary flow rate. The stick was then read with a portable colorimetric meter (Shetty et al., 2011). They then securely wore a Polar H10 (single-lead ECG cheststrap) heart rate monitor. The HRM connected to their personal smartphone via Bluetooth. To ease data collection and improve compliance, we implemented a custom iOS application, in which the participant could enter all their information, and the HRM could log its data. The interface of this application is presented in Figure 2. For the entire class, users remained within 15 meters of their phone, to prevent the HRMs from disconnecting. Roughly one minute after the main portion of the workout, prior to cooling down, the participants gauged their rate of perceived exertion, and again measured their amylase. Finally, participants noted if any unusual circumstances befell the workout or their preparation for it; these include illness, injury, a high temperature in the gym, or

Figure 1: Data collection equipment. Left: pouches of saliva swabs. Center: point-of-care salivary amylase meter, with a single saliva swab. Right: Polar H10 cheststrap heart rate monitors.

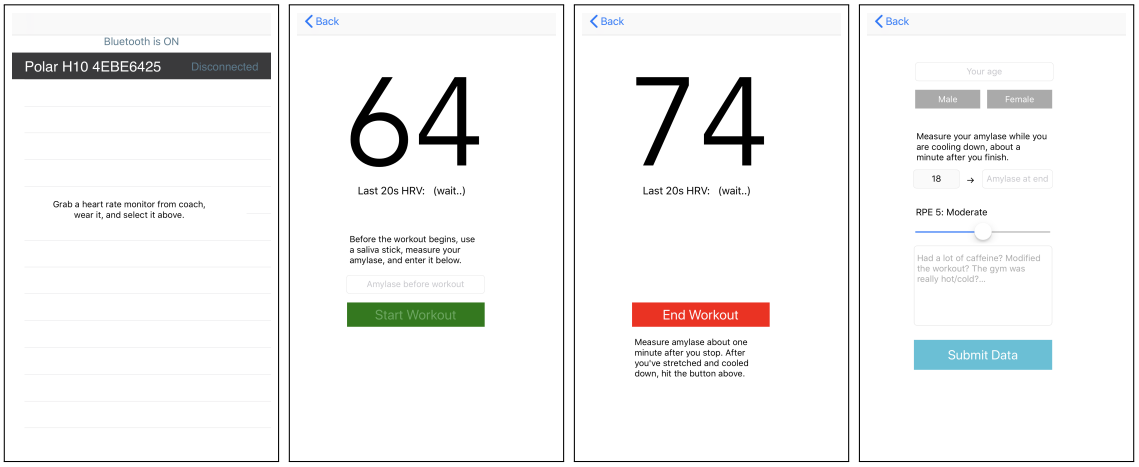

Figure 2: Screens of the iOS data collection application used in the study. **A**: The user connects to a Bluetooth heart rate monitor. **B**: Heart rate is prominently displayed. Starting amylase is recorded, and the green "start workout" button is pressed. **C**: Heart rate is continuously updated during the workout, after which the red "end workout" button is pressed. **D**: The remainder of the data (age, gender, ending amylase, and RPE) are collected.

an atypically large dose of caffeine. Finally, users submitted all of the data through the application. Using the start and end times of the data collection, we ensured that data was collected continuously.

We excluded samples on the following grounds. (1) User error: the user did not follow proper protocol (e.g. forgot to measure their amylase at the cessation of exercise.) (2) HRM error: the heart rate monitor prematurely disconnected from the phone. (3) Application error: an unknown technical error was encountered while using the iOS application. (4) Too much noise: the heart rate monitor slipped or malfunctioned, resulting in excessively noisy measurement.

3. Basic Data Preprocessing and Analysis

In total, 71 workouts were successfully recorded from 19 participants. Though small by the standards of modern machine learning, this is relatively large compared to previous studies. It is enough to cover a wide variety of exercise plans performed at varying degrees of intensity. In this section, we review the salient features of the dataset, and gain some intuition for the machine learning model.

3.1. Noise

The RR interval data are noisy. The noise consists primarily of isolated ectopic beats which spike below or above the otherwise-smooth curve. Ectopic beats are known to impede calculation of heart rate variability (Lippman et al., 1994). Accordingly, we eliminated the

Total workouts	85
User errors	3
HRM errors	2
App. errors	6
Too much noise	3
Dataset size	71

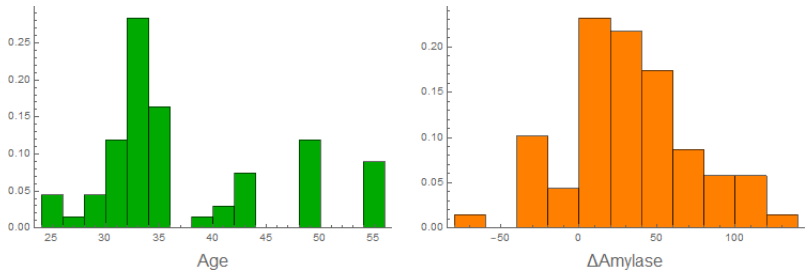

Figure 3: Basic statistics of the dataset. As the table shows, substantial number of errors were encountered during data collection. The mean age is 37. The mean absolute magnitude of ΔAmylase is 40. The mean absolute deviation of ΔAmylase is 29.4.

ectopic beats with the following noise filtering algorithm. First, we calculate a median filter on the entire RR interval sequence using a window size of 24. Next, we measure the relative deviation of the original RR interval from the filtered value. If this is more than 20%, then we consider the RR interval ectopic and replace it with the filtered value. If more than 15% of the beats are ectopic, then we exclude the entire sample from subsequent analysis. Three samples were excluded in this manner.

The ΔAmylase measurements are similarly noisy. Immediately following intense exercise, participants were sometimes forgetful or simply too tired to completely adhere to the measurement protocol. On occasion, saliva measurements were performed in duplicate. Based on these measurements, we estimate that the inherent noise in the ΔAmylase values is roughly 10. For similar reasons, RPE was essentially ignored; participants rarely changed it from its default value of 5. No such omissions were made for age or gender.

3.2. Heart Rate and ΔAmylase

As expected, there does not seem to be a discernible relationship between basic heart rate metrics and ΔAmylase, even when controlling for age. Figure 6 illustrates the lack of correlation. This is consistent with the consensus that heart rate is not a reliable measure of the response to exercise, at least across individuals.

Heart rate variability (HRV) measures the activity of the parasympathetic nervous system in terms of the variation of time between heartbeats. Higher variation corresponds to higher PS activity. The most commonly used HRV statistics are RMSSD (root mean square of successive differences) and SDNN (standard deviation of normal-to-normal intervals). See the equations in Section 4.3 for formal definitions, or Shaffer et al. (2014); Shaffer and Ginsberg (2017) for reviews of the subject. In Figure 6, we display, for a variety of workouts, SDNN computed over sliding windows of 24 beats. Other metrics behaved similarly. In this section, we simply note that HRV, in of itself, is not related to ΔAmylase. All the workouts have similar minimum, maximum, and mean SDNN. However, there is a dynamic relationship between HR and HRV, which we will examine in the next section.

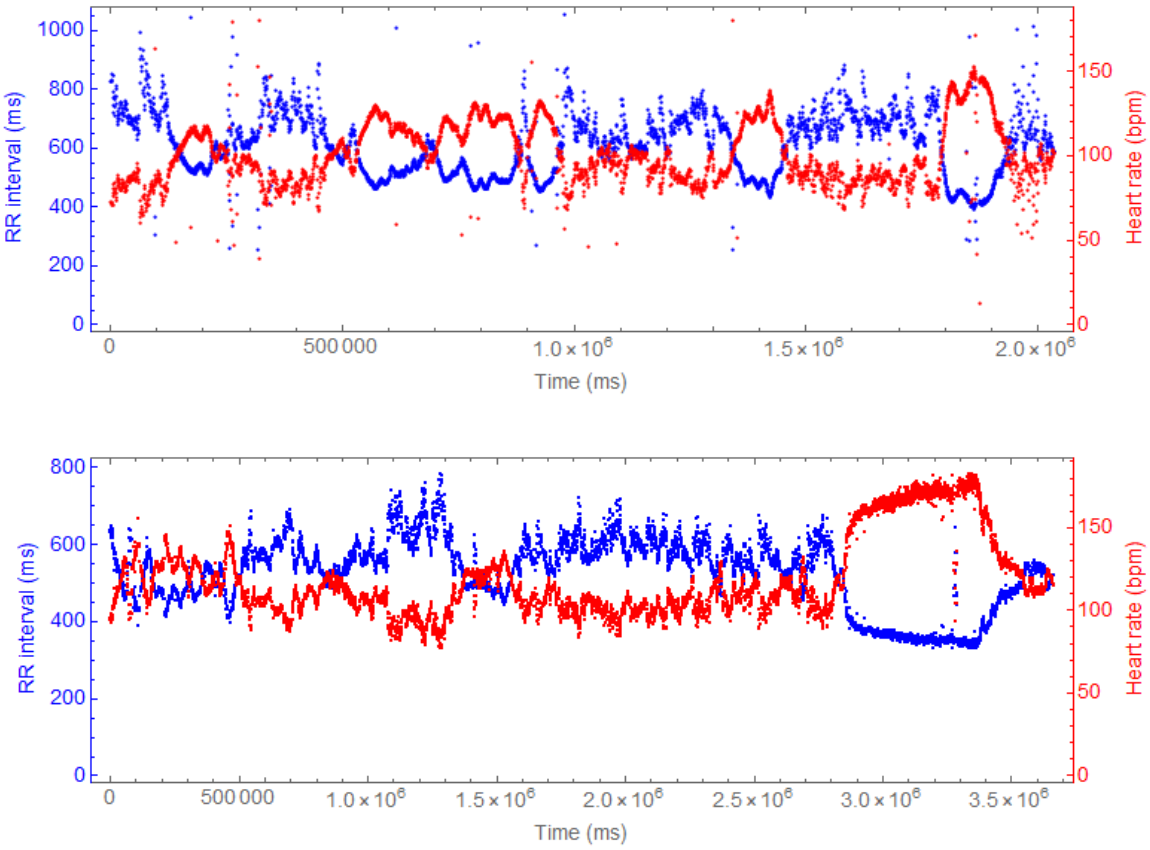

Figure 4: *Top*: a tachogram of the raw heartbeat data of a single workout. RR intervals are recorded, and converted to instantaneous heart rate via the equation HR = 60000/RR. The workout is 30 jump-rope double-unders and 15 dumbbell snatches, for as many rounds as possible within 10 minutes. The main working portion is near the end of the session, and is preceded by a substantial warmup. The data are presented without any smoothing or noise filtering. The large amount of noise may be partially attributed to jumping. *Bottom*: the previous tachogram processed by the noise filtering algorithm.

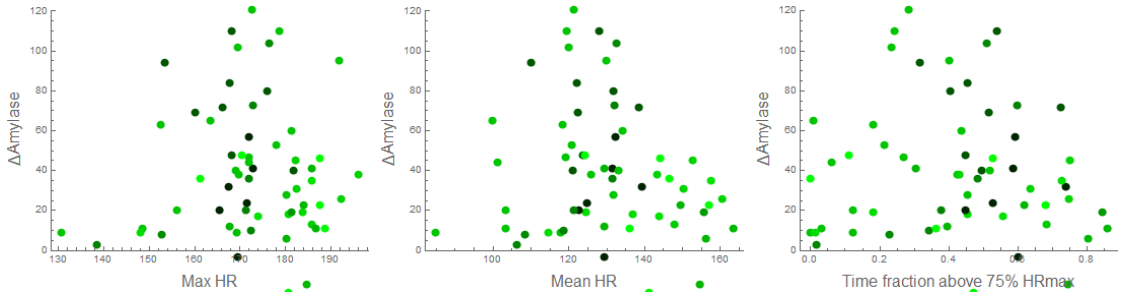

Figure 5: Various heart rate metrics, age and ΔAmylase do not seem to correlate. Bright green denotes young age, and black denotes old age. Five workouts resulted in negative ΔAmylase; these remain in the dataset, but are clipped off the plot for visual clarity. Maximum heart rate is estimated as HRmax $= 208 - 0.7 \cdot$ Age (Tanaka et al., 2001). Heart rate metrics are not a good measures of sympathetic response, even taking age into account.

4. Machine Learning Algorithm

4.1. Problem formulation

Let the length-T sequence of interbeat RR intervals be $x = [x_1, \ldots, x_T]$. Each $x_t \in \mathbb{R}^+$ has millisecond units. So, if R-peak t occurs half a second after R-peak $t-1$, then $x_t = 500$. Let $\Delta\text{Amylase} \in \mathbb{R}$ (now abbreviated as just Δ) be the difference in amylase incurred during exercise. Our goal is to find a function f which minimizes the following mean absolute error in predicting Δ:

$$\mathbf{E}\,|f(x) - \Delta|$$

We also consider the induced pairwise comparison problem.

$$\mathbf{P}\left((\Delta < \Delta') \implies (f(x) < f(x'))\right)$$

In our present context, these problems are challenging for the following reasons:

- The sequences are long — typically longer than 7000 elements — and are highly nonstationary, with long-range dependencies. This precludes the use of many RNNs, which are typically trained on short segments, lest they become a serial computational bottleneck.
- The signal-to-noise ratio is somewhat low. As discussed in Section 3.1, there is noise in both the inputs and the outputs. There is only a single point of supervision Δ for each sequence x.
- Age, gender, and starting time are excluded from the predictive model. We do this to avoid overfitting, and are interested to see if ΔAmylase can be estimated from heartbeat data alone.

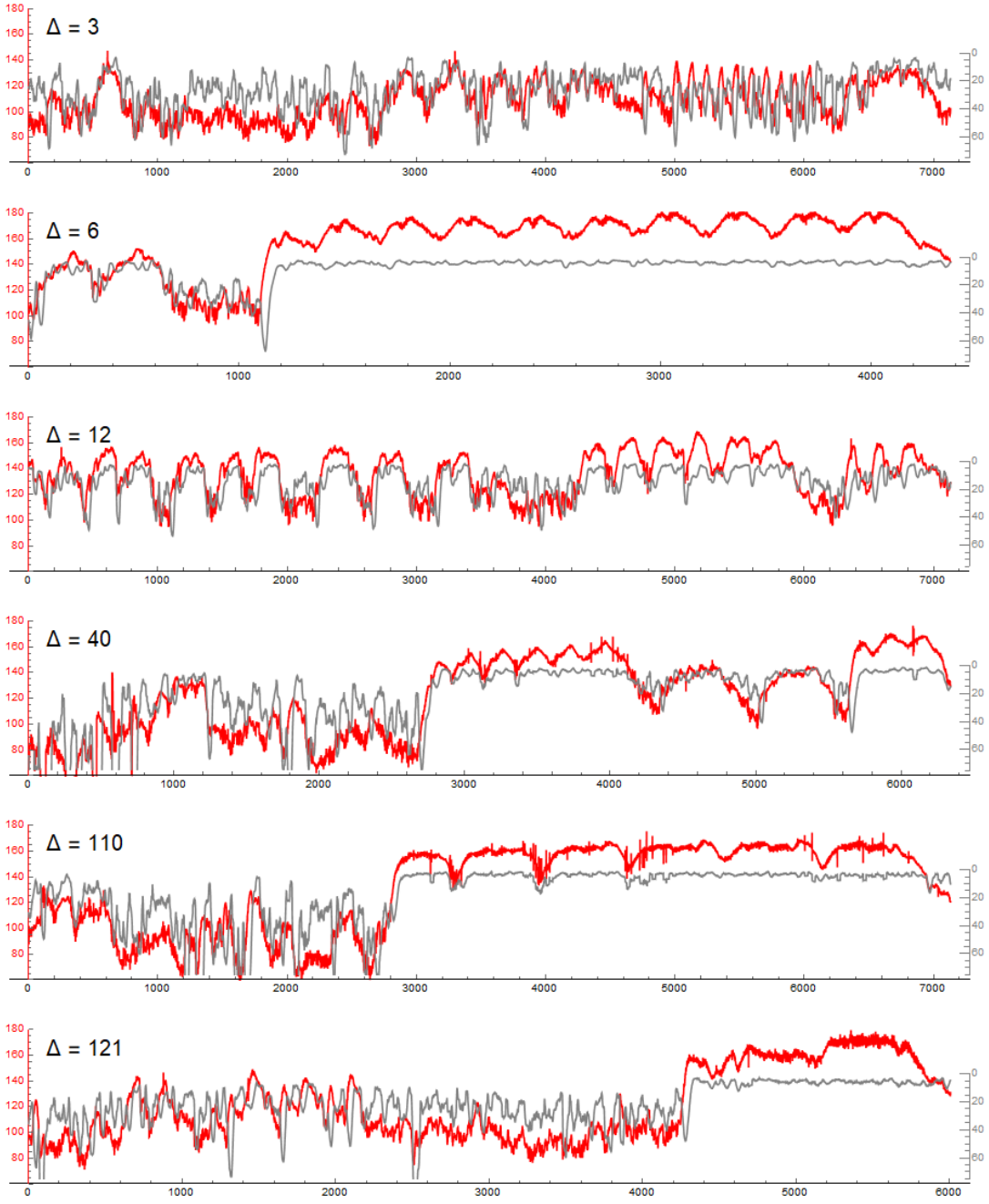

Figure 6: Heart rate (red) and SDNN (a measure of parasympathetic activity, gray) compared for workouts with different amylase changes (given by Δ in the top left). The vertical axis of SDNN is flipped for visual clarity. All of these workouts, except for the one with $\Delta = 6$, illustrate an interesting pattern. When Δ is low, changes in HR are mirrored by changes in SDNN. When Δ is high, HR and SDNN become decoupled: SDNN remains flat while HR may continue to increase. This manifests as the red line spiking above the gray line. HR changes that do not coincide with SDNN changes may be sympathetically driven.

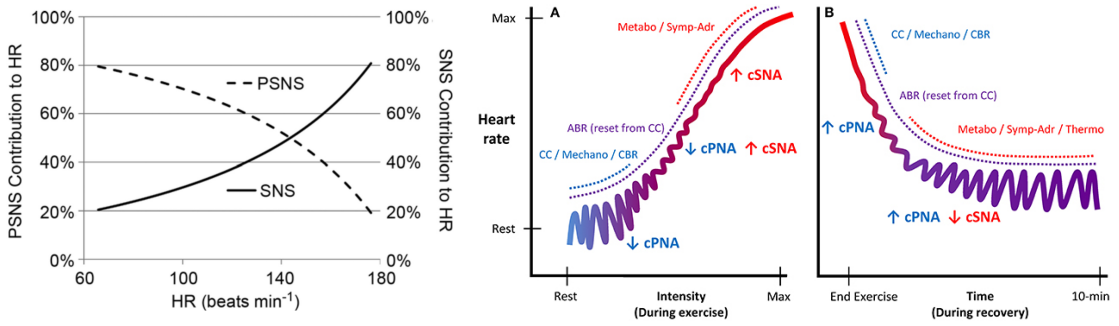

Figure 7: Schematic relationship of the parasympathetic (PS) and sympathetic (S) nervous systems during intense exercise. The left figure, from White and Raven (2014), shows that PS withdraws mostly at moderate intensities, and S activates mostly at high intensities. The right figure, from Michael et al. (2017a), shows that PS is faster-acting than S; it withdraws before S activates, and reactivates before S withdraws.

4.2. Key Intuitions

Let us motivate the design of our machine learning model by examining the relationship between HR and HRV in Figure 6. Large gaps between HR and (inverse) HRV seem to be a necessary condition for high ΔAmylase. That is, high ΔAmylase seems to be contingent upon sudden increases in HR despite no decrease in HRV. In workouts with low ΔAmylase, changes in HR seem to be mirrored by changes in HRV. It makes some visual sense to subtract (inverted) HRV to obtain the S activity that would explain the unaccounted HR changes. This idea makes some physiologic sense as well, due to the complementary relationship between the PS and S systems (Figure 7) . The physiology suggests another telltale clue: HRV increasing while HR remains elevated.

The overall model takes the form $f(x) = s(h(x) - (h \circ p \circ v \circ z)(x))$, where z are squared differences of x, v is a generalized HRV metric, p is the pretrained parasympathetic layer, $h(r) = 60000/r$ converts from RR intervals to instantaneous HR, $\tilde{x} = h(x) - h(p)$ is the residual, and s is the sympathetic layer. We describe the model architecture below.

4.3. Pretrained Parasympathetic Layer

To eliminate the parasympathetic influence on heart rate, it is tempting to simply subtract an existing HRV metric from heart rate. However, this is not quantitatively satisfactory for multiple reasons. First, it is not clear which HRV metric to use. Second, doing this in the manner of Figure 6 would involve translating and scaling the metric; these operations would have to be calibrated to the data. Finally, simply subtracting HRV may not be the best way of removing information from the signal. Ideally, the residual should be independent of parasympathetic activity, in that predicting the former from the latter should not be possible. Rather than attempting to ameliorate these issues by hand, we employ machine learning.

Squared differences. We observe that the most commonly-used HRV metrics can be expressed in terms of the squared differences $z_{i,j} = (x_i - x_j)^2$. This is obvious for RMSSD and PNN50. In the following equation for SDNN, $\mu = \frac{1}{m}\sum_{i\le m} x_i$ is the mean, and we invoke the usual decomposition of variance.

$$\begin{aligned}
\text{RMSSD}^2 =& \frac{1}{m}\sum_{i<m}(x_{i+1}-x_i)^2 = \frac{1}{m}\sum_{i<m} z_{i,i+1} \\
m^2\cdot\text{SDNN}^2 =& m\sum_{i\le m}(x_i-\mu)^2 = m\sum_{i\le m} x_i^2 - m^2\mu^2 \\
=& \frac{1}{2}\sum_{i,j} x_i^2 + x_j^2 - 2x_ix_j = \frac{1}{2}\sum_{i,j} z_{i,j} \\
\text{PNN50} =& \frac{1}{m}\sum_{i\le m}\mathbf{1}\left(|x_{i+1}-x_i|\ge 50\right) = \frac{1}{m}\sum_{i\le m}\mathbf{1}\left(z_{i,i+1}\ge 50^2\right)
\end{aligned}$$

Put another way, the squared differences are sufficient statistics for HRV. However, these statistics do not capture all the information in the heartbeat signal. This restriction is beneficial because if the (purportedly) parasympathetic model overlaps too much with remaining layers, it becomes less interpretable, and may fail its purpose of isolating a useful residual. We will soon discuss why squared differences, rather than the more typical mean and moments x_ix_j, are being used.

Generalized HRV metric. The architecture of this part is as follows. For each window of size m, with stride 1, compute z. Optionally apply a nonlinearity for statistics like PNN50. (We do not do this.) Then, take a linear combination of the entries or z. We initialize this to the constant $1/(2m^2)$ matrix, to compute SDNN. We use a window size of $m = 24$, which is considered an ultra-short-term measure of HRV.

Parasympathetic contribution. Let us momentarily ignore the unit conversion h. To make $x - p(v)$ difficult to predict from v, we pretrain p to predict x from v, by minimizing mean squared error. In this way, $\tilde{x}$ becomes the unpredictable part of x. We expect this prediction will be difficult and that $x - p$ will be substantial. The HRV metrics could have also been written in terms of the mean and second moments x_ix_j, but this data would allow x to be easily recovered. The squared differences do not reveal much about the mean of x. This makes $p(v(z(\cdot)))$ a kind of autoencoder.

Now let us examine the use of h. Because x and p have units of RR intervals, there is an implicit bias for p to more closely fit low-intensity periods, which is when we expect more parasympathetic activity. This is because when HR is high, RR intervals are small, so the squared error is limited. For the sympathetic layer, we want the opposite numerical tendency, so we use units of heart rate. Computing $h(x - p)$ leads to near division-by-zero where x is close to p. Instead, we use $h(x) - h(p)$.

Finally, we discuss the actual architecture of $p(v)$. Since the parasympathetic system is high-frequency and fast-acting, we use 3 layers of convolutions of size 4 with linear activation.

4.4. Sympathetic Layer

The sympathetic layer is designed to be sensitive to the sharp increases and prolonged elevations coincident with sympathetic response. These sequential, spatially-local patterns can be recognized by one-dimensional convolutions. To increase the receptive field (i.e. to allow for some amount of sequential dependence), we employ dilated convolutions (Oord et al., 2016). We stack a block of dilated convolutions, each followed by a standard convolution with a stride length of 2, which halves the output dimension. To mute the output in uninteresting regions, we use gated convolutions, which doubles the number of convolution parameters (Dauphin et al., 2017). Together, the convolution layers output a single response value for each segment The cumulative sympathetic response is just the sum of the responses for each segment. This locality assumption is plausible, but could be reexamined in future work.

Compared to the parasympathetic layer, the sympathetic layer has a generic architecture. This is appropriate because the theory of how the sympathetic nervous system affects heartbeat intervals is much less well-developed. Though the chosen architecture seems to make the sympathetic layer work well, it is possible that other choices may work better.

4.5. Implementation details

We split the dataset into training and evaluation sets of equal size. To avoid imbalances, we reject splits where the means of ΔAmylase differ by more than 5. (Due to the small size of the dataset, this is a potential concern.) We pad or left-truncate the variable-length sequences to a uniform length of 7000.

We implement our model in TensorFlow 1.13. For both pretraining and training, we use the Adam optimizer with the default learning rate. For pretraining, we use the entire training set in each batch; for training, we use just a single example. We run both pretraining and training for 60,000 steps. During training, we freeze the parasympathetic layer's parameters. For both the regression and comparison problems, we use ℓ_1 loss. To examine the potential of overparametrization, we informally varied the number of parameters in the sympathetic layer.

4.6. Related Work

Most work on classifying cardiologic signals starts with nearly-continuous ECG waveforms sampled at approximately 200Hz (Hannun et al., 2019; Lehman et al., 2018). The goal is often to detect arrythmia or determine risk of myocardial infarction. We use coarser RR interval data for the following reasons. It isn't possible to access raw waveforms from consumer-grade monitors; the Bluetooth standard provides only for heart rate and RR intervals. During intense exercise, the waveform is likely to be extremely noisy, since even the extracted RR intervals are noisy. Lastly, sympathetic response is relatively slow compared to parasympathetic response and other ECG dynamics, so the time scale of RR intervals is more appropriate.

Most algorithms for assessing autonomic function from RR intervals originate from the field of signal processing. The algorithms are handcrafted, not learned, for the sake of simplicity, interpretability, and computational efficiency. Even though RR interval data is

relatively abundant, machine learning algorithms which process them are somewhat uncommon. In 2002, the PhysioNet challenge involved generating artificial RR interval sequences and discriminating them from real ones (Moody, 2002). Tsipouras et al. (2005) classify heartbeats using a handcrafted classifier which sequentially operates on windows of RR intervals. Gjoreski et al. (2017) employ a 7-layer fully-connected ReLU network upon the raw RR intervals. (Faust et al., 2018) apply an LSTM to detect atrial fibrillation. Asl et al. (2012) extracts time-series features before using a neural network.

Our extraction of a sympathetic residual should not be confused with (and in fact is diametrically opposed to) residual networks in deep learning (He et al., 2016). The layers of these networks have skip connections which add the original input to their transformation of the input. With a skip connections, the parasympathetic layer would output $x + p$ rather than just p. Skip connections seem to ease optimization because, if many such layers are stacked, then each individual layer can be very close to an identity transformation, with each p modeling a slight change (or "residual") of the input. By contrast, the point of our approach is to eliminate the extraneous parasympathetic component from the original input.

Various consumer devices calculate proprietary scores for workouts. Some of these scores purportedly measure the response to intense, anaerobic exercise. These scores encode some intuitions about exercise intensity and may have some motivational purpose. Since they do not correspond to any real physiologic quantity, it is difficult to assess the validity of these scores relative to a gold standard. It is also difficult to assess the relevance of these scores to health outcomes. It is easier to reason about the outcomes associated with amylase response, since it is part of the neuroendocrine system.

5. Machine Learning Results

First, we examine the results of pretraining. The quantitative error incurred during pretraining is not pertinent, so we examine some of the qualitative aspects of the learned features. In Figure 8, we see that parasympathetic contribution roughly accords with SDNN in Figure 6, but is more sparse. In Figure 9, we see that the learned HRV metric is substantially different than the known ones. Now we examine the accuracy of the algorithm on the regression and comparison problems. As a baseline method, we consider a plain CNN, whose architecture is the same as the sympathetic layer. We also consider what happens if we didn't pretrain the parasympathetic layer, but merely subtracted RMSSD from HR. For the comparison problem, we consider using maximum heart rate as a ranking. We find that our algorithm is superior to all of these baseline approaches. This suggests that both our modeling effort and pretraining is worthwhile. However, the accuracy of all the methods is still relatively poor. We believe this is simply due to the small amount of training data.

6. Discussion

Before discussing the our results, we recognize the following limitation: *the validity of ΔAmylase as a marker of exercise intensity is outside the scope of our study.* This question is examined by previous works mentioned in the introduction. Importantly, the validity of ΔAmylase may depend on the patient population. For example, cardiac rehabilitation

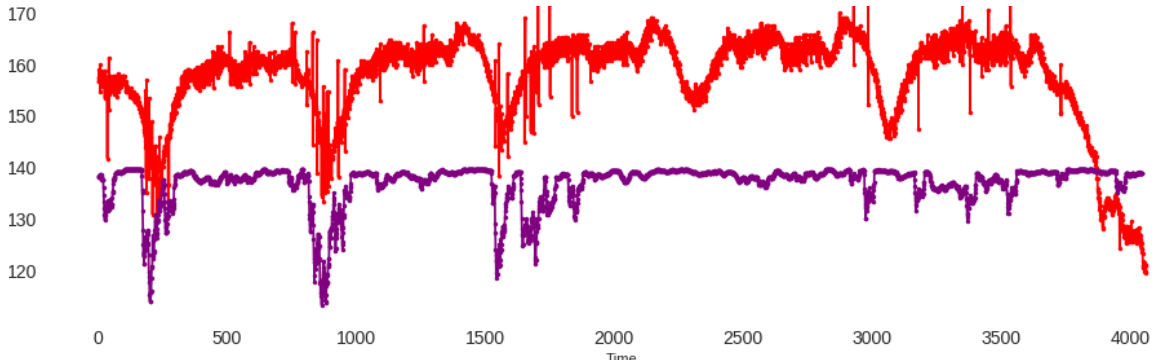

Figure 8: Example of heart rate $60000/x$ in red and parasympathetic contribution $60000/p$ in purple.

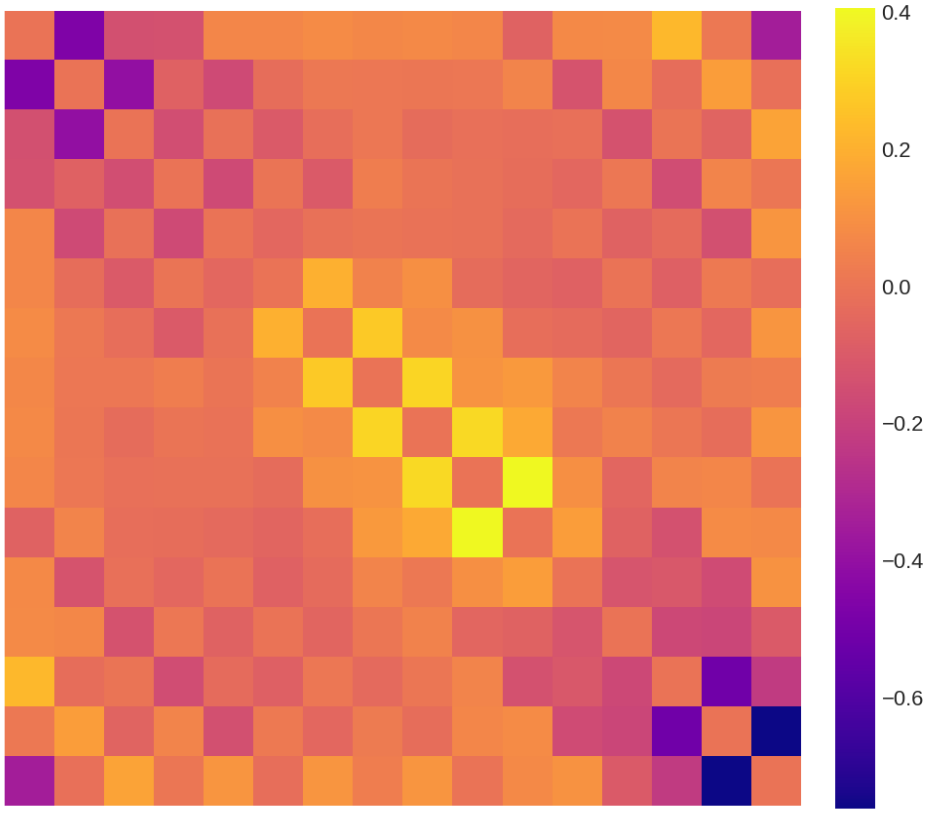

Figure 9: The matrix of weights on the squared differences $z_{i,j}$. Each entry was initialized to the same value, but after training, there is clearly nonuniformity. The algorithm seems to take advantage of the flexibility afforded by the additional parameters, with a nontrivial weight pattern, and both positive and negative weights.

Algorithm	Training Regression Error	Evaluation Regression Error
Plain CNN	8.02	26.29
No Pretrained PS	12.28	19.21
Our Algorithm	10.04	15.12

Algorithm	Training Comparison Error	Evaluation Comparison Error
Max HR	0.48	0.52
Plain CNN	0.11	0.53
No Pretrained PS	0.37	0.41
Our Algorithm	0.28	0.34

Figure 10: Quantitative evaluation of the algorithm. Our algorithm is superior in both the regression and comparison problems.

patients are often prescribe β-receptor antagonists, which inhibit both heart rate and likely amylase response. We are presently focused on the purely quantitative prediction task.

Our main result is that Δ*Amylase is (weakly) discernible solely from heartbeat data.* This is supported by the nontrivial regression and comparison accuracy of our model. A secondary result is that *it seems easier to predict* Δ*Amylase after attempting to subtract parasympathetic contribution.* It is possible that that this is simply due to the larger number of parameters in the pretrained model. However, this is unlikely, since the sympathetic layer already has a large number of parameters, and its performance was not substantially affected by adding more parameters.

Considering the small size of our dataset, this work should be considered a pilot study. We identify, formalize, and initiate study of a machine learning problem, but do not adequately solve it. Accordingly, we offer a methodological suggestions for future studies: *Point-of-care devices are not recommended for large-scale use.* Delicate use of the devices is necessary, and probably would not occur without careful instruction and/or supervision.

7. Conclusions and Future Work

This paper made the following contributions. (1) We initiated the study of a new supervised learning problem. (2) We collect a dataset large enough to conduct a pilot study. (3) We generalize HRV metrics to allow them to express parasympathetic contribution to heart rate. (4) We train and evaluate a physiologically-informed machine learning model. It outperforms baseline methods used by practitioners.

Due to the relatively small size of our dataset, our quantitative results should be considered preliminary. We are interested to see if they generalize to a larger-scale study.

References

Inger-Lise Aamot, Siv Hege Forbord, Trine Karlsen, and Asbjørn Støylen. Does rating of perceived exertion result in target exercise intensity during interval training in cardiac

rehabilitation? a study of the borg scale versus a heart rate monitor. *Journal of science and medicine in sport*, 17(5):541–545, 2014.

Kazunori Akizuki, Syouichirou Yazaki, Yuki Echizenya, and Yukari Ohashi. Anaerobic threshold and salivary α-amylase during incremental exercise. *Journal of physical therapy science*, 26(7):1059–1063, 2014.

Babak Mohammadzadeh Asl, Ahmad R Sharafat, and S Kamaledin Setarehdan. An adaptive backpropagation neural network for arrhythmia classification using rr interval signal. *Neural Network World*, 22(6):535, 2012.

Romeo B Batacan, Mitch J Duncan, Vincent J Dalbo, Patrick S Tucker, and Andrew S Fenning. Effects of high-intensity interval training on cardiometabolic health: a systematic review and meta-analysis of intervention studies. *Br J Sports Med*, 51(6):494–503, 2017.

Olga L Bocanegra, Miguel M Diaz, Renata R Teixeira, Silvio S Soares, and Foued S Espindola. Determination of the lactate threshold by means of salivary biomarkers: chromogranin a as novel marker of exercise intensity. *European journal of applied physiology*, 112(9):3195–3203, 2012.

Silke Boettger, Christian Puta, Vikram K Yeragani, Lars Donath, Hans-Josef Mueller, Holger HW Gabriel, and Karl-Juergen Baer. Heart rate variability, qt variability, and electrodermal activity during exercise. *Medicine & science in sports & exercise*, 42(3): 443–448, 2010.

Jos A. Bosch, Enno C.I. Veerman, Eco J. de Geus, and Gordon B. Proctor. -amylase as a reliable and convenient measure of sympathetic activity: dont start salivating just yet! *Psychoneuroendocrinology*, 36(4):449 – 453, 2011. ISSN 0306-4530. doi: https://doi.org/10.1016/j.psyneuen.2010.12.019. URL http://www.sciencedirect.com/science/article/pii/S0306453011000072.

Martin Buchheit, Paul B Laursen, and Saïd Ahmaidi. Parasympathetic reactivation after repeated sprint exercise. *American journal of physiology-heart and circulatory physiology*, 293(1):H133–H141, 2007.

Felipe Calvo, José L Chicharro, Fernando Bandrés, Alejandro Lucía, Margarita Pérez, Julián Álvarez, Luis L Mojares, Almudena F Vaquero, and Julio C Legido. Anaerobic threshold determination with analysis of salivary amylase. *Canadian Journal of Applied Physiology*, 22(6):553–561, 1997.

Daniel L Carl, Pierce Boyne, Bradley Rockwell, Myron Gerson, Jane Khoury, Brett Kissela, and Kari Dunning. Preliminary safety analysis of high-intensity interval training (hiit) in persons with chronic stroke. *Applied physiology, nutrition, and metabolism*, 42(3): 311–318, 2016.

José L Chicharro, Alejandro Lucía, Margarita Pérez, Almudena F Vaquero, and Rosario Ureña. Saliva composition and exercise. *Sports medicine*, 26(1):17–27, 1998.

Yann N Dauphin, Angela Fan, Michael Auli, and David Grangier. Language modeling with gated convolutional networks. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 933–941. JMLR. org, 2017.

Zuzana de Jong, Marten Munneke, Aeilko H. Zwinderman, Herman M. Kroon, Annemarie Jansen, Karel H. Ronday, Dirkjan van Schaardenburg, Ben A. C. Dijkmans, Cornelia H. M. Van den Ende, Ferdinand C. Breedveld, Theodora P. M. Vliet Vlieland, and Johanna M. W. Hazes. Is a long-term high-intensity exercise program effective and safe in patients with rheumatoid arthritis?: Results of a randomized controlled trial. *Arthritis & Rheumatism*, 48(9):2415–2424, 2003. doi: 10.1002/art.11216. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/art.11216.

VN De Oliveira, A Bessa, RPMS Lamounier, MG De Santana, MT De Mello, and FS Espindola. Changes in the salivary biomarkers induced by an effort test. *International journal of sports medicine*, 31(06):377–381, 2010.

James L Devin, Michelle M Hill, Marina Mourtzakis, Joe Quadrilatero, David G Jenkins, and Tina L Skinner. Acute high intensity interval exercise reduces colon cancer cell growth. *The Journal of physiology*, 2019.

Robert Edelberg. Effect of vasoconstriction on galvanic skin response amplitude. *Journal of applied physiology*, 19(3):427–430, 1964.

Øyvind Ellingsen, Martin Halle, Viviane Conraads, Asbjørn Støylen, Håvard Dalen, Charles Delagardelle, Alf-Inge Larsen, Torstein Hole, Alessandro Mezzani, Emeline M Van Craenenbroeck, et al. High-intensity interval training in patients with heart failure with reduced ejection fractionclinical perspective. *Circulation*, 135(9):839–849, 2017a.

Oyvind Ellingsen, Martin Halle, Eva Prescott, and Axel Linke. Response by ellingsen et al to letters regarding article, "high-intensity interval training in patients with heart failure with reduced ejection fraction". *Circulation*, 136(6):611–612, 2017b. doi: 10.1161/CIRCULATIONAHA.117.029145. URL https://www.ahajournals.org/doi/abs/10.1161/CIRCULATIONAHA.117.029145.

Adrian D Elliott, Kanchani Rajopadhyaya, David J Bentley, John F Beltrame, and Edoardo C Aromataris. Interval training versus continuous exercise in patients with coronary artery disease: a meta-analysis. *Heart, Lung and Circulation*, 24(2):149–157, 2015.

Oliver Faust, Alex Shenfield, Murtadha Kareem, Tan Ru San, Hamido Fujita, and U Rajendra Acharya. Automated detection of atrial fibrillation using long short-term memory network with rr interval signals. *Computers in biology and medicine*, 102:327–335, 2018.

James P Fisher, Ahmed M Adlan, Alena Shantsila, J Frederik Secher, Henrik Sørensen, and Niels H Secher. Muscle metaboreflex and autonomic regulation of heart rate in humans. *The Journal of physiology*, 591(15):3777–3788, 2013.

Martin Gjoreski, Hristijan Gjoreski, Mitja Luštrek, and Matjaž Gams. Deep affect recognition from rr intervals. In *Proceedings of the 2017 ACM International Joint Conference*

on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers, pages 754–762. ACM, 2017.

Guido Grassi and Murray Esler. How to assess sympathetic activity in humans. *Journal of hypertension*, 17(6):719–734, 1999.

Amanda L Hannan, Wayne Hing, Vini Simas, Mike Climstein, Jeff S Coombes, Rohan Jayasinghe, Joshua Byrnes, and James Furness. High-intensity interval training versus moderate-intensity continuous training within cardiac rehabilitation: a systematic review and meta-analysis. *Open access journal of sports medicine*, 9:1, 2018.

Awni Y Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H Tison, Codie Bourn, Mintu P Turakhia, and Andrew Y Ng. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature medicine*, 25(1):65, 2019.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. In *European conference on computer vision*, pages 630–645. Springer, 2016.

Charlotte Jelleyman, Thomas Yates, Gary O'Donovan, Laura J Gray, James A King, Kamlesh Khunti, and Melanie J Davies. The effects of high-intensity interval training on glucose regulation and insulin resistance: a meta-analysis. *Obesity reviews*, 16(11):942–961, 2015.

Charlotte Lauren Jelleyman. *High-intensity physical activity for improving glucose regulation: can science justify IT?* PhD thesis, Department of Cardiovascular Sciences, 2018.

Prince J Kannankeril, Francis K Le, Alan H Kadish, and Jeffrey J Goldberger. Parasympathetic effects on heart rate recovery after exercise. *Journal of investigative medicine*, 52(6):394–401, 2004.

Holly S Kessler, Susan B Sisson, and Kevin R Short. The potential for high-intensity interval training to reduce cardiometabolic disease risk. *Sports medicine*, 42(6):489–509, 2012.

Eri Koibuchi and Yoshio Suzuki. Exercise upregulates salivary amylase in humans. *Experimental and therapeutic medicine*, 7(4):773–777, 2014.

Docent Juhani Koistinen and Docent Tomi Laitinen. Effect of physical exercise on autonomic regulation of heart rate. 2004.

Eric P. Lehman, Rahul G. Krishnan, Xiaopeng Zhao, Roger G. Mark, and Li-wei H. Lehman. Representation learning approaches to detect false arrhythmia alarms from ecg dynamics. In Finale Doshi-Velez, Jim Fackler, Ken Jung, David Kale, Rajesh Ranganath, Byron Wallace, and Jenna Wiens, editors, *Proceedings of the 3rd Machine Learning for Healthcare Conference*, volume 85 of *Proceedings of Machine Learning Research*, pages 571–586, Palo Alto, California, 17–18 Aug 2018. PMLR. URL http://proceedings.mlr.press/v85/lehman18a.html.

TL Li and Michael Gleeson. The effect of single and repeated bouts of prolonged cycling and circadian variation on saliva flow rate, immunoglobulin a and-amylase responses. *J sports Sci*, 22(11-12):1015–1024, 2004.

NEAL Lippman, KENNETH M Stein, and BRUCE B Lerman. Comparison of methods for removal of ectopy in measurement of heart rate variability. *American Journal of Physiology-Heart and Circulatory Physiology*, 267(1):H411–H418, 1994.

Theresa Mann, Robert Patrick Lamberts, and Michael Ian Lambert. Methods of prescribing relative exercise intensity: physiological and practical considerations. *Sports medicine*, 43(7):613–625, 2013.

Errol B Marliss and Mladen Vranic. Intense exercise has unique effects on both insulin release and its roles in glucoregulation: implications for diabetes. *Diabetes*, 51(suppl 1): S271–S283, 2002.

Scott Michael, Kenneth S Graham, and Glen M Davis. Cardiac autonomic responses during exercise and post-exercise recovery using heart rate variability and systolic time intervalsa review. *Frontiers in physiology*, 8:301, 2017a.

Scott Michael, Ollie Jay, Kenneth S Graham, and Glen M Davis. Higher exercise intensity delays postexercise recovery of impedance-derived cardiac sympathetic activity. *Applied Physiology, Nutrition, and Metabolism*, 42(8):834–840, 2017b.

Scott Michael, Ollie Jay, Kenneth S Graham, and Glen M Davis. Longer exercise duration delays post-exercise recovery of cardiac parasympathetic but not sympathetic indices. *European journal of applied physiology*, 117(9):1897–1906, 2017c.

Zoran Milanović, Goran Sporiš, and Matthew Weston. Effectiveness of high-intensity interval training (hit) and continuous endurance training for vo 2max improvements: a systematic review and meta-analysis of controlled trials. *Sports medicine*, 45(10):1469–1481, 2015.

Jeffrey P Moak, David S Goldstein, Basil A Eldadah, Ahmed Saleem, Courtney Holmes, Sandra Pechnik, and Yehonatan Sharabi. Supine low-frequency power of heart rate variability reflects baroreflex function, not cardiac sympathetic innervation. *Heart Rhythm*, 4(12):1523–1529, 2007.

GB Moody. Rr interval time series modeling: the physionet/computers in cardiology challenge 2002. In *Computers in Cardiology*, pages 125–128. IEEE, 2002.

Urs M Nater and N Rohleder. Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: current state of research. *Psychoneuroendocrinology*, 34(4): 486–496, 2009.

Aaron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. Wavenet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499*, 2016.

Henry T Peng, Erin Savage, Oshin Vartanian, Shane Smith, Shawn G Rhind, Catherine Tenn, and Stephen Bjamason. Performance evaluation of a salivary amylase biosensor for stress assessment in military field research. *Journal of clinical laboratory analysis*, 30(3): 223–230, 2016.

Hugo F Posada-Quintero, Natasa Reljin, Craig Mills, Ian Mills, John P Florian, Jaci L VanHeest, and Ki H Chon. Time-varying analysis of electrodermal activity during exercise. *PloS one*, 13(6):e0198328, 2018.

Gustavo A Reyes del Paso, Wolf Langewitz, Lambertus JM Mulder, Arie Van Roon, and Stefan Duschek. The utility of low frequency heart rate variability as an index of sympathetic cardiac tone: a review with emphasis on a reanalysis of previous studies. *Psychophysiology*, 50(5):477–487, 2013.

Nicolas Rohleder and Urs M Nater. Determinants of salivary α-amylase in humans and methodological considerations. *Psychoneuroendocrinology*, 34(4):469–485, 2009.

Fred Shaffer and JP Ginsberg. An overview of heart rate variability metrics and norms. *Frontiers in public health*, 5:258, 2017.

Fred Shaffer, Rollin McCraty, and Christopher L Zerr. A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. *Frontiers in psychology*, 5:1040, 2014.

Vivek Shetty, Corwin Zigler, Theodore F Robles, David Elashoff, and Masaki Yamaguchi. Developmental validation of a point-of-care, salivary α-amylase biosensor. *Psychoneuroendocrinology*, 36(2):193–199, 2011.

J. Kevin Shoemaker, Stephen A. Klassen, Mark B. Badrov, and Paul J. Fadel. Fifty years of microneurography: learning the language of the peripheral sympathetic nervous system in humans. *Journal of Neurophysiology*, 119(5):1731–1744, 2018. doi: 10.1152/jn.00841.2017. URL https://doi.org/10.1152/jn.00841.2017. PMID: 29412776.

Theresa A Strzelczyk, Rebecca J Quigg, Pamela B Pfeifer, Michele A Parker, and Philip Greenland. Accuracy of estimating exercise prescription intensity in patients with left ventricular systolic dysfunction. *Journal of Cardiopulmonary Rehabilitation and Prevention*, 21(3):158–163, 2001.

Hirofumi Tanaka, Kevin D Monahan, and Douglas R Seals. Age-predicted maximal heart rate revisited. *Journal of the american college of cardiology*, 37(1):153–156, 2001.

Bianca Lee Thomas, Nicolaas Claassen, Piet Becker, and Margaretha Viljoen. Validity of commonly used heart rate variability markers of autonomic nervous system function. *Neuropsychobiology*, 78(1):14–26, 2019.

Markos G Tsipouras, Dimitrios I Fotiadis, and D Sideris. An arrhythmia classification system based on the rr-interval signal. *Artificial intelligence in medicine*, 33(3):237–250, 2005.

Jessica L Unick, Sarah Gaussoin, Judy Bahnson, Richard Crow, Jeff Curtis, Tina Killean, Judith G Regensteiner, Kerry J Stewart, Rena R Wing, John M Jakicic, et al. Validity of ratings of perceived exertion in patients with type 2 diabetes. *Journal of novel physiotherapy and physical rehabilitation*, 1(1), 2014.

Ake B. Vallbo, Karl-Erik Hagbarth, and B. Gunnar Wallin. Microneurography: how the technique developed and its role in the investigation of the sympathetic nervous system. *Journal of Applied Physiology*, 96(4):1262–1269, 2004. doi: 10.1152/japplphysiol.00470.2003. URL https://doi.org/10.1152/japplphysiol.00470.2003. PMID: 15016790.

Marieke van Dooren, Joris H Janssen, et al. Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology & behavior*, 106(2):298–304, 2012.

Kassia S Weston, Ulrik Wisløff, and Jeff S Coombes. High-intensity interval training in patients with lifestyle-induced cardiometabolic disease: a systematic review and meta-analysis. *Br J Sports Med*, 48(16):1227–1234, 2014.

Michael A Wewege, Dohee Ahn, Jennifer Yu, Kevin Liou, and Andrew Keech. High-intensity interval training for patients with cardiovascular diseaseis it safe? a systematic review. *Journal of the American Heart Association*, 7(21):e009305, 2018.

Daniel W White and Peter B Raven. Autonomic neural control of heart rate during dynamic exercise: revisited. *The Journal of physiology*, 592(12):2491–2500, 2014.