Cross-Modal Health State Estimation

Nitish Nag
University of California, Irvine
Irvine, California
nagn@uci.edu

Vaibhav Pandey
University of California, Irvine
Irvine, California
vaibhap1@uci.edu

Preston J. Putzel
University of California, Irvine
Irvine, California
pputzel@uci.edu

Hari Bhimaraju
University of California, Irvine
Irvine, California
haribhimaraju@gmail.com

Srikanth Krishnan
University of California, Los Angeles
Los Angeles, California
srikanthkrishnan@mednet.ucla.edu

Ramesh Jain
University of California, Irvine
Irvine, California
jain@ics.uci.edu

ABSTRACT

Individuals create and consume more diverse data about themselves today than any time in history. Sources of this data include wearable devices, images, social media, geo-spatial information and more. A tremendous opportunity rests within cross-modal data analysis that leverages existing domain knowledge methods to understand and guide human health. Especially in chronic diseases, current medical practice uses a combination of sparse hospital based biological metrics (blood tests, expensive imaging, etc.) to understand the evolving health status of an individual. Future health systems must integrate data created at the individual level to better understand health status perpetually, especially in a cybernetic framework. In this work we fuse multiple user created and open source data streams along with established biomedical domain knowledge to give two types of quantitative state estimates of cardiovascular health. First, we use wearable devices to calculate cardiorespiratory fitness (CRF), a known quantitative leading predictor of heart disease which is not routinely collected in clinical settings. Second, we estimate inherent genetic traits, living environmental risks, circadian rhythm, and biological metrics from a diverse dataset. Our experimental results on 24 subjects demonstrate how multi-modal data can provide personalized health insight. Understanding the dynamic nature of health status will pave the way for better health based recommendation engines, better clinical decision making and positive lifestyle changes.

CCS CONCEPTS

• **Information systems** → **Multimedia information systems**; **Multimedia and multimodal retrieval**; • **Social and professional topics** → **Personal health records**; • **Applied computing** → **Systems biology**; **Consumer health**; **Health care information systems**; **Health informatics**; **Bioinformatics**; • **Computer systems organization** → **Embedded and cyber-physical**
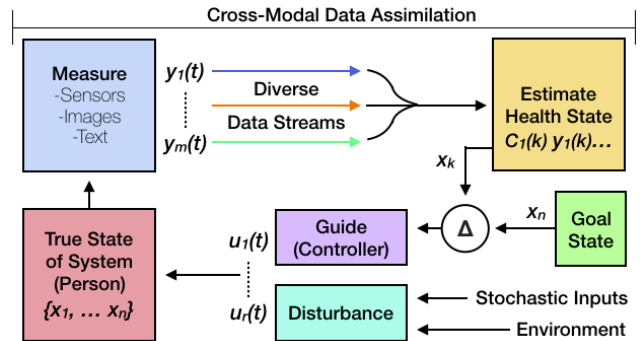

Figure 1: Cross-modal measurements are essential for state estimation in cybernetic feedback systems. This estimation impacts the eventual guidance from the controller (physician or automated system) to help reach a goal state.

systems; • **General and reference** → *Estimation*; • **Networks** → *Cyber-physical networks*;

KEYWORDS

Personal Health Navigation; Cross-Modal Data; ; Health Situation; Cybernetic Health; Multimedia; Wearables; Health Informatics

1 INTRODUCTION

"To live effectively is to live with adequate information."

- Norbert Weiner, 1950

A century ago, the largest contributor to mortality and morbidity was infectious disease. Infection is an episodic problem, where a categorical diagnosis (i.e. malaria) is made after a patient feels unwell and arrives at a medical care facility (where the data is gathered to confirm the disease). Treatment is usually prescribed based on evidence based rules to solve the problem, and the patient is not monitored anymore. Globally, chronic diseases have emerged as the 21st century major contributor to health burden. When compared to infectious disease, there are fundamental differences. There is
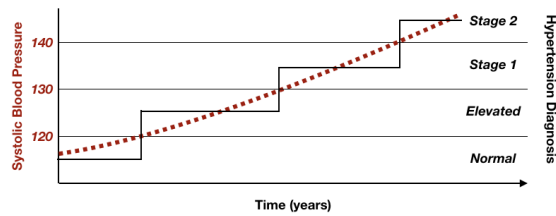
Figure 2: True physiological average blood pressure (red dashed) vs. clinical assessment (black solid) show the comparison between health state assessments that are quantitative vs. categorical.

no single event that leads to the disease, rather slow changes in the operating function of the body. If we are to apply some type of control input to keep people on a healthy trajectory via feedback loops in cybernetic systems, we must be able to continuously estimate this state. Hence there is a clear distinction between the classification versus quantified estimation problem in health. Due to this reason, there is compelling need for progress in health state estimation.

To illustrate this need we describe the situation of hypertension (high blood pressure). First, patients are unable to feel the disease as it slowly builds up over time, and thus do not even know they are being affected. Second, the diseases stem from both daily actions and environmental exposures, not just a single source. Third, these diseases are not truly categorical in nature, but are rather declines in organ function over time. In the example of hypertension, clinical practice uses cutoff thresholds to decide when to change the labeled blood pressure status of an individual, when in reality, the average pressure is increasing over time as shown in Figure 2. Ultimately, individuals, clinicians, and in general cybernetic systems (Figure 1), make decisions based on the method of determining health state. What we measure is what we control.

Individuals create diverse data streams about themselves on a daily basis. Much of this data can be leveraged to provide perpetual insight into the health of individuals. Fitness devices and wearable sensors, ambient sensors, images, video, audio, digital human computer interactions, and IoT devices provide a plethora of data that is routinely collected, but so far has been difficult to use for common real world health applications. Because people are unable to feel their health change over time from the multitude of factors affecting them, we need to develop methods to quantify and report health status using continuously collected multi-modal data sources. If we are able to track changes before permanent organ dysfunction, we may be able to correct course and prevent or delay onset of chronic diseases. Finally, medical practice needs to shift from using episodic categorical definitions of health status, to a continuous quantitative measurement.

With the rapidly increasing availability of low cost sensors in the last decade, there has been an explosion in the amount of continuously collected multi-modal data. This is especially relevant in field of health with the advent of wearable, IoT, and ambient sensors. Many of these low cost sensors such as accelerometers, light sensors, microphones, heart rate monitors, and barometers, to name a few, produce continuous streams of data usable in a wide range of scenarios. While this allows measurement of the user state

continuously, the downside of these sensors is that the measurements are either very noisy, produce information overload, produce non-actionable metrics, or are not directly related to the attribute we want to measure. These limitations have so far proven to be a barrier in using low cost sensors for real world decision making for health.

Furthermore, enabling data assimilation will require intelligible understanding of how sensor information relates to the health status. Existing medical and biological scientific domain knowledge must be used to guide the data assimilation and conversion matrices from signal to state. Intelligibility is also an important attribute for developing health estimation systems. Knowing *why* an individual has a certain health status will be paramount to explanation, recommendation, and treatment. Data driven methods can be components in a large system, but will have difficult time explaining the reason for the classification. This is why other complimentary methods need to be used that take advantage of domain knowledge.

To summarize, we believe that cross-modal health state estimation will be a fundamental centerpiece for the following needs of future:

- **Early Detection:** Insight into health status changes in the prodromal state, where health state can be readily altered towards wellness.
- **Continuous Monitoring:** Understanding and assisting the individual in all aspects of life, everyday.
- **Quantitative Real-Time Assessment:** Shifting health assessment to a dynamic quantitative measurement rather than categories of normal versus abnormal.
- **Reduced Cost:** Through trickle down technology, we anticipate more data types available through devices, reducing the barriers and cost for health assessment.

Clinical need for measuring cardiorespiratory fitness (CRF) is in high demand, but at the moment it is only captured in high need care through expensive lab tests. In 2013, the American Heart Association and the American College of Cardiology jointly released guidelines for the prevention and treatment of coronary artery disease stating CRF is a leading risk factor for cardiovascular disease, the most significant cause of death in humans. Flatly stated by the AHA, "It is currently the only major risk factor not routinely assessed in clinical practice" [35]. The reason for not measuring this value for patients is due to the burdensome cost in time, inconvenience, and resources to gather this data directly. We take this as motivation to see if we can use lower cost wearable devices to accomplish this task. We compare how different wearable devices can provide observability into our own bodies. CRF levels change throughout our lives from effects of our lifestyle. A more refined and accurate reflection of cardiovascular health state would take into account additional information like the environment, stress, and genetic background. We address this challenge through adding additional data sources such as images and geospatial sensors.

For the aforementioned reasons, we focus the scope of our work on cardiovascular health state estimation by studying the following research questions:

- **Research Question 1 (RQ1):** What is the quality of CRF estimation via the combination of multi-modal data and domain knowledge from different wearable devices?

- **Research Question 2 (RQ2):** What total cardiovascular health information can we elucidate from images, wearables, surveys, social media, Internet of Things (IoT), and environmental sensors? How can we assimilate this data in a useful way for individuals and health providers?

At the time of writing, there have been no investigations we found about cardiovascular health state measurements in the mutlimedia research community. Broadly, the motivation for this work in the Brave New Ideas track is to open the frontier into personalized health state estimation from multi-modal data. Further rigorous research in this field will look into expanding to other health domains, improve quality metrics, tackle performance issues, and much more. We hope ultimately to create research opportunities that allow us to effectively be informed about our health throughout life.

2　RELATED WORK

Health state estimation and tracking has been an important field in medical literature and computer science. There has been a strong call by the medical science community to use continuous multimodal data for tracking individual health [29, 30, 36, 43, 48]. Most modern metrics that are used to understand patient health were derived from longitudinal studies of large cohorts to see what led to morbidity and mortality. Outcomes of these studies were then retrospectively analyzed with linear regression to predict future outcomes for new patients. Modern epidemiology efforts are beginning to use modern data collection tools such as social multimedia and wearable devices [14]. These efforts include the United States Precision Medicine Initiative led by President Obama [16], Mobile Sensor Data to Knowledge [23], and Alphabet's Verily division [12]. These research efforts may take decades before we have data available for meaningful insight, as they largely depend upon outcomes of mortality before they become sufficiently powerful.

Within the field of cardiology, the Framingham study laid the foundation for most modern clinical guidelines by the American Heart Association (AHA) and American College of Cardiology [18, 50]. AHA has also called for the specific metric of CRF as the most powerful predictor of cardiovascular health that is not routinely measured (mostly due to cost of expensive and laborious lab testing) [35]. Technically speaking, activity which measures general movement patterns (such as through wearable accelerometers) is a different risk factor than aerobic exercise work capacity (which is CRF). CRF has a much stronger established relationship with true cardiovascular health [49]. Widespread use of standard wearable accelerometers that measure steps or higher semantic activities like walking, jogging, biking are indicative of activity only, hence the need for wearables that can give estimates of CRF.

Wearable devices have been used to estimate energy expenditure through various computational approaches such as deep learning [51], knowledge based regression [25], and data filtering and segmentation techniques [1]. Energy expenditure provides insight into the total amount of activity performed by an individual, but does not provide maximal work output to estimate CRF.

Computational research in CRF prediction began in the 1970's with the formulation of exercise stress scoring metrics based on then newly available chest strap based heart rate monitors [7, 8]. Recently, contextual understanding improved the performance of heart rate based CRF estimation, and were further refined by calibrating custom algorithmic parameters for a particular user [2, 3]. Heart rate data has also been used to derive additional features, such as vagal tone (commonly referred to has heart rate variability) and respiratory rate, to provide regression analysis more features for prediction [41]. Improvements in accelerometer based CRF prediction have been achieved through body placement optimization [34]. The only known research at this point that has attempted use of multimodal data for CRF prediction has been done by Firstbeat Corporation which uses both heart rate and speed information with a proprietary algorithm to filter periods of heart rate that are indicative of steady state metabolism [15].

Multifactorial cardiovascular health risks have been investigated in many of the large epidemiologic studies such as the Framingham study. Conclusions from these large studies are used in current clinical practice through the AtheroSclerotic CardioVascular Diease (ASCVD) calculator [18]. This pooled cohort algorithm was based on linear regression analysis for four separate cohorts of individuals female blacks, male blacks, female whites, male whites. Other than ethnicity and gender, they take into account age, systolic and diastolic blood pressure, cholesterol (Total, HDL, LDL), smoking history, diabetes (binary field: yes or no), medication history (hypertension, statin, aspirin only), and is only applicable for patients in the age range of 40-79. The limitations of this calculator include the requirement of invasive blood data and non-consumer based lab processing. No integrations of environment, lifestyle, social determinants, or biological parameters that test real world function such as CRF are used in any clinical setting at the moment. Individual parameters such as local air and noise pollution have established as risks, but are not shown in any relevant way to clinicians. Wearable devices are not used in any clinical setting for cardiovascular disease presently.

3　CYBERNETIC HEALTH STATE ESTIMATION

In the simplest terms, cybernetics is about setting goals and devising action sequences to accomplish and maintain those goals in the presence of noise and disturbances [31]. This is enabled by the availability of sensors that can estimate the system state from observations to perpetually feed this information back to the system. This generates new control signals as required to move toward the desired goal or destination. Cybernetic in health has 4 main components: Measurement, Estimation, Guidance, Action as shown in figure 1 [20]. These four parts synthesize how we can produce a navigational system for improving health.

The mathematical model in classic systems theory states that:

$$X[k + 1] = A[k]X[k] + B[k]U[k]$$

$$Y[k] = C[k]X[k] + D[k]U[k]$$

Where X, U, and Y are the system true state, inputs, and measured output vectors respectively. A, B, C, and D are matrices that provide the appropriate transformation of these variables at a given time k. Human health can be described by a state system, and the previous state and the inputs into the system play a role in determining health at time $k+1$. Inputs into the human cybernetic system can be defined as anything which changes gene expression or physical actions in the body (from a molecular interactions to coarse movements). Thus

a body is continuously exposed to these inputs which may or may not be within the controllability of an individual. The inputs beyond the control of an individual are referred to as external disturbance and the rest can be viewed as controllable inputs u. The true health state of an individual at a time k is represented by X, which is in reality difficult to obtain and always estimated. What we do get are the observable output variables. The state estimation challenge is in interpreting the observables to understand the underlying true state. If we solely focus on this, the challenge of state estimation is represented in the matrix C, with observables as Y and our unknown state as X.

Estimating health impairment in individuals who seem to be healthy is inherently difficult due to: 1)poor sensing ability of developing adverse outcomes with current clinical methods and 2)the long lead time to developing full blown chronic disease. By the time current clinical measurements such as cholesterol, blood pressure, or glucose metabolism are beyond the normal range, the user has already been in a dysfunctional health state for quite some time. Capturing the change in health state earlier (before true dysfunction begins) is paramount to keeping people healthy and preventing them from slipping into a diseased state. Clinical researchers refer to this as the prodromal state. Multimedia work in understanding, vision, classifiers, intent, and sentiment analysis can greatly expand the capability for higher resolution understanding of an individuals health state. In our following experimental work, we focus on this specific aspect in the domain of cardiovascular health.

4 EXPERIMENTAL APPROACH

Transforming this data into semantically meaningful information is the first step in using data for an end goal. In the application of health, an additional step is needed to take these information bits into the domain of biological variables. A sufficient set of biological variables can provide an overview of how the human system is operating. This flow is shown in Figure 3.

4.1 RQ1: Estimation of CRF Bio-Variable

CRF represents the integrated biological performance of delivering oxygen from the atmosphere via the lungs and blood to the mitochondria to perform physical work (Work = Force x Distance). This essentially quantifies the functional capacity of the respiratory, cardiovascular, metabolic, and type-1 fiber musculature. CRF is usually measured through breath captured maximal oxygen consumption (VO2max) during a maximal exercise effort of several minutes. Because the efficiency of muscular work produced per unit oxygen consumed is directly related to the physics of adenosine triphosphate synthesis and breakdown to adenosine diphosphate, we can use power output on bicycle ergometry to directly calculate VO2max [26, 27]. VO2max is measured in mL oxygen consumed / minute / kilogram of bodyweight and is a direct measurement of CRF. For the purposes of this paper, they are equivalent.

We use a multi step process to extract meaningful information from the wearable devices (Figure (4). Forces against a bicycle motion in real world activity are divided into three main components: wind, gravity, and friction. Effort by a rider can be measured by duration of exercise or heart rate based effort. We use three methods to estimate power output, and test these methods against the ground truth of known power output from device 8.

Active Time Based Training Effect - TIME: We use devices 1 and 2 and instances where we only have accelerometer, time, or cadence data to estimate how much time the user is actively exercising. We base this estimate from the increased exercise volume (time) leading to increased CRF [21].

Heart Rate Based Training Effect - TRIMP: Devices 4,5,6, and 7 have heart rate sensors we use to predict not only exercise volume, but also intensity. Intensity of exercise is calculated by the established Training Impulse (TRIMP) method [7].

Work Against Gravity - VAM: We use devices 3,5,6,and 7 to give us both horizontal and vertical velocity. Devices 3,5 and 7 use GPS to give latitude,longitude and altitude. Device 6 uses a barometer for altitude and wheel magnet for horizontal velocity. Vertically Ascended Meters (VAM) is the z-axis velocity in meters/hour. For all instances where the rider is going uphill, we calculate the Newtonian physical work done against gravity. The horizontal velocity is less climbing uphill, and thus we assume a minimal component of wind resistance.

We test these estimation methods to compare performance in prediction of CRF with different situations of sensor derived information. First, we use a global prediction model by using a subset of 50% of subjects. We use an individual model for instances where a user would be given a calibration device of a power meter for a given period of time, and measure how well we can model future CRF prediction after the calibration device was removed. For both instances we use 70% of the data for training.

4.2 RQ2: Multi-Factorial Approach to Health

The true state of the cardiovascular system will depend upon many different controlled inputs and disturbances. In this section we provide an example of how varied these sources can be, and how they may be integrated to provide a dashboard of the cardiovascular state to a user or health expert. The total state of cardiovascular health may be summarized into some sub-states such as circulation, metabolism, stress, vascular perfusion, electrical activity, and valvular function for example. As shown in Figure 3, these sub-states are composed from evidence based relationships with bio-variables. We describe the methods used to gather and extrapolate these data relationships below:

Images of the user were used to derive several biological features. We use OpenCV facial landmarks detection to determine width-to-height ratio as shown in Figure 5 [32]. We calculate a proxy for genetic testosterone levels through this ratio [24]. Higher testosterone levels are positively correlated with better cardiovascular circulation and metabolism [33]. We also use the images for ethnicity detection and gender identification.

Location of living for each user was determined through histograms of GPS coordinates at the beginning of the activities. These locations were mapped to zip codes in the United States.

Environmental Zip code and county based average income, cardiac deaths, community crime risk [45], air [44], light[47], and noise pollution [46] data were then mapped to each user. Established relationships that affect cardiovascular health have been reported for PM2.5 air pollution [42], light pollution and noise pollution [28].

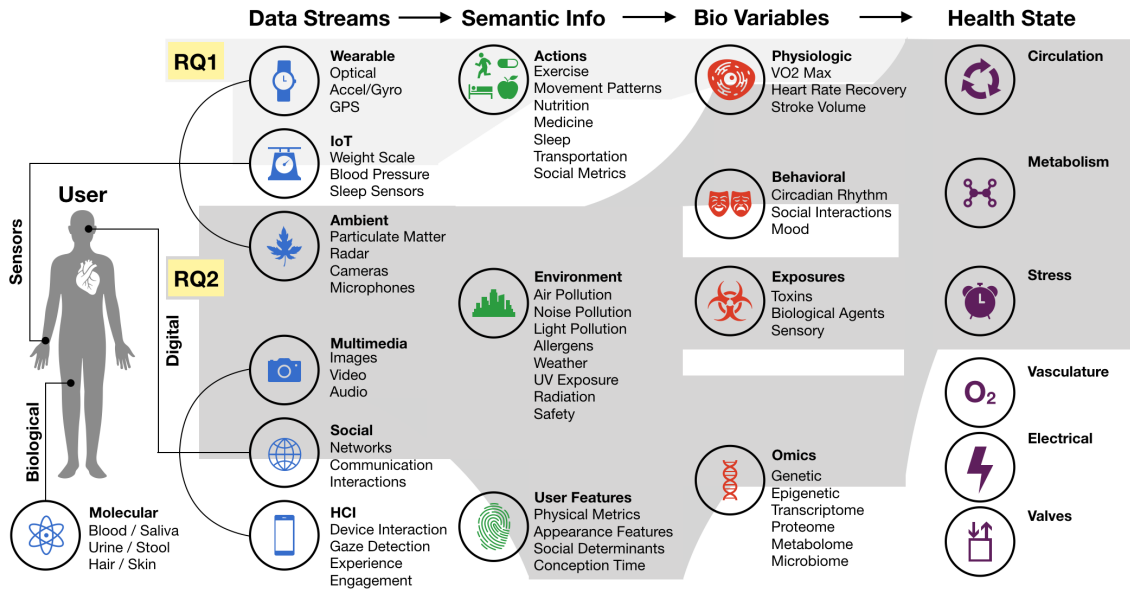**Circadian** light exposure during exercise were derived from physics

Figure 3: Intermediate steps of transforming cross-modal data into a bio-variables or health state (for cardiovascular health in this work). Research question 1 delves into taking wearable data streams and producing a single biological variable of VO2 Max (CRF equivalent). Research question 2 takes a much larger set of data streams to produce multiple biological variables that can then be used to approximate a health state of the individual. Unshaded additional data streams, information, biological variables, and health states are not addressed in detail but shown to demonstrate future potential.
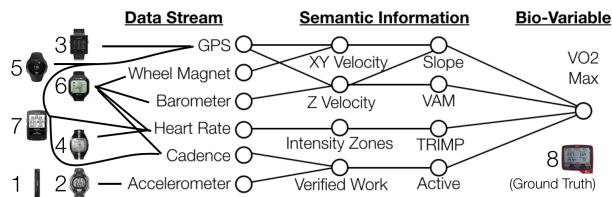

Figure 4: Wearable devices used for comparison in the experiments and respective feature extraction: 1. Timex Ironman, 2. Fitbit Flex2, 3. Garmin VivoActive, 4. Polar FT7, 5. Suunto Spartan, 6. Polar RCX5, 7. Garmin Edge 520, 8. SRM-PC8 (contains all sensors and used for ground truth).
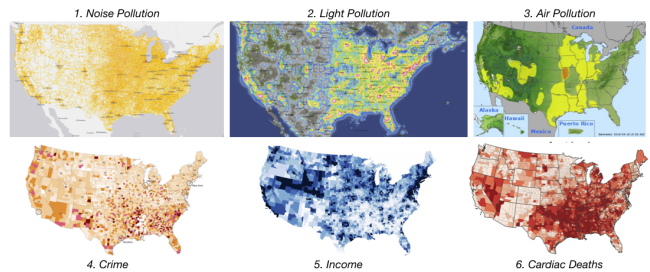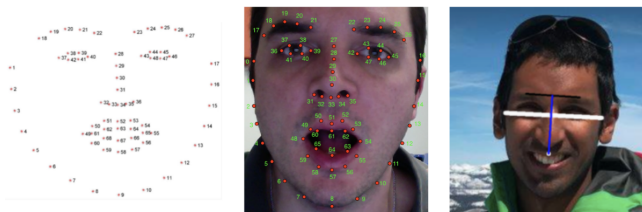

Figure 5: Images of people provide insight into their health state. Visual features of facial width-height ratio used as a validated proxy for basal genetic testosterone levels.

based models of earth rotation to determine natural light exposure. Patterns of weekly variability in exercise habits and time zone changes were also calculated from user data, and light pollution at the living location. Circadian disruption has been shown in humans

Figure 6: Visualization of various geo-spatial data sources used in our health state estimation.

to cause cardiovascular impairment [37].

Social Media networks were used to not only gather the images, but also professional status and educational attainment of the user from LinkedIn. The combination of education, professional status, and zip code average income was used to estimate financial status [6]. Education links to cardiovascular disease as a proxy for other risk factors have been studied [11, 22].

Surveys were given to users to obtain their age, smoking status, height, weight, and waist circumference. These measurements can also be automated with IoT devices, such as connected weight scales. Body Mass Index (BMI) [50] and Waist-to-Height Ratio (WHR) [4] were derived from these values.

Wearables (specifically device 8) from RQ1 were used to estimate bio-variables that have relationships with cardiovascular disease and heart functionality. These bio-variables include heart rate recovery [9], heart left ventricle stroke volume [5], heart rate drift[10],

kilojoules of work [19] in addition to the CRF, TRIMP, and active time.

Inherent ASCVD risk was calculated from a combination of established risk factors due to ethnicity, age, basal testosterone, and smoking status. This is an established risk of potential for a hard cardiovascular event [18] within the next 10 years for that individual. We define high friction risk as factors that require dramatic life change to alter (such as moving to a new home or acquiring a higher educational degree). In our case this relates to the variables of education attainment and income in addition to factors related to living location which include crime, local incidence of cardiac death, air, noise, and light pollution. Circadian Rhythm Disruption is a normalized sum of light pollution in living location, time zone changes (hours changed relative to GMT in last 4 weeks), and exercise habit variability (average exercise start time difference from previous day in last 4 weeks). Circulation capability of the heart is a normalized sum of CRF, HRR, SV, TRIMP in the last 4 weeks. These factors capture the ability of the heart to pump blood throughout the body. Metabolism summary was calculated as the normalized sum of exercise work (in kJ), HRD, active time, BMI and WHR. These factors capture the ability of the individual to maintain high resting basal metabolic rate and resist fatigue. These summaries does not reflect any absolute risk, just relative risk to others in our sample subjects. They are meant to observe how an individual's health status is longitudinally changing over time, or as a cross-sectional comparison to others in the same subject population. The overall heart score is an equal weighted average of both these relative metrics and the inherent ASCVD risk.

Multi-modal data assimilation and visualizations have been used extensively to maintain the health state of jet engines and other mechanical devices [40]. By placing various sensors on the engine, engineers and pilots are able to monitor the status of an engine in real-time and understand when to take precaution or perform an action to ensure the safety and longevity of the engine. We present a similar view of health data in Figure 9 for individual use and Figure 10 for professional/expert use.

5 EXPERIMENTAL RESULTS

The dataset used in the experiments includes sensor data streams at one second resolution from eight wearable devices collected on 24 male cycling athletes over an average of 5 years in the United States. Athletes also had strain gauges installed on their bicycles to measure physiologic true power output. The total dataset includes 31,776 activities and 70,178 hours of exercise data. Social media outlets of Instagram and LinkedIn were also used to gather an image dataset of 50 images per athlete and general demographic background information. Environmental data was sourced from government or open source databases.

5.1 RQ1: CRF Bio-Variable Estimation

Per second power output values collected from the strain gauges were used as ground truth in our experiments for bio-variable estimation. We used a rolling average of maximum 4 minute power output per day over 42 days to generate the ground truth for our experiments. We trained two sets of linear regression models for each feature, a global model and a personal model. The global model was

trained using the data collected from a subset of subjects and tested on the remaining subjects. The personal model for each individual was trained on a 70% training subset for the subject and tested on the remaining 30% subset.

VAM models are trained to predict average power output (normalized by body weight) in 4 minute windows in an activity using the VAM in the time window, and the maximum estimated power output is then used to compute a continuous daily estimate for VO2 Max. We trained models with varying slope thresholds to identify the impact of slope on estimate accuracy. As the slope increases, the effect of other resistance factors (such as wind, rolling resistance) decreases and the model performs better (Table 1). We choose which model to use based on the maximum slope observed in the 4 minute windows, for example if in a ride the maximum slope observed in a 4 minute interval is 5.3%, we would choose the model trained on intervals where slope is greater than 5%. Since we are predicting body weight normalized power using VAM, none of the two metrics are greatly influenced by individual parameters. This is reflected in similar global and individual model performances for VAM (fig 8).

TRIMP captures the work done by an individual's heart in the last 42 days. We trained linear regression model to predict an individual's VO2 Max value based on their total TRIMP score in past 42 days. This metric proved to be more effective in a personal model than a global model as different individuals have different heart rate response to same exercise intensity (fig 8).

Active time is the actual amount of time the individual was actively putting in effort in past 42 days. We obtained this metric using cadence values collected at per second resolution. We trained linear regression model to predict an individual's VO2 Max value based on their total activity time in past 42 days. Similar to TRIMP, this metric performs better in a personalized model than in global model as different individuals would have a different response to the same exercise volume (fig. 8).

Combination models have outperformed their constituent models in all our experiments as shown by the error plots in fig. 8. The estimates from the previous models were combined using a weighted average, where weights for a model estimate are inverse of the model's training error. The error in estimates for these models are reported in fig. 8 and discussed in this section. We can see from the plot that the best model in terms of average error and variance in error utilizes all available data streams.

We also performed an experiment to find out the optimum time to be considered for aggregating the metrics while estimating CRF values. We plotted the test error for the global models utilizing one metric in fig. 7. We can see that while there is some variation in mean error, the 95% confidence intervals overlap for all time windows and we cannot find an optimum time window to use in our experiments based solely on the data. Therefore we have used the clinically recommended period of 42 days to aggregate the past exercise events.

5.2 RQ2: Cross-Modal Heart Health State

Even while referencing established bioscience research we still have no way to validate ground truth until decades into the future when people die. So comparisons are not the best way to experimentally validate this question. Performance comparisons for this type of

Table 1: Slope based optimization of VAM models

Slope threshold (%)	Test Set RMSE (Rel. Power)	Training Set RMSE (Rel. Power)	Training R Squared	Size of training set
0+	0.726	0.665	0.381	16810792
1+	0.620	0.537	0.527	10728610
2+	0.557	0.473	0.593	7952334
3+	0.488	0.424	0.655	6236507
4+	0.451	0.391	0.695	4884243
5+	0.420	0.363	0.732	3509724
6+	0.405	0.344	0.760	2294199
7+	0.395	0.328	0.781	1400488
8+	0.365	0.318	0.793	781841
9+	0.347	0.317	0.797	423315

Figure 7: Comparison of sensor prediction performance based on changing the memory in the model in determining health state. Based on a p-value of 0.05, there were no statistically significant differences in choosing the memory value. Thus, we chose the established standard of 42 days for our model [21].
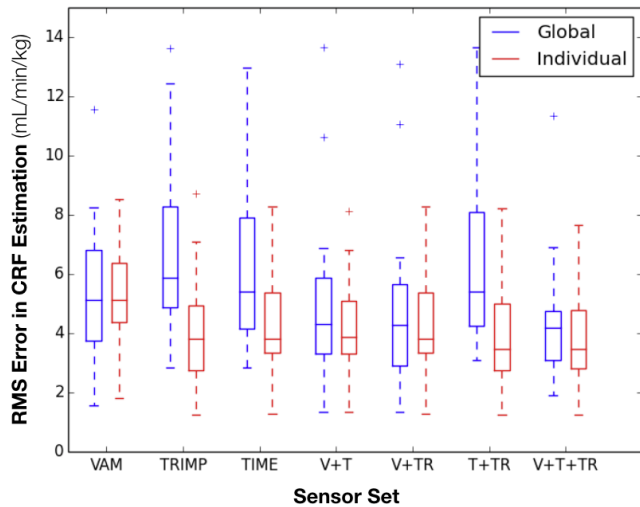

Figure 8: Comparison of sensor prediction performance. Global models are trained on all data collected from a subset of subjects. Individual models are trained on data collected from one subject. Global models can be used for estimation before the individual model for a person is calibrated.

experiment will need to be validated through large scale data collection and monitoring in prospective studies as mentioned in the related works. This research question largely poses a beginning for how multi-factorial health states can begin to surface for use.
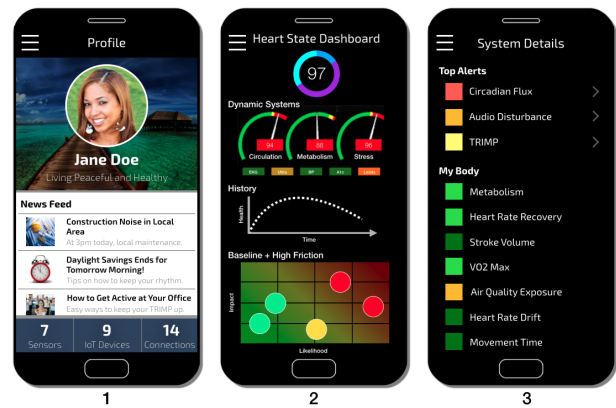

Figure 9: Screen 1 we show how health state can fuel recommendations. Inspired by jet engine dashboards, screen 2 can give a live snapshot of the health state. Screen 3 gives a comprehensive list of all bio-variables and health states being tracked, with a ranking system to provide relevant results at the top.

Using the approach described in section 4, we assimilate various biological parameters for each of the 24 subjects as shown in Figure 10. We find that even though most of these subjects are all cycling athletes, they have a wide range in both their bio-variables and environmental exposures. Current day primary care doctors would not be able to see this when a patient visits.

This data assimilation can be used inform personal health state as shown in Figure 9. The combination of sensors, IoT devices, and environmental data connections can provide a rich experience to interact with meaningful health insights. This can be for individual use, or for use in when a user visits a health care provider. One step further would be to link this with their electronic medical record system. Looking at this data panel across the subject pool, we discover some interesting trends in Figure 10. As expected, as the age increases the overall heart health state decreases, since age is a large factor of cardiac health (age range in panel is 18-57). As age increases, we also see a reduction in crime and noise pollution, suggesting that older individuals live in safer and quieter neighborhoods. We also see circadian rhythm disruption maximally in the middle ages (20-29), suggested a more erratic lifestyle for those in their twenties. Circulation and metabolism scores also trend (including VO2 Max / CRF) lower as age increases. Although this is a small sample size to make any strong conclusions, we can begin to see the power of using this cross data analysis for health.

6 CONCLUSIONS

In this paper we propose an approach to leverage cross-modal data to estimate a needed health variable or health state in the context of cybernetics. Specifically in the focus of cardiovascular health, we estimate CRF from various wearable devices. From our experimental results we can see that increasing the number of data streams provide increased performance characteristics in achieving this goal. Furthermore, we show that total health of an individual is much greater than any single biological variable, and that we need to integrate a diverse array of data types to more better understand
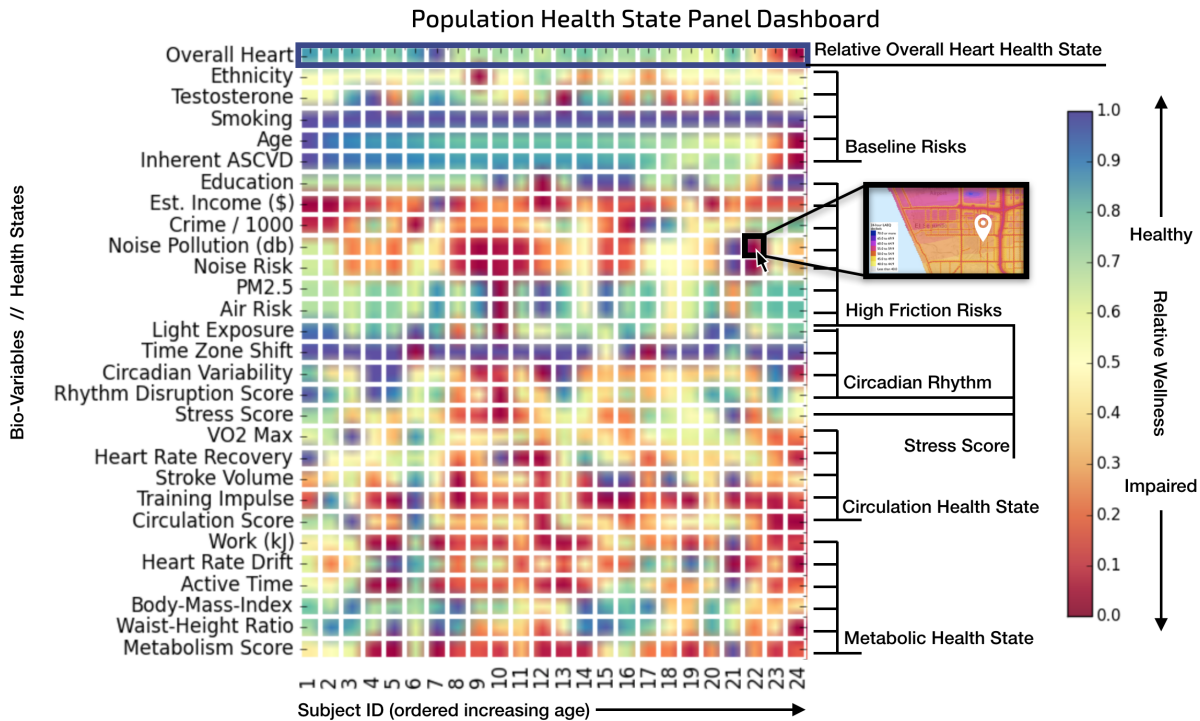
Figure 10: Heat map of bio-variables and summary scores that affect each individual subject. This type of visualization integrates cross-modal data in a manner that a clinician, hospital, public health agency, or any expert can use to monitor health of a patient panel. Clicking on a certain box would pull up further insights and details.

the total health state of a particular individual or organ system. Ideally we have some actionable or semantically meaningful dashboard as shown in Figures 9 and 10, which a user or a health expert can reference to get an "engine check" of the health state in real time.

Implications: The proposed utility of this work is to open the concept of using a diverse array of data streams to improve health state estimation. In the ideal case, this also lowers the cost for instantaneous health assessment, and provides increased value for individuals to purchase sensors like wearables and IoT devices. Individuals may be more motivated to track their health state, especially if it will be used in professional clinical decision making or influencing daily actions. This also provides the user with instant feedback with results from their lifestyle modification, medicine, environment and more. Perhaps this may be used as a tool to encourage healthy habits, or to avoid dangerous environments.

Limitations: Estimation in its initial iteration may not be accurate, but it is assumed to improve over time with refining of equations, algorithms, feature extraction methods, learning methods, as well as with improvements in hardware technology. We use linear methods in this paper as a starting point for obvious further advancement with advanced learning and predictive methods. Baseline comparisons are also difficult when studying individual subjects, and will require statistical methods for n of 1 studies [13, 17, 38, 39]. Wearable devices and similar low cost sensors are currently better used as a screening tool to identify if a user is at risk or their health state is changing, but clinical gold standard testing (which is more

expensive) might be used to confirm the true state of the user if the situation is critical. More work will need to be done to ensure the robustness of the estimations are good enough to use alone for important clinical decisions from more validated clinical and biomedical research. This work at the current stage cannot give a validated prediction window of when an adverse event (eg. myocardial infarction) may happen.

Future Directions: We hope to show how wearable devices, Iot, images, along with other data types can potentially be used in lieu of expensive sensors to estimate health status. For a single device, we show how multiple types of sensors improves prediction quality. Beyond a single device, we show that an assimilation of more diverse data with domain knowledge can further illustrate a wide view of health states. Additionally, privacy and security methods must evolve concurrently for such systems to function in the real world. Estimation is just one part of the cybernetic health paradigm. Ultimately, work must be done to ensure that reliable and useful systems are developed to guide people towards better health. Our dataset will be available in the public domain to encourage other researchers. We invite others to participate in building the foundational blocks of health state estimation and cybernetic health so that we can all enjoy a more informed and healthier life.

REFERENCES
[1] Fahd Albinali, Stephen Intille, William Haskell, and Mary Rosenberger. 2010. Using wearable activity type detection to improve physical activity energy expenditure estimation. *Proceedings of the 12th ACM international conference on Ubiquitous computing - Ubicomp '10* (2010), 311. https://doi.org/10.1145/1864349.1864396

[2] Marco Altini, Pierluigi Casale, Julien Penders, and Oliver Amft. 2015. Personalized cardiorespiratory fitness and energy expenditure estimation using hierarchical Bayesian models. *Journal of Biomedical Informatics* 56 (2015), 195–204. https://doi.org/10.1016/j.jbi.2015.06.008

[3] Marco Altini, Pierluigi Casale, Julien Penders, and Oliver Amft. 2016. Cardiorespiratory fitness estimation in free-living using wearable sensors. *Artificial Intelligence in Medicine* 68 (2016), 37–46. https://doi.org/10.1016/j.artmed.2016.02.002

[4] Margaret Ashwell and Sigrid Gibson. 2016. Waist-to-height ratio as an indicator of early health risk: Simpler and more predictive than using a matrix based on BMI and waist circumference. *BMJ Open* 6, 3 (2016). https://doi.org/10.1136/bmjopen-2015-010159

[5] Per-Olof Astrand, T Edward Cuddy, Bengt Saltin, and Jesper Stenberg. 1964. Cardiac Output During Submaximal and Maximal Work. *J Appl Physiol* 19 (1964), 268–274.

[6] August B. Hollingshead. 1975. Four Factor Index of Social Status. *YALE JOURNAL OF SOCIOLOGY* 8 (1975). https://s3.amazonaws.com/academia.edu.documents/30754699/yjs_fall_2011.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1526782500&Signature=zI74sfIS0kLlqQq4aYRtpnLp2NI%3D&response-content-disposition=inline%3B%20filename%3DAugust_B._Hollingshead_s_Four_Fa

[7] E W Banister and T W Calvert. 1980. Planning for future performance: implications for long term training. *Canadian journal of applied sport sciences. Journal canadien des sciences appliquees au sport* 5, 3 (9 1980), 170–6. http://www.ncbi.nlm.nih.gov/pubmed/6778623

[8] Jill Borresen and Michael Ian Lambert. 2009. The Quantification of Training Load, the Training Response and the Effect on Performance. *Sports Medicine* 39, 9 (9 2009), 779–795. https://doi.org/10.2165/11317780-000000000-00000

[9] Christopher R. Cole, Eugene H. Blackstone, Fredric J. Pashkow, Claire E. Snader, and Michael S. Lauer. 1999. Heart-Rate Recovery Immediately after Exercise as a Predictor of Mortality. *New England Journal of Medicine* 341, 18 (1999), 1351–1357. https://doi.org/10.1056/NEJM199910283411804

[10] E F Coyle and J. González-Alonso. 2001. Cardiovascular Drift during Prolonged Exercise: New Perspectives. *Exercise and Sport Sciences Reviews* 29, 2 (2001), 88–92. https://doi.org/10.1097/00003677-200104000-00009

[11] Irene R. Dégano, Jaume Marrugat, Maria Grau, Betlem Salvador-González, Rafel Ramos, Alberto Zamora, Ruth Martí, and Roberto Elosua. 2017. The association between education and cardiovascular disease incidence is mediated by hypertension, diabetes, and body mass index. *Scientific Reports* 7, 1 (2017), 1–8. https://doi.org/10.1038/s41598-017-10775-3

[12] E. Ray Dorsey and William J. Marks Jr. 2017. Verily and Its Approach to Digital Biomarkers. *Digital Biomarkers* 94080 (2017), 96–99. https://doi.org/10.1159/000476051

[13] Naihua Duan, Richard L. Kravitz, and Christopher H. Schmid. 2013. Single-patient (n-of-1) trials: a pragmatic clinical decision methodology for patient-centered comparative effectiveness research. *Journal of Clinical Epidemiology* 66, 8 (8 2013), S21–S28. https://doi.org/10.1016/J.JCLINEPI.2013.04.006

[14] Aleksandr Farseev and Tat-Seng Chua. 2017. Tweet Can Be Fit: Integrating Data from Wearable Sensors and Multiple Social Networks for Wellness Profile Learning. *ACM Transactions on Information Systems* 35, 4 (8 2017), 1–34. https://doi.org/10.1145/3086676

[15] Firstbeat. 2014. Automated Fitness Level (VO 2 max) Estimation with Heart Rate and Speed Data. *Firstbeat* (2014), 1–9.

[16] Collins Francis. 2015. A New Initiative on Precision Medicine. *The New England Journal of Medicine* 372, 9 (2015), 1–3. https://doi.org/10.1056/NEJMp1002530

[17] Nicole B. Gabler, Naihua Duan, Sunita Vohra, and Richard L. Kravitz. 2011. N-of-1 Trials in the Medical Literature: A Systematic Review. (2011), 761–768 pages. https://doi.org/10.2307/23053842

[18] David C. Goff, Donald M. Lloyd-Jones, Glen Bennett, Sean Coady, Ralph B. D'Agostino, Raymond Gibbons, Philip Greenland, Daniel T. Lackland, Daniel Levy, Christopher J. O'Donnell, Jennifer G. Robinson, J. Sanford Schwartz, Susan T. Shero, Sidney C. Smith, Neil J. Stone, and Peter W.F. Wilson. 2014. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American college of cardiology/American heart association task force on practice guidelines. *Circulation* 129, 25 SUPPL. 1 (2014). https://doi.org/10.1161/01.cir.0000437741.48606.98

[19] M T Hamilton, D G Hamilton, and T W Zderic. 2007. Role of low energy expenditure and sitting in obesity, metabolic syndrome, type 2 diabetes, and cardiovascular disease. *Diabetes* 56, November (2007), 2655–2667. https://doi.org/10.2337/db07-0882.CVD

[20] Ramesh Jain. 2018. A Navigational Approach to Health. *Arxiv* (5 2018). http://arxiv.org/abs/1805.05402

[21] Andrew M. Jones and Helen Carter. 2000. The Effect of Endurance Training on Parameters of Aerobic Fitness. *Sports Medicine* 29, 6 (2000), 373–386. https://doi.org/10.2165/00007256-200029060-00001

[22] Yasuhiko Kubota, Gerardo Heiss, Richard F. Maclehose, Nicholas S. Roetker, and Aaron R. Folsom. 2017. Association of educational attainment with lifetime risk of cardiovascular disease the atherosclerosis risk in communities study. *JAMA Internal Medicine* 177, 8 (2017), 1165–1172. https://doi.org/10.1001/jamainternmed.2017.1877

[23] Santosh Kumar, Gregory Abowd, William T Abraham, Mustafa Absi, Timothy Hnat, Syed Monowar Hossain, Zachary Ives, Jacqueline Kerr, Benjamin Marlin, Susan Murphy, James M Rehg, and Inbal Nahum-shani. 2017. Pervasive Health: Center of Excellence for Mobile Sensor Data-to-Knowledge (MD2K). *Pervasive Computing, IEEE* (2017), 18–22.

[24] Carmen E. Lefevre, Gary J. Lewis, David I. Perrett, and Lars Penke. 2013. Telling facial metrics: Facial width is associated with testosterone levels in men. *Evolution and Human Behavior* 34, 4 (2013), 273–279. https://doi.org/10.1016/j.evolhumbehav.2013.03.005

[25] Jonathan Lester, Carl Hartung, Laura Pina, Ryan Libby, Gaetano Borriello, and Glen Duncan. 2009. Validated caloric expenditure estimation using a single body-worn sensor. *Proceedings of the 11th international conference on Ubiquitous computing - Ubicomp '09* (2009), 225. https://doi.org/10.1145/1620545.1620579

[26] Alejandro Lucia, Alfredo Santalla, Margarita Pérez, Luis Miguel Chicharro, Alejandro Luc, Jes S Hoyos, Margarita Pé Rez, and José L Chicharro. 2002. Kinetics of VO2 in professional cyclists. *Med. Sci. Sports Exerc* 34, 2 (2002), 320–325. https://www.researchgate.net/publication/11532623

[27] Thomas Maier, Lucas Schmid, Beat Müller, Thomas Steiner, and Jon Wehrlin. 2017. Accuracy of Cycling Power Meters against a Mathematical Model of Treadmill Cycling. *International Journal of Sports Medicine* 38, 06 (6 2017), 456–461. https://doi.org/10.1055/s-0043-102945

[28] Thomas Münzel, Frank P. Schmidt, Sebastian Steven, Johannes Herzog, Andreas Daiber, and Mette Sørensen. 2018. Environmental Noise and the Cardiovascular System. *Journal of the American College of Cardiology* 71, 6 (2 2018), 688–697. https://doi.org/10.1016/j.jacc.2017.12.015

[29] Nitish Nag, Vaibhav Pandey, and Ramesh Jain. 2017. Health Multimedia: Lifestyle Recommendations Based on Diverse Observations. *Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval* (2017), 99–106. https://doi.org/10.1145/3078971.3080545

[30] Nitish Nag, Vaibhav Pandey, Hyungik Oh, and Ramesh Jain. 2017. Cybernetic Health. *arXiv* arXiv:1705, May (5 2017). http://arxiv.org/abs/1705.08514

[31] Norbert Wiener. 1948. *Cybernetics: Communication And Control In The Animal And The Machine.* https://archive.org/details/CyberneticsOrCommunicationAndControlInTheAnimalAndTheMachineNorbertWiener

[32] OpenCV. [n. d.]. Facial Landmark Detection. ([n. d.]). https://www.learnopencv.com/facial-landmark-detection/

[33] P. M. Oskui, W. J. French, M. J. Herring, G. S. Mayeda, S. Burstein, and R. A. Kloner. 2013. Testosterone and the Cardiovascular System: A Comprehensive Review of the Clinical Literature. *Journal of the American Heart Association* 2, 6 (2013), e000272–e000272. https://doi.org/10.1161/JAHA.113.000272

[34] J. Pärkkä, M. Ermes, K. Antila, M. Van Gils, A. Mänttäri, and H. Nieminen. 2007. Estimating intensity of physical activity: A comparison of wearable accelerometer and gyro sensors and 3 sensor locations. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings* (2007), 1511–1514. https://doi.org/10.1109/IEMBS.2007.4352588

[35] Robert Ross, Steven N. Blair, Ross Arena, Timothy S. Church, Jean Pierre Després, Barry A. Franklin, William L. Haskell, Leonard A. Kaminsky, Benjamin D. Levine, Carl J. Lavie, Jonathan Myers, Josef Niebauer, Robert Sallis, Susumu S. Sawada, Xuemei Sui, and Ulrik Wisløff. 2016. *Importance of Assessing Cardiorespiratory Fitness in Clinical Practice: A Case for Fitness as a Clinical Vital Sign: A Scientific Statement from the American Heart Association.* Vol. 134. e653–e699 pages. https://doi.org/10.1161/CIR.0000000000000461

[36] Sanjiv Sam Gambhir, T Jessie Ge, Ophir Vermesh, and Ryan Spitler. 2018. Toward achieving precision health. *Sci. Transl. Med* 10 (2018). http://stm.sciencemag.org/content/scitransmed/10/430/eaao3612.full.pdf

[37] F. A. J. L. Scheer, M. F. Hilton, C. S. Mantzoros, and S. A. Shea. 2009. Adverse metabolic and cardiovascular consequences of circadian misalignment. *Proceedings of the National Academy of Sciences* 106, 11 (2009), 4453–4458. https://doi.org/10.1073/pnas.0808180106

[38] P. Sedgwick. 2014. What is an "n-of-1" trial? *BMJ* 348, apr10 1 (4 2014), g2674–g2674. https://doi.org/10.1136/bmj.g2674

[39] Stephen Senn. 2017. Sample size considerations for *n* -of-1 trials. *Statistical Methods in Medical Research* (9 2017), 096228021772680. https://doi.org/10.1177/0962280217726801

[40] D L Simon, S Garg, G W Hunter, T H Guo, and K J Semega. 2004. Sensor Needs for Control and Health Management of Intelligence Aircraft Engines. August (2004), 1–15. http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA426767

[41] Juhani Smolander, Tanja Juuti, Marja Liisa Kinnunen, Kari Laine, Veikko Louhevaara, Kaisa Männikkö, and Heikki Rusko. 2008. A new heart rate variability-based method for the estimation of oxygen consumption without individual laboratory calibration: Application example on postal workers. *Applied Ergonomics* 39, 3 (2008), 325–331. https://doi.org/10.1016/j.apergo.2007.09.001

[42] Qinghua Sun, Xinru Hong, and Loren E. Wold. 2010. Cardiovascular effects of ambient particulate air pollution exposure. *Circulation* 121, 25 (2010), 2755–2765. https://doi.org/10.1161/CIRCULATIONAHA.109.893461

[43] Eric Topol. 2015. *The patient will see you now: the future of medicine is in your hands*. Basic Books.

[44] United States of America. [n. d.]. AirNow - Environmental Protection Agency. ([n. d.]). https://www.airnow.gov/

[45] United States of America. [n. d.]. Centers for Disease Control and Prevention. ([n. d.]). https://www.cdc.gov/

[46] United States of America - Depart of Transportation. [n. d.]. National Transportation Noise Map | Bureau of Transportation Statistics. ([n. d.]). https://www.bts.gov/newsroom/national-transportation-noise-map

[47] United States of America - NOAA. [n. d.]. Light Pollution Map. ([n. d.]). https://www.lightpollutionmap.info

[48] Christopher Paul Wild. 2005. Complementing the genome with an "exposome": The outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiology Biomarkers and Prevention* 14, 8 (2005), 1847–1850. https://doi.org/10.1158/1055-9965.EPI-05-0456

[49] Paul T Williams. 2001. Physical fitness and activity as separate heart disease risk factors : a Meta-Analysis. *Med Sci Sports Exerc.* 33, 5 (2001), 754–761.

[50] Peter W. F. Wilson, Ralph B. D'Agostino, Lisa Sullivan, Helen Parise, and William B. Kannel. 2002. Overweight and Obesity as Determinants of Cardiovascular Risk. *Archives of Internal Medicine* 162, 16 (2002), 1867. https://doi.org/10.1001/archinte.162.16.1867

[51] Jindan Zhu, Amit Pande, Prasant Mohapatra, and Jay J Han. 2015. Using Deep Learning for Energy Expenditure Estimation with Wearable Sensors. *17th International Conference on E-health Networking, Application & Services (HealthCom) - IEEE* (2015), 501–506. https://doi.org/10.1109/HealthCom.2015.7454554