Université de Montréal

Predicting Viral Respiratory Tract Infections Using Wearable Garment Biosensors

par

Oussama Jlassi

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Mémoire présenté en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en Informatique

14 Octobre 2023

© Oussama Jlassi, 2023

Université de Montréal

Faculté des arts et des sciences

Ce mémoire intitulé

Predicting Viral Respiratory Tract Infections Using Wearable Garment Biosensors

présenté par

Oussama Jlassi

a été évalué par un jury composé des personnes suivantes :

Miklos Csuros
(président-rapporteur)

Philippe Dixon
(directeur de recherche)

Jonathan Tremblay
(membre du jury)

Résumé

Les infections virales des voies respiratoires (IVVRs) causées par certains virus comme la grippe et le COVID-19 ont un impact significatif sur la santé publique et l'économie mondiale. Ces infections touchent un nombre important de personnes dans le monde et exercent une pression immense sur les systèmes de santé. Pour atténuer les effets néfastes des IVVRs, il est important de développer des techniques de détection précoce capables d'identifier les personnes infectées même si elles ne présentent aucun symptôme. Une telle détection permet un isolement et traitement rapide, ce qui réduit le risque de transmission et permet des interventions de santé publique ciblées pour limiter la propagation de l'infection.

Les méthodes de détection actuelles telles que la réaction en chaîne par polymérase (RCP) démontrent une sensibilité et une spécificité élevées, atteignant des taux de détection de 100% avec certaines méthodes de test disponibles dans le marché. De plus, les approches actuelles d'apprentissage automatique pour la détection des IVVRs, montrent des résultats prometteurs ; cependant, les méthodes actuelles reposent souvent sur l'apparition des symptômes, exigent un équipement coûteux et un personnel formé, et fournissent des résultats relativement retardés.
Notre projet vise à étudier la faisabilité de l'utilisation d'un algorithme d'apprentissage automatique entraîné sur des données physiologiques provenant de biocapteurs portables lors d'un protocole de test de marche sur escalier pour prédire le niveau d'inflammation associé aux IVVRs. De plus, l'étude vise à identifier les indicateurs les plus prédictifs des IVVRs.

Des participants en bonne santé ont été recrutés et inoculés avec un vaccin antigrippal vivant pour induire une réponse immunitaire. Au cours d'une série de tests d'escalier contrôlés cliniquement, des physiomarqueurs tels que la fréquence respiratoire et la fréquence cardiaque ont été meusurés à l'aide de biocapteurs portables. Les données collectées ont été utilisées pour développer un modèle de prédiction en ayant recours aux algorithmes d'apprentissage automatique, combinés avec un réglage d'hyperparamètres et en écartant un participant à la fois lors de l'entraînement du modèle.

L'étude a développé avec succès un modèle prédictif qui démontre des résultats prometteurs dans la prédiction du niveau d'inflammation lié au vaccin induit. Notamment, les caractéristiques de variabilité de la fréquence cardiaque (VFC) dérivées du biocapteur portable présentaient le potentiel le plus élevé pour détecter le niveau d'inflammation, atteignant une sensibilité de 70% et une spécificité de 77%.

Les implications du modèle de prédiction développé sont importantes pour les cliniciens et le grand public, notamment en termes d'autosurveillance et d'intervention précoce. Grâce aux algorithmes d'apprentissage automatique et des physiomarqueurs utilisés, en particulier les caractéristiques de VFC, cette approche a le potentiel de faciliter l'administration en temps opportun des traitements appropriés, atténuant ainsi l'impact des futures épidémies des IVVRs. L'intégration de biocapteurs portables et d'algorithmes d'apprentissage automatique fournit une stratégie innovante et efficace de détection précoce, permettant une intervention rapide et réduisant la charge sur les systèmes de santé.

Mots clés : Infections virales des voies respiratoires ; Détection précoce ; Biocapteurs portables ; Apprentissage automatique ; Physiomarqueurs ; Test d'escalier de 3 minutes ; Variabilité de la fréquence cardiaque

Abstract

Viral respiratory tract infections (VRTIs) caused by certain viruses like influenza and COVID-19, significantly impact public health and the global economy. These infections affect a large number of people worldwide and put immense pressure on healthcare systems. To mitigate the detrimental effects of VRTIs, it is crucial to urgently develop accurate early detection techniques that can identify infected individuals even if they do not exhibit any symptoms. Timely detection allows for prompt isolation and treatment, reducing the risk of transmission and enabling targeted public health interventions to limit the spread of the infection.

Current detection methods like polymerase chain reaction (PCR) demonstrate high sensitivity and specificity, reaching 100% detection rates with some commercially available testing methods. Additionally, current machine learning approaches for automatic detection show promising results; however, current methods often rely on symptom onset, demand expensive equipment and trained personnel, and provide delayed results.
This study aims to investigate the feasibility of utilizing a machine learning algorithm trained on physiological data from wearable biosensors during a stair stepping task protocol to predict the level of inflammation associated with VRTIs. Additionally, the study aims to identify the most predictive indicators of VRTIs.

Healthy participants were recruited and inoculated with a live influenza vaccine to induce an immune response. During a series of clinically controlled stair tests, physiomarkers such as breathing rate and heart rate were monitored using wearable biosensors. The collected data were employed to develop a prediction model through the utilization of gradient boosting machine learning algorithms, which were combined with hyperparameter tuning and a leave-one-subject-out approach for training.

The study successfully developed a predictive model that demonstrates promising results in predicting the level of inflammation related to the induced VRTI. Notably, heart rate variability (HRV) features derived from the wearable biosensor exhibited the highest potential

in detecting the level of inflammation, achieving a sensitivity of 70% and a specificity of 77%.

The implications of the developed prediction model are significant for clinicians and the general public, particularly in terms of self-monitoring and early intervention. By leveraging machine learning algorithms and physiomarkers, specifically HRV features, this approach holds the potential to facilitate the timely administration of appropriate treatments, thereby mitigating the impact of future VRTI outbreaks. The integration of wearable biosensors and machine learning algorithms provides an innovative and effective strategy for early detection, enabling prompt intervention and reducing the burden on healthcare systems.

Keywords: Viral respiratory tract infections; Early detection; Wearable biosensors; Machine learning; Physiomarkers; 3-minute stair test; Heart rate variability

Contents

List of tables

List of figures

List of acronyms and abbreviations

IVVRs	Infections virales des voies respiratoires
COVID-19	Coronavirus disease of 2019
RCP	Réaction en chaîne par polymérase
VFC	Variabilité de la fréquence cardiaque
VRTIs	Viral respiratory tract infections
PCR	Polymerase chain reaction
HRV	Heart rate variability
RSV	Respiratory syncytial virus
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid

ELISAs	Enzyme-linked immunosorbent assays
RPA	Recombinase polymerase amplification
LAMP	Loop-mediated isothermal amplification
HDA	Helicase-dependent amplification
RT-LAMP	Reverse transcription loop-mediated isothermal amplification
ml	Millilitre
CNN	Convolutional neural network
AUC	Area under the curve
ECG	Electrocardiogram
EDA	Electrodermal activity
SpO2	Oxygen saturation
IL	Interleukin
MCP	Monocyte chemotactic protein

IP-10	Interferon-gamma induced protein-10
IFNγ	Interferon
RRI	R-wave to R-wave interval
LF-HRV	Low-frequency HRV
CRP	C-reactive protein
HF-HRV	High-frequency HRV
ICC	Intraclass correlation coefficient
AI	Artificial intelligence
ML	Machine learning
RHR	Resting heart rate
BR	Breathing rate
HR	Heart rate
MV	Minute ventilation

LAIV	Live attenuated influenza vaccine
CoV	Coefficient of variance
cm	Centimetres
rpm	Respiration per minute
MSE	Mean squared error
3MSST	3-minute stair-stepping test
IRS	Immune response score
SVM	Support vector machine
RF	Random forest

Dedications

I am deeply grateful to dedicate this work to my dear parents. As many sentences and expressions, as eloquent as they are, they cannot express my gratitude and appreciation.
You knew how to instill in me a sense of responsibility, optimism and self-confidence in the face of the difficulties of life.
I owe you what I am today and what I will be tomorrow, and I will always do my best to remain your pride.
To my siblings, whom I love so much and to whom I wish success in life.

Oussama

Acknowledgements

This research project could not have been achieved without the guidance of my supervisor and collaborators.

First, I would like to express my immense gratitude to my supervisor Dr. Philippe Dixon, who spared no effort to provide me with more than enough guidance leading to the completion of this work, I deeply appreciate his extraordinary human qualities, his availability and readiness throughout the whole period of this project.

Additionally, I would like to thank Dr. Amir Hadid, my professional collaborator, for the help and feedback he provided me throughout my project as well as his availability to answer my questions and clarify some key concepts in this project.

I would also like to thank Dr. Dennis Jensen, my academic collaborator, who was always ready to provide valuable insights and comments throughout this project.

In addition, I'm thankful to the jury members for agreeing to evaluate my thesis and for their insightful comments.

Lastly, my family members, especially my sister, deserve special thanks for the immense support and encouragement throughout the project and the writing of this thesis.

Thank you to all those people who made my project a success.

Chapter 1

Introduction

1.1. Background and Rationale

VRTIs have negative impacts on healthcare systems and economies worldwide. The emergence of infectious diseases, such as influenza and the recent COVID-19 pandemic, has underscored the urgent need for early detection methods to mitigate the impact of these outbreaks. Current detection approaches encompass molecular methods like PCR and antigen tests, as well as machine learning algorithms combined with wearable biosensors; however, these methods have limitations.

Molecular methods provide precise detection of viral pathogens; however, their requirements for specialized equipment, highly trained personnel, and clinical settings pose challenges for scalability and accessibility. Moreover, the time required for obtaining test results may impede swift intervention and control measures. On the other hand, machine learning algorithms have demonstrated promising results in VRTI detection, but previous studies predominantly relied on patient-reported symptoms, rendering them unreliable for early detection. Furthermore, these studies have not adequately explored the prediction of the level of inflammation associated with VRTIs.

It is known that individuals with VRTIs may show decreased physical function, suggesting that a physical exercise protocol may be a powerful method to detect infection. To date, the use of a reproducible physical exercise in conjunction with quantitative biomarker measurements and machine learning predictive algorithms for VRTI detection has not been explored.

1.2. Aims of the Study

To address the previously mentioned limitations, we introduce the WE SENSE study which is a controlled, prospective longitudinal study that seeks to revolutionize the early detection of VRTIs by leveraging advanced wearable biosensors including smart shirts, wrist watches, and rings. Novel infection detection techniques, such as inflammatory biomarker mapping, PCR testing, and app-based symptom tracking, are employed. The study aims to identify subtle patterns indicative of infection within a vast dataset of wearable sensor data. Ultimately, the WE SENSE study aspires to provide a foundation for rapid intervention and outbreak control. As part of the overarching WE SENSE study, our sub-study focuses on a controlled, repeatable task, the 3-minute stair test. Our sub-study aims to determine whether objective data can replace or complement self-reported symptoms for VRTI detection by utilizing quantitative biomarker measurements and machine learning algorithms. We hypothesize that physiomarkers, particularly HRV, Breathing Rate (BR), and Minute Ventilation (MV) features derived from wearable devices will exhibit a strong correlation with the level of inflammation associated with VRTIs.

1.3. Thesis Outline

This thesis begins with a review of the literature (Chapter 2) in which the background and impact of VRTIs, their current detection methods, as well as related predictors, are introduced.

The third chapter, presented in the form of an article, is the core of the study in which, methodology (including the study design, data preprocessing and machine learning model development and testing), results, and discussion are presented. This chapter is a manuscript draft that will be submitted to the journal JMIR Public Health and Surveillance.

Additional insights about our approach, its implication on public health, and the future directions related to our project are presented in the fourth chapter.

Chapter 2

Literature review

2.1. Background and impact of viral respiratory tract infections

Infectious diseases that cause VRTIs, such as influenza, rhinoviruses, respiratory syncytial virus (RSV), and the recently identified severe acute respiratory syndrome coronavirus 2 (SARS CoV 2) which caused the COVID-19 pandemic have substantial implications for global health. The impact of these contagious illnesses spans across all age groups and necessitates considerable resources from healthcare systems and economies globally [**1**]. VRTIs often emerge with mild symptoms resembling the common cold, and they can have serious consequences on individuals' respiratory health, particularly among vulnerable groups such as children, older adults, and those with underlying medical conditions. As such, VRTIs significantly contribute to the yearly rise in sickness and mortality rates worldwide [**2, 3, 4**].

The ongoing COVID-19 pandemic has further emphasized the catastrophic ramifications of VRTIs, resulting in an unparalleled impact on public health, overwhelming healthcare systems, and inflicting significant economic burdens on nations [**5**]. As of June 2023, the total number of COVID positive cases in Canada reached 4686867 and the total number of deaths is 52804. The peak number of hospitalized patients reached nearly 11000 in 2022 [**6**]. Even after the COVID-19-related outbreak, vulnerable people might still be at high risk when catching common influenza.
One of the difficulties associated with VRTIs is their ability to manifest across a range of clinical expressions spanning from mild symptoms to severe respiratory distress. Moreover, individuals infected with VRTIs can carry the virus without experiencing symptoms (asymptomatic cases), unwittingly transmitting the infection to others. This plays a substantial role in the rapid transmission and potential for outbreaks of these infections.

2.2. Current methods for VRTI detection

Molecular-based and emerging machine-learning-based approaches to VRTI detection are now described.

2.2.1. Molecular-based methods

Molecular-based methods for VRTI detection, such as PCR, antigen, and isothermal nucleic acid amplification, refer to diagnostic techniques that rely on analyzing molecular components.

PCR tests are a type of test that uses a special technology to make more copies of a specific gene. This technology relies on the activity of an enzyme called deoxyribonucleic acid (DNA) polymerase. Detecting VRTI involves converting viral ribonucleic acid (RNA) into complementary DNA through reverse transcription. Afterwards, PCR amplification is carried out to quantitatively detect the fluorescent reaction, determining whether the patient carries the virus. [**7, 8**].

In addition to PCR methods, diagnostic techniques known as Antigen methods detect the presence of specific antigens, such as viral proteins or components. These methods typically utilize immunoassays like lateral flow or enzyme-linked immunosorbent assays (ELISAs) to identify and measure the target antigens accurately. To do this, a patient's nasal swab or blood sample is applied to a test strip or plate, interacting with labelled antibodies specific to the target antigen. The resulting reaction generates a visible signal, providing clear evidence of whether the target antigen is present or not. Antigen methods are often preferred due to their ability to deliver rapid results and frequent use in diagnosing viral infections, including respiratory viruses such as influenza or SARS-CoV-2 [**9**].

Apart from PCR and antigen tests, there are ongoing advancements in the development of techniques for isothermal nucleic acid amplification. These methods allow the specified reaction to occur at a constant temperature without requiring specialized equipment like thermal cyclers. Isothermal nucleic acid amplification techniques encompass various approaches such as loop-mediated isothermal amplification (LAMP), recombinase polymerase amplification (RPA) and helicase-dependent amplification (HDA). These techniques can be executed using portable devices like tablets, potentially mitigating the transmission of VRTIs within communities [**10**].
LAMP is a highly efficient nucleic acid amplification technique, allowing for rapid amplification of target DNA within just one hour. It is widely recognized as the preferred diagnostic test among isothermal amplification methods, with enhanced specificity due to the use

of four to six distinct primers binding to specific regions of the target DNA [**11**]. When dealing with RNA templates, the RT-LAMP reaction achieves amplification in a single step by incorporating a reverse transcription step. Traditional clinical diagnosis of VRTIs often requires specialized equipment and takes one or more days for completion, prompting the search for faster and more convenient techniques. To address this need, researchers have introduced RT-LAMP tests that are capable of identifying SARS-CoV-2 in as little as 30 minutes [**12, 13**]. While some early approaches exhibited lower sensitivity, later studies, such as the one by Zhu et al., implemented a one-step RT-LAMP method combined with a nanoparticle-based biosensor assay, achieving rapid and accurate detection of VRTIs in approximately 1 hour with 100% specificity and sensitivity [**14**]. Similarly, Yang's team presented a one-step RT-LAMP method capable of detecting three genes (ORF1ab, N, and E genes) to facilitate swift COVID-19 diagnosis, demonstrating a specificity of 99% and sensitivity comparable to RT-PCR [**15**].

Lu et al. [**16**] investigated many commercially available methods. The BINAXNOW INFLUENZA A&B and the FilmArray utilize different technologies to detect influenza viruses. The BINAXNOW test utilizes a chromatographic immunoassay. While the FilmArray utilizes RT-PCR technology. Both tests target Influenza A and B viruses and provide rapid results. The BINAXNOW test can deliver results within 15 minutes while the FilmArray takes 45 minutes. The sensitivity and specificity of these tests differ. When using rapid influenza diagnostic tests that rely on nasopharyngeal swabs on a population of 202 participants aged >= 16 years old, the BINAXNOW test has resulted in a sensitivity ranging from 70% to 89% for Flu A and 50% to 69% for Flu B whereas the FilmArray has a sensitivity of 97.1% and a specificity of 99.3% [**17**]. In contrast, the Xpert® Xpress SARS-CoV-2 test is designed specifically for detecting the SARS-CoV-2 virus. It utilizes real-time RT-PCR molecular technology and provides 100% sensitivity and specificity. Results can be obtained within 45 minutes using this test. What sets this test apart is its focus on COVID-19 diagnosis. While the BINAXNOW and FilmArray tests cover a broader range of VRTIs, including Influenza A and B, and RSV, the Xpert® Xpress SARS-CoV-2 test is highly specific to the SARS-CoV-2 virus. In summary, these diagnostic tests showcase diverse technologies, target pathogens, and performance characteristics. Molecular methods, especially PCR tests, have become the "gold standard for COVID-19 detection" [**18**] given their high sensitivity and specificity (100% with the Xpert® Xpress SARS CoV 2).
The downsides of the previously mentioned methods are numerous. While antigen swabs are easy self-test home kits, they are generally less sensitive than molecular tests like PCR. Their accuracy can be influenced by the timing of the test since they are typically more accurate when performed during the peak of viral shedding, which may be a few days after symptom onset. The disadvantages of PCR and LAMP methods are related to the

reliance on expensive laboratory equipment and highly trained clinicians. They also take a considerable amount of time to get the results ready. In addition, they require patients to be present at the clinic to perform the tests. Furthermore, PCR and LAMP methods are specific to strains that have the targeted sequences. Moreover, these methods do not provide an early detection given that in most cases, patients undergo these tests when in doubt of having a VRTI or after developing symptoms.

2.2.2. Machine learning methods

Machine learning methods have emerged as promising tools for VRTI detection. These approaches utilize algorithms that can learn patterns and relationships from large clinical and molecular datasets to classify whether an individual has a VRTI. Due to the recent proliferation of machine learning methods, notably deep learning methods, and the COVID-19 outbreak, most studies presented focused on the detection of COVID-19.

Several studies have proposed automatic detection methods for COVID-19 using different approaches and data sources, notably X-ray images. Narin et al. [**19**] utilized deep convolutional neural networks, achieving a sensitivity of 90.6% and specificity of 96.0% with the InceptionV3 model [**20**]. Minaee et al. developed Deep-COVID, employing deep transfer learning. Their ResNet50 model achieved a sensitivity of 98.0% and specificity of 89.6% for COVID-19 detection [**21**]. El Asnaoui et al. [**22**] focused on automated methods for pneumonia detection using deep learning and X-ray images. They reported a sensitivity of 94.61% and specificity of 98.02% with the Mobilenet V2 model. Ibrahim et al. [**23**] developed the Deep-Chest model for classification and diagnosis of COVID-19, pneumonia, and lung cancer. Using the deep learning VGG19 model combined with a convolutional neural network (CNN) architecture, they reported a sensitivity of 98.05% and a specificity of 99.5%.

In contrast to the image-based approaches, Quer et al. developed a machine learning model that incorporated wearable biosensors and patient-reported symptoms for COVID-19 detection. Their study reported an area under the curve (AUC) of 0.72 when considering metrics such as resting heart rate, sleep, and activity, as well as an AUC of 0.80 when relying solely on reported symptoms [**24**]. Mishra et al. developed a model using smartwatch data, reporting a 63% detection rate for symptomatic patients [**25**].

While these methods demonstrated promising results when relying on X-ray images and decent metrics using wearable biosensors, they are still limited in terms of reliance on special equipment to capture X-ray images, and they rely on patient-reported symptoms, which makes them unreliable for early detection, especially with asymptomatic patients.

2.3. Wearable biosensors for health monitoring

2.3.1. Overview of wearable biosensors

Wearable biosensors continuously monitor a person's health in real-time using a variety of technologies, and they provide digital measurements that are neutral and reliable [**26**]. These metrics, which provide a basic knowledge of a person's health and activity levels, include heart rate, breathing rate, minute ventilation, and acceleration offering a general understanding of an individual's health and activity levels [**27**]. Some examples of wearable devices for health monitoring are now described.

The Apple Watch (Apple Inc, USA) is a wearable with many sensors measuring various health and fitness parameters. The accelerometer detects movement and measures steps taken and calories burned. The gyroscope enhances the watches' ability to track movement and orientation accurately, providing valuable data for activities and workouts. Users can also take an electrocardiogram (ECG) using the electrical heart sensor, which gives insights into their heart rhythm and detects irregularities.

The Fitbit Sense (Google LLC, USA) is purposely engineered for comprehensive health monitoring. It incorporates an optical heart rate sensor to track resting heart rates. The wearable also has an accelerometer as well as a gyroscope to monitor physical activity levels. It also includes an electrodermal activity (EDA) sensor, which offers valuable insights into bodily responses towards emotional fluctuations or stressful situations. Also, it incorporates a skin temperature sensor to measure temperature fluctuations.

The Empatica wearable by Empatica Inc. in the USA is a device capable of monitoring and tracking various physiological signals. Using advanced sensors to measure parameters like EDA and skin temperature enables this device to detect potential health conditions or environmental influences. In addition to this functionality, it measures heart rate and heart rate variability providing information about cardiovascular health and stress levels. Furthermore, an accelerometer is included with this device to monitor activity levels.

The Astroskin/Hexoskin (Carré Technologies Inc, Canada) is a smart shirt that measures and tracks various physiological parameters. It utilizes integrated sensors to monitor vital signs such as heart rate, breathing rate, minute ventilation, and heart rate variability. The shirt also measures activity levels, providing data on steps taken, calories burned, and overall movement patterns. Additionally, it is a technology that provides accurate and reliable sensor readings during movement [**28**]. It also tracks sleep quality, providing

insights into sleep duration and stages.

The Oura Ring (developed by Oura Health in Finland) is a wearable biosensor capable of monitoring and analyzing health parameters. By assessing essential physiological parameters like heart rate, heart rate variability, and body temperature, it offers insightful data for users' well-being assessment. The wearable provides insights into cardiovascular health, stress levels, and sleep quality. It also tracks sleep patterns and stages, offering detailed information on sleep duration, efficiency, and disturbances. Additionally, the ring incorporates an accelerometer to monitor activity levels and movement throughout the day.

These biosensors have notable advantages such as the ability to continuously and automatically monitor and transmit the physiomarkers. Additionally, their small size and light weight make them portable allowing them to ease the burden of having fewer medical staff and use more hospital space for emergency or response care [**29**]; however, they also have some limitations such as the reliance on batteries which for some devices they tend to last only one day (Apple Watch). Users might also forget to remove and charge their devices while not in use, which can make them miss measuring valuable physiomarker data.

2.3.2. Applications of wearable biosensors for VRTI and inflammation detection

Ding et al. reviewed wearable technology comprehensively for monitoring COVID-19 patients [**30**]. The study examined various devices, including smartwatches, rings, wrist-worn bands, earbuds, and flexible skin-like e-tattoos, evaluated for their ability to measure oxygen saturation (SpO2) that is typically measured using a pulse oximeter (a small, device that can be clipped onto a person's finger, earlobe, or another area with good blood flow). Additionally, they evaluated the effectiveness of chest/abdominal straps, vests, facial masks (with humidity sensors), and flexible patches in measuring respiratory rate (using sensors such as accelerometers to detect chest or abdominal movements) [**30, 31**].
While the previously mentioned technologies provide reliable measures and convenient insights into vital signs like heart rate, respiratory rate, and SpO2, not all of them are suited for everyday use and it can be cumbersome and socially awkward to carry such devices during daily activities. Several other studies have focused on the detection of COVID-19 using wearable devices and patient-reported symptoms [**32**]. Tayal et al. [**33**] and Jiang et al. [**34**] developed innovative devices capable of measuring various physiological parameters. Tayal et al.'s device assessed SpO2, body temperature and respiration as well as pulse rates, while Jiang et al.'s integrated chest strap measured pulse pressure waves, blood pressure, ECG, HRV, lung volume and cough frequency.

In terms of patient-reported diagnosis, multiple studies employed popular wearable devices such as Fitbit, Apple Watch, Ava-bracelet, and Oura ring. Nestor et al. [**35**], Rich et al. [**36**], Natarajan et al. [**37**], Alavi et al. [**38**], Quer et al. [**24**], Mason et al. [**39**], and Conroy et al. [**40**] evaluated the accuracy of patient-reported symptoms in predicting COVID-19 cases. Nestor et al. and Rich et al. reported moderate accuracy, with Nestor et al. achieving an AUC of 60% and Rich et al. demonstrating a sensitivity of 68%. Natarajan et al. analyzed a large sample size (2745) and reported a reasonable AUC of 77%. Alavi et al. achieved good accuracy, with a sensitivity of 89% and specificity of 73%. Quer et al. reported a sensitivity of 72% and specificity of 73%, while Mason et al. achieved relatively good accuracy with a sensitivity of 82% and specificity of 63%. Conroy et al. demonstrated moderate accuracy, reporting a sensitivity of 62% and specificity of 88% using the Oura ring and Garmin.
These studies utilized various wearable devices and methodologies to assess VRTIs. While each study showed different levels of sensitivity and specificity, the results collectively highlight the potential of wearable technology in aiding the detection and monitoring of respiratory infections.

2.4. Biomarkers for inflammation detection

2.4.1. Introduction to biomarkers

In recent years, there have been significant advancements made in the use of molecular-based substances, such as cytokines, chemokines, and various immune cells status, for inflammation detection. These techniques involve analyzing specific inflammatory biological markers (biomarkers), which are substances in the body that are measurable in the biological media such as the human tissue, cells or fluids, and they can indicate the earliest subtle changes in the immune system objectively [**41**].
McClain et al. [**42**] conducted a study using a human viral challenge to examine the expression patterns of peripheral blood cytokines during an influenza A (H3N2) infection. Their findings highlight the distinct and meaningful cytokine expression patterns associated with symptomatic influenza in humans. The biomarkers and prediction rules included in their comparative study were chosen based on their wide-ranging clinical use or their potential to be applied in real-world clinical settings to differentiate between bacterial and viral respiratory infections and to identify cases of fever without a known cause. [**43, 44**].

Overall, biomarkers offer several advantages, including objective assessment, providing a reliable and valid measurement of a particular condition or disease. They are less

biased compared to questionnaires, allowing for a more accurate evaluation. Biomarkers also enable the study of disease mechanisms and provide a homogeneous understanding of risk or disease. However, there are certain disadvantages to consider. Timing is crucial when using biomarkers, as the sample collection timing can impact the results' accuracy. Biomarker analyses can be expensive, requiring financial resources. Storage of samples is another concern, as the longevity of samples needs to be ensured for future analyses. Laboratory errors can also affect the reliability of biomarker results. Finally, there is an ethical responsibility associated with the use of biomarkers, as their interpretation and application should be made carefully and ethically [**41**].

2.4.2. Commonly used biomarkers for inflammation detection

Several biomarkers have been studied for inflammation detection. McClain et al. conducted a compelling study, revealing that 53% of human subjects who were exposed to the influenza virus experienced symptoms [**42**]. However, their inflammatory biomarkers demonstrated increasing levels as early as 12 to 29 hours post-inoculation. The used biomarkers are monocyte chemotactic protein (MCP)-1, interleukin (IL)-6, IL-8, IL-15 and interferon-gamma induced protein-10 (IP-10). Meanwhile, Davey R et al. [**45**] found that the severity of influenza H1N1 and IL-6, IL-10, and Interferon (IFNγ) are connected. In another study, Lee N et al. [**46**] found that IL-6 demonstrated early increasing levels in the immune system between 5 to 12 hours after being exposed to the VRTI which is in this case the influenza virus [**47, 44**]. The increase of this biomarker preceded symptom onset by nearly 2 days. In their study, they also found other important biomarkers that showed increasing levels in the immune systems after nearly 29 to 36 hours post-inoculation. These biomarkers are MCP-1 and the chemoattractants IP-10. Importantly, at times that correspond to average symptom onset, but still well before symptom peak levels (45 to 60 hours post inoculation), other cytokines such as IFNγ, IL-8 and IL-15 begin to manifest in highly inflamed people [**46**].

2.5. Heart rate variability as a physiological marker for VRTI and inflammation during post-exercise conditions

2.5.1. Heart rate variability and inflammation

Heart rate variability quantifies the dissimilarity in temporal gaps separating each heartbeat. This parameter, influenced by both sympathetic and parasympathetic branches within the autonomic nervous system, unveils oscillations in successive heartbeats [**48**]. Quantifying the fluctuations in R-wave to R-wave intervals (RRI), also known as HRV, has

been considered useful for monitoring autonomic activity, particularly cardiac parasympathetic modulation [**48**]. By extracting HRV data, physicians, clinical researchers, and health experts establish vital benchmarks for monitoring an individual's physiological parameters regarding optimal health status across varied research activities. The data assists in identifying indicators that could significantly influence the parasympathetic and sympathetic nervous systems.

Numerous studies have extensively investigated the association between HRV and inflammation. They revealed that HRV and inflammation are inversely correlated [**49**]. Due to its established credibility as a measure of cardiac vagal regulation, HRV is expected to have an inverse connection with inflammatory markers. Numerous studies have provided evidence supporting this anticipated inverse association. For instance, in a study involving 1,601 healthy young individuals, Haarala et al. [**50**] observed that reduced low-frequency HRV (LF-HRV) was linked to elevated levels of C-reactive protein (CRP). Frederick et al. [**51**] found that substantial HRV decreases preceded CRP elevations with a 90.9% positive predictive value within 72 hours after intense activity. Meanwhile, Timothy M et al. [**52**] found significant inverse relationships between high-frequency HRV (HF-HRV) and the biomarker IL-6. Additionally, LF-HRV was significantly inversely correlated with IL-6, fibrinogen and CRP.

2.5.2. Heart rate variability during post-exercise

Assessing HRV after exercising provides an understanding of an individual's autonomic nervous system and overall cardiovascular health and gives a better idea about the balance between the sympathetic and parasympathetic branches of the autonomic nervous system [**53**]. Analyzing HRV after exercise allows us to evaluate how quickly the body returns to a state of balance after stress, which is a crucial indicator of cardiovascular resilience and recovery. During exercise, HRV is influenced by demands on the body making it less reliable for gauging autonomic function [**54**]. After engaging in activity the heart rate gradually goes back to its resting state because the parasympathetic nervous system gets reactivated and the sympathetic activity decreases [**55**].

Post-exercise HRV is commonly used in non-invasive assessment procedures for the determination of cardiovascular parasympathetic function. [**56, 57, 58, 59, 60**]. HRV has also been employed as a tool to investigate post-exercise autonomic and parasympathetic activity in studies conducted by Goldberger et al. [**61**], and Stanley et al. [**62**]. During recovery from exercise, HRV demonstrates a time-dependent recovery and eventually goes back to pre-exercise levels [**62**]. Recovery conditions such as posture have also been shown to affect

HRV recovery, with a more upright posture slowing recovery. Few studies investigated test-retest reliability of recovery during acute post-exercise, and they have reported moderate to good relative reliability, such as intraclass correlation coefficient (ICC) values between 0.58 and 0.91 during active recovery [**55**] and 0.69 to 0.92 during static recovery [**63**] for various HRV measures.

2.6. Research gap

Molecular detection techniques are widely considered to be the method, for detecting viruses due, to their exceptional sensitivity and specificity.[**64**]; however, these methods do have their drawbacks. They require samples of purity and specialized laboratory equipment. Additionally, trained professionals are needed to conduct the tests. It can also take several hours for results to be available [**65**]. One other major downside of such methods is the requirement of physical presence in clinics to conduct the necessary molecular screening tests. In previous research studies, there has been a heavy reliance on patient-reported symptoms; however, this approach poses limitations in effectively detecting asymptomatic or mildly symptomatic infections [**38, 66**]. The importance of overcoming this limitation stems from its direct implications for public health. Investigations into SARS-CoV-2 outbreaks have shed light on the substantial contribution made by asymptomatic individuals in driving infection rates—sometimes accounting for up to 69% [**67**]. It is, therefore, crucial to recognize and tackle these instances of unnoticed infection so as to effectively control the transmission of the virus [**68**]. The current machine learning-based methods either rely on patient-reported symptoms or have average results for early detection of patients who developed symptoms. Additionally, until this point, no prior research studies have attempted to detect the level of inflammation. They are only dedicated to detecting the presence of VRTI. Detecting the actual level of inflammation could provide flexibility advantages to making the necessary decisions towards mitigating the negative effects of these VRTIs. Our study aims to overcome these limitations by leveraging quantitative biomarker measurements (instead of relying on patient-reported symptoms) and combining wearable biosensors and machine learning algorithms during a series of easily reproducible physical activity settings for early detection of VRTIs.

Chapter 3

Predicting Viral Respiratory Tract Infections Using Wearable Garment Biosensors

Oussama Jlassi[1,2], Amir Hadid[3], Philippe C. Dixon[2,4], Dennis Jensen[3]

1. Department of computer science and operational research, Université de Montréal, Montreal, Canada

2. Research centre of the CHU Sainte-Justine, Montreal, Canada

3. Department of Kinesiology and Physical Education, McGill University, Montreal, Canada

4. School of Kinesiology and Physical Activity Sciences, Université de Montréal, Montreal, Canada

This manuscript is written by Oussama Jlassi and Philippe Dixon. It is intended to be submitted to the journal of JMIR Public Health and Surveillance after the main manuscript related to the study led by another team member gets published.

All the co-authors contributed to this study.

Oussama Jlassi was involved in all the steps of conducting the literature review, data analysis, the interpretation of the results, and the drafting of the manuscript.

Philippe Dixon assisted in developing the research questions, as well as code and manuscript reviews.

Amir Hadid established the study design and was involved during the data collection phase, he also provided feedback during the writing of the manuscript.

Dennis Jensen, the lead researcher, supervised the whole project and provided insights and comments during the writing of the manuscript.

3.1. Abstract

Background: Viral Respiratory Tract Infections (VRTIs) pose a major threat to public health. Early detection of these infections and implementation of suitable preventive measures are crucial to contain their transmission effectively. Machine learning approaches for automatic, early detection are promising; however, current approaches generally rely on the onset of symptoms, require expensive equipment, trained personnel, and, often, results are not quickly available. Thus, the primary aim of this study is to determine if a machine learning algorithm trained on physiological data from wearable biosensors during a stair stepping task protocol (3 minutes stair test, 2 minutes recovery) could be used to predict the level of inflammation. The secondary aim is to determine the most predictive indicators of the VRTI.

Methods: Healthy participants were recruited and inoculated with a live influenza vaccine to induce an immune response. Physiomarkers, including breathing rate and heart rate, during a series of clinically controlled stair tests, were monitored via a wearable biosensor. These data were collected to develop a prediction model using gradient boosting machine learning algorithms combined with hyperparameter tuning and a leave-one-subject-out method to train the models.

Results: The study successfully developed a predictive model that accurately predicts the level of inflammation in individuals. Notably, the features derived from heart rate variability (HRV) demonstrated the highest potential in detecting the level of inflammation with a sensitivity of 70% and a specificity of 77%. The findings indicate a correlation between the physiomarkers collected during the clinically controlled stair tests and the inflammatory response associated with VRTIs.

Impact: The developed prediction model and its association with the stair tests have significant implications for clinicians and the general public regarding self-monitoring and early intervention. By leveraging machine learning algorithms and utilizing physiomarkers, particularly HRV features, this approach has the potential to assist in administering appropriate treatments promptly, thus, mitigating the impact of future outbreaks.

Keywords: Viral respiratory tract infections; Early detection; Wearable biosensors; Machine learning; Physiomarkers; 3-minute stair test; Heart rate variability.

3.2. Introduction

VRTIs like COVID-19 and influenza have become major public health issues as they spread rapidly and can cause severe symptoms or even death. It is vital to establish a system for early detection of VRTIs, allowing for treatment of these infections in their initial stages and effectively preventing widespread transmission. The integration of advanced technologies like artificial intelligence (AI) and machine learning (ML) can play a significant role in developing and implementing early detection protocols.

Current detection methods involve molecular testing such as antigen methods, isothermal nucleic acid amplification, and polymerase chain reaction (PCR) which became the gold standard for COVID-19 detection [**64**]. For example, Lu et al. [**16**] investigated many commercially available methods, including FilmArray (BioFire Diagnostics Inc. Salt Lake City, USA), a testing system that detects VRTI, such as Influenza A and B, and respiratory syncytial virus (RSV), using rapid testing PCR technology. This test provides results within 45 minutes and reports a sensitivity of 97.1% and a specificity of 99.3%.

There have also been studies that leveraged ML algorithms to detect VRTIs. For example, Narin et al. [**19**] proposed an automatic detection method for Covid-19 using X-ray images and deep convolutional neural networks. Their study utilized the InceptionV3 model [**20**], achieving a sensitivity of 90.6% and specificity of 96.0%. In another study, Quer et al. [**24**], developed ML models to detect COVID-19. Although no sensitivity and specificity metrics were presented, the study reported an area under the curve (AUC) of 0.72 when relying on resting heart rate (RHR), sleep, and activity metrics. They also used a model that only relies on patient-reported symptoms, resulting in an AUC of 0.80. This latter, more predictive model, may be of limited practical use for early detection of VRTIs as it relies on symptoms. Moreover, it cannot be used to detect infection in asymptomatic individuals.

Biological markers (biomarkers) can be used to objectively measure the immune response in cases of inflammation. They are also effective in detecting inflammation early on as they show increased levels in the host immune system several hours before symptoms begin to manifest [**41**]. Several research studies have identified monocyte chemotactic protein 1 (MCP-1), certain interleukins (IL 6, IL 8, IL 10, IL 15), interferon gamma induced protein 10 (IP 10), interferon gamma (IFNγ) and tumour necrosis factor alpha (TNFα) as being strongly associated with inflammation detection [**42, 45, 46, 47, 69, 70**].

Additionally, research has found that there is an inverse correlation between inflammation and HRV [**49, 50, 51, 52**]. This means that as inflammation increases, HRV decreases.

Other studies report that HRV may act as a physiomarker of cardiac wellness during post-exercise periods [**55, 59, 62**], suggesting that exercise protocols could be a key component of early VRTI detection.

Overall, previous studies mostly relied on the onset of symptoms, required special equipment, highly trained clinicians, and with some methods, results are not quickly available. Additionally, to date, no prior research has attempted to detect the level of inflammation but rather focused only on detecting the presence/absence of VRTIs. Thus, we have developed the Wearable Sensors for Early Detection and Tracking of Viral Respiratory Tract Infection (WE SENSE) protocol [**68**]. The overall objective of WE SENSE is to leverage wearable biosensors combined with ML algorithms and quantitative objective immune response measurements to allow for early detection of VRTIs. In the present study, we explore the use of a controlled exercise protocol within this paradigm. Specifically, the aims of this study are to (1) assess the potential of a 3-minute stair-stepping task to predict the level of inflammation using ML and wearable biosensors and (2) identify the most critical wearable sensor-derived metric indicators (features) of such infections.

3.3. Methods

3.3.1. Participants

We recruited healthy adults (aged 18 to 59 years). Exclusion criteria were pre-existing respiratory tract infections or other underlying health conditions. Additionally, we only recruited participants who did not intend to receive a COVID-19 or seasonal influenza vaccine during the study period, and participants had to be clear of any other inflammation within 7 days prior to the start of the study. We conducted the data collection from December 2021 to February 2022. Participants were recruited through word-of-mouth and posted advertisements in Montreal, Canada. Further details about the inclusion, exclusion, and withdrawal criteria are provided in the study protocol [**68**].

3.3.2. Study design and data collection

The study period lasted 12 days and was divided into two distinct phases: the pre-inoculation period (considered as day -7 to day 0) and the post-inoculation period (day 0 to day 4). On day -7, participants continuously wore a Astroskin (10 participants) / Hexoskin (45 participants) garment (Carré Technologies, Inc. Montreal, Canada) to monitor their physiomarkers (Figure 3.1 A). Blood samples were also collected (Figure 3.1 B), and a 3-minute stair-stepping test (3MSST) followed by 2 minutes of recovery was conducted (Figure

3.1 C). The Astroskin/Hexoskin garment was used to collect the following physiological markers relevant to respiratory health:

- **Heart Rate (HR):** The number of heartbeats per minute (bpm)
- **Breathing Rate (BR):** The number of respirations per minute (rpm)
- **Minute Ventilation (MV):** The total volume of air breathed in a minute, derived from the breathing rate and the tidal volume, which is the volume of air inspired in the last inspiration (ml/minute)
- **RR Interval (RRI):** The time duration between two consecutive heart beats on an electrocardiogram (ECG) (seconds)

The data format of the measured physiomarkers is a 1-D array counting 300 rows (1 data point every second) per physiomarker per visit. Each data point is an integer number for HR, BR, and MV, and a floating point value for RR. Although the Astroskin shirt also measures blood oxygen concentration, we limited analyses to signals that were common to both garments (the exact same raw data is provided by both biosensors). The accuracy and reliability of the smart shirts have been previously established [**71, 72, 28**]. On day 0, blood samples, and a stair test were performed to provide additional baseline data and assess participants' physiological responses during exertion. Participants were then inoculated with a live attenuated influenza vaccine (LAIV): FluMist® [**73**]. The post-inoculation period (5 days) involved daily blood sample collection to monitor the participants' response to the inoculated influenza vaccine. Participants performed the stair test 3 additional times during the post-inoculation period (days 1, 2 and 3).

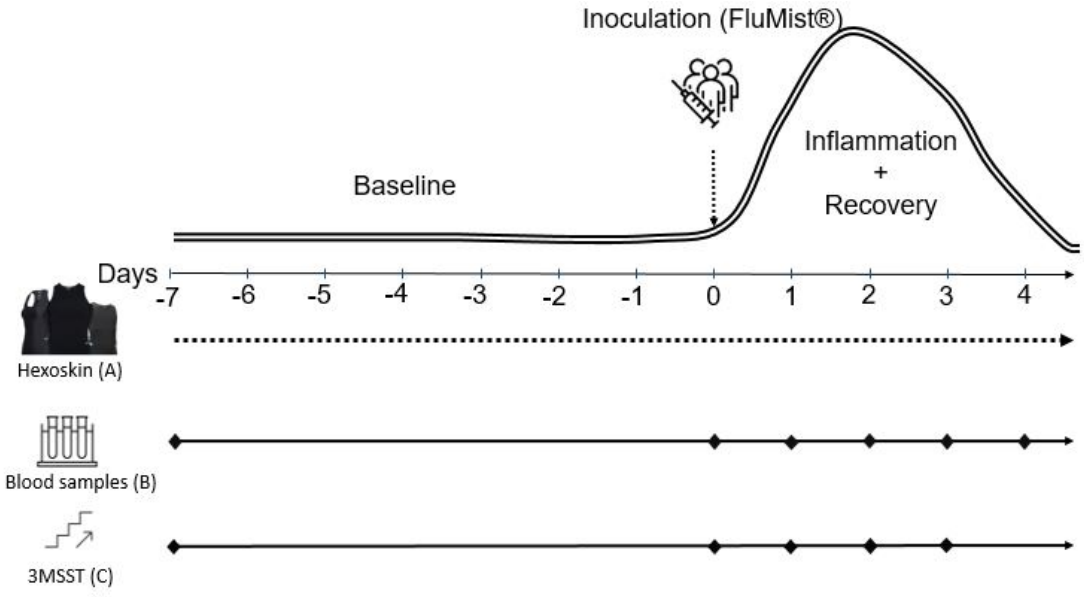

Figure 3.1. Study design

The study incorporates continuous physiomarker monitoring (A). The inoculation takes place

on day 0, with nasal swabs and blood samples collected on days -7 and 0, as well as during the post-inoculation period (B). The 3-minute stair test (C) was conducted on days -7, 0, 1, 2 and 3. Idealized inflammation response curve shown.

3.3.3. Calculating the immune response score

Blood samples facilitated the measurement of biomarkers and immune-related factors, aiding in identifying baseline levels and subsequent changes induced by the influenza vaccine. The inflammation level was obtained by calculating the immune response score (IRS) which is the objective of our prediction model. This score is composed of the mean value of 8 measured biomarkers (IFNγ, IL-6, IL-8, IL-10, IL-15, IP-10, MCP-1, and TNFα) during each visit (equation (3.3.1)). A detailed definition of each biomarker is presented in Table 3.1. These biomarkers were chosen due to their ability to detect inflammation in the host immune system [**42, 45, 46, 47, 69, 70**]. These biomarkers were measured using the ELISA method which involves the use of specific antibodies that bind to the target biomarker. Additionally, the Cytokine Arrays method was used which is a microarray technology that can simultaneously measure the levels of multiple cytokines in a single sample [**74**].

$$\text{IRS} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3.3.1}$$

Where:

$$n = \text{Number of measured biomarkers (8)}$$
$$x_i = \text{Individual biomarker measurement at index } i$$

We then used 4 IRS scores obtained from the baseline period (2 scores per visit), to calculate the IRS' coefficient of variance (CoV) for each participant (equation (3.3.2)), allowing us to assess the individual variation of the IRS values.

$$CoV = \left(\frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 + \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}{\frac{1}{n}\sum_{i=1}^{n} X_i + \frac{1}{n}\sum_{i=1}^{n} Y_i} \right) \tag{3.3.2}$$

Where:

$$X_i = \text{IRS values at the first baseline visit (i = 1, 2)}$$
$$Y_i = \text{IRS values at the second baseline visit (i = 1, 2)}$$
$$\bar{X} = \text{Mean of the IRS values at the first visit}$$
$$\bar{Y} = \text{Mean of the IRS values at the second visit}$$
$$n = \text{Number of IRS values (pairs of values)}$$

We used these CoV values to apply a 0-CoV normalization (equation (3.3.3)) to the original IRS:

$$\text{Normalized IRS} = \text{IRS} - \text{average CoV} - 1 \tag{3.3.3}$$

By normalizing the IRS values, the individual baseline variations among all participants are eliminated, thus, providing generalizable scores of inflammation levels.

Table 3.1. Definitions of the used biomarkers

Biomarker	Definition
IFNγ	Interferon-gamma (IFNγ) is a cytokine that plays a crucial role in the immune response to viral infections and various immune system functions.
IL-6	Interleukin-6 (IL-6) is a proinflammatory cytokine involved in the regulation of immune responses and acute-phase reactions.
IL-8	Interleukin-8 (IL-8) is a chemokine that attracts and activates neutrophils, playing a key role in the inflammatory response.
IL-10	Interleukin-10 (IL-10) is an anti-inflammatory cytokine that helps regulate and limit immune responses.
IL-15	Interleukin-15 (IL-15) is involved in immune system functions, including the development and activation of T and natural killer cells.
IP-10	Interferon-gamma-induced protein 10 (IP-10) is a chemokine that recruits immune cells to sites of infection and inflammation.
MCP-1	Monocyte chemoattractant protein-1 (MCP-1) is a chemokine that attracts monocytes to sites of injury and infection.
TNFα	Tumor necrosis factor-alpha (TNFα) is a proinflammatory cytokine that plays a key role in the body's response to infection and inflammation.

3.3.4. Physiological parameter monitoring (stair test)

The stair test consists of stepping for 3 minutes (with a consistent pace of 30 steps/minute) up and down a step having a height of 20 cm [**68**]. We chose this controlled clinical test as it's been previously used to uncover information about a participant's physiological response to exercise-induced stress [**75, 76, 77, 78**]. Following the test, participants recovered in a seated position for 2 minutes. Physiomarkers were monitored during the entire session.

3.3.5. Data processing and feature extraction

First, the 5-minute period of interest (stair test and recovery) was segmented based on annotated timestamps, resulting in 238 visits having 4 signal data files per visit and 300 data points (integer numbers) per signal (Figure 3.3 A). We then plotted all the sensor data for each visit to assess the signal quality and identify missing or noisy segments through a visual inspection (Figure 3.3 B). Examplars of noisy and typical heart rate signals are shown in Figure 3.2. During this step, data were removed due to noisy signals from BR (24

visits), missing or noisy HR signal (49 visits), noisy MV signals (3 visits), and missing or noisy RRI signal (68 visits). Next, we performed data cleaning. This involved identifying inconsistencies and errors in the dataset that were flagged by Hexoskin's systems and the data cleaning protocols (detailed in Table 3.2 [**79, 80, 81**] and Figure 3.3 C). The data points having such inconsistencies and errors are discarded from the record of the concerned visit. Finally, we performed artifact removal for the RRI data following 3 rules (Acar, Karlsson and Malik rules) [**81**] (Figure 3.3 D).

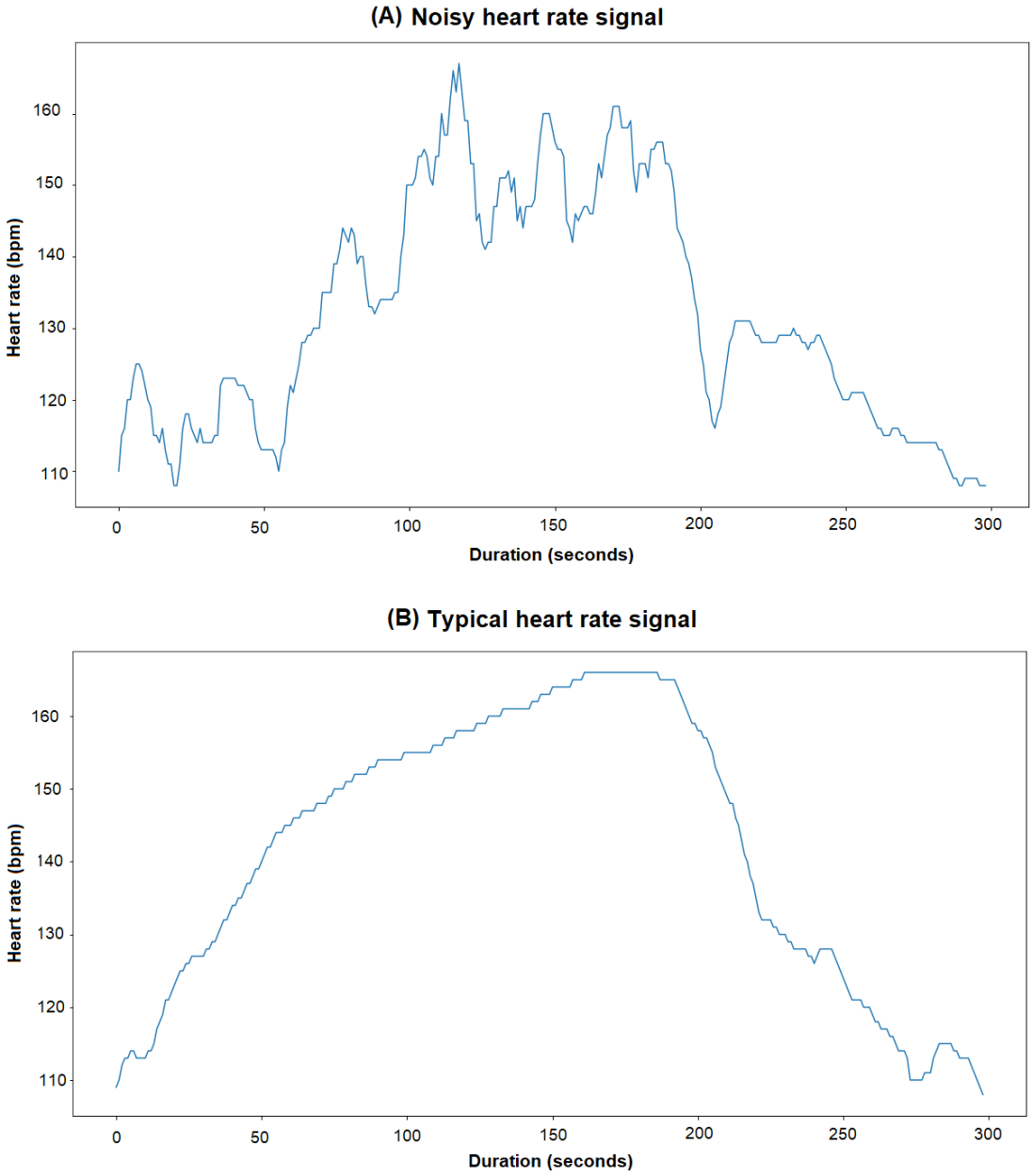

Figure 3.2. Visual inspection of the data.

Table 3.2. Data cleaning rules

Signal	Cleaning rules [**79, 80, 81**]
BR	• Keep data points where the BR quality flag is 0 or 2 • Discard data points where BR is outside the range [4, 60] respirations per minute (rpm)
MV	• Keep data points where MV is under 160000 ml/minute
RRI	• Discard data points where RRI quality is 128 • Discard data points where RRI is outside the range [0.25, 2] seconds
Abbreviations: RR (breathing rate), MV (minute ventilation), RRI (R wave to R wave interval). Data quality flags 0, 2 and 128 correspond to "good quality respiration", "abdominal channel disconnected" (we can use the thoracic channel alone), and "unreliable RRI signal", respectively [**82**]	

We extracted the HRV features during the first minute of recovery using the Flirt Python package [**81**]. Flirt offers tools and algorithms specifically designed for extracting HRV-related features from the RRI signal, including statistical (e.g. min, max, mean), time-domain (e.g. skewness, kurtosis, number of peaks), and frequency-domain features (e.g. low frequency to high frequency ratio) (Figure 3.3 E). Furthermore, we calculated additional features to further enhance the analysis (Figure 3.3 F). These features specifically targeted the BR, HR as well as the MV signals, and included the increase or decrease from the baseline during specific time intervals of the stair test, such as minutes 0 to 3 and 3 to 4 (equation (3.3.4)). An AUC feature was also calculated (using the Trapz function provided in the Numpy Python package [**83**]) for the 3 minutes of activity and the subsequent minute of recovery. A full list of features is provided in Table 3.3.

$$\mathrm{f} = \sum_{i=1}^{5} x_i - \sum_{j=1}^{5} (x-1)_j \tag{3.3.4}$$

Where:

$$\mathrm{f} = \text{The calculated feature}$$

$$x_i = \text{Last 5 data points in minute } x$$

$$(x-1)_j = \text{First 5 data points in minute } x-1$$

Finally, we concatenated the cleaned data, extracted and calculated features, and consolidated all relevant information for a particular visit into a cohesive record. Afterwards, we merged all visits together resulting in a final dataset having 170 rows (visits) and 40 columns (39 features and the corresponding IRS score for each visit) (Figure 3.3 G).

Table 3.3. Full feature list

Feature name	Definition	Source
signal_inc_0_1	Difference of the signal from minute 0 to minute 1	Equation (3.3.4)
signal_inc_1_2	Difference of the signal from minute 1 to minute 2	Equation (3.3.4)
signal_inc_2_3	Difference of the signal from minute 2 to minute 3	Equation (3.3.4)
signal_dec_3_4	Difference of the signal from minute 3 to minute 4	Equation (3.3.4)
signal_auc_3	Area under the curve of the signal during the 3 minutes of activity	Numpy
signal_auc_3_4	Area under the curve of the signal during the first minute of recovery	Numpy
rmssd	Magnitude of the differences between consecutive heartbeats	Flirt
lf_hf_ratio	Low frequency to high-frequency ratio	Flirt
std_hr	Standard deviation of the heart rate	Flirt
peaks	Number of RRI peaks	Flirt
n_above_mean	Number of RRI peaks above the mean	Flirt
skewness	Skewness (assymetry) of the data distribution	Flirt
range_nni	*max(NN_intervals) - min(NN_intervals)	Flirt
kurtosis	Kurtosis of the data distribution	Flirt
sdnn	Standard Deviation of the NNI	Flirt
cvnni	Coefficient of Variation of the NNI	Flirt
mean_hrv	Heart rate variability mean value	Flirt
min_hrv	Heart rate variability min value	Flirt
max_hrv	Heart rate variability max value	Flirt
vlf_hrv	Very Low-Frequency power	Flirt
lf_hrv	Low-Frequency power	Flirt
hf_hrv	High-Frequency power	Flirt
lineintegral	Are under the curve of the heart rate variability	Flirt
iqr	Dispersion of the heart rate variability measures within the middle half of the data	Flirt
entropy	entropy of the heart rate variability	Flirt
median_nni	*Median of NN intervals	Flirt
mean_nni	*Mean of NN intervals	Flirt
signal can be heart rate, breathing rate, minute ventilation Abbreviations: RRI (R-wave to R-wave Interval) * NNI (Normal to Normal Interval) * RMSSD (Root Mean Square of Successive Differences)		

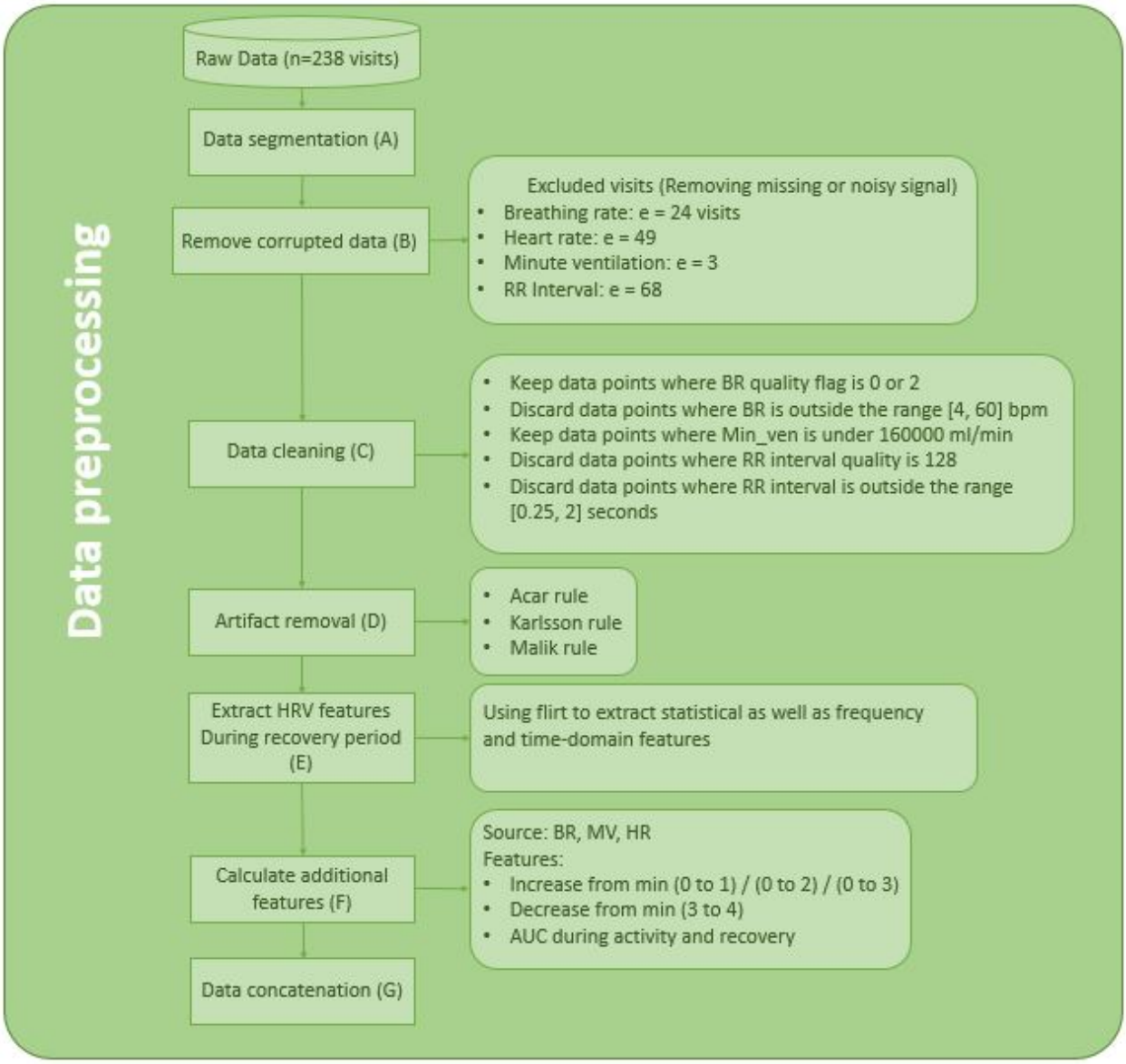

Figure 3.3. Diagrammatic representation of the data preprocessing approach

Raw data are segmented (A), inspected (B), cleaned (C), and artifacts are removed (D). Afterwards, features are calculated and concatenated (E, F, G).

3.3.6. Machine learning model development

Selecting an appropriate model for our task is a critical step. Our objective was to predict the inflammation response score using a combination of physiological features obtained from wearable garment biosensors, making it a regression task. Several models can be applied for such tasks, each with distinct characteristics and trade-offs. Linear Regression, a classic method that models relationships using a linear equation, offers simplicity and interpretability; however, it cannot model complex non-linear relationships. Random forests,

an ensemble method, combines decision trees by using random feature selection and data subsets, offering high predictive accuracy and resistance to overfitting. XGBoost, another ensemble technique, builds decision trees sequentially, optimizing for errors made by previous trees, this model provides exceptional predictive performance and fine-tuned control over regularization. Finally, neural networks, a deep learning approach, use interconnected layers of artificial neurons to capture complex, non-linear relationships and are best suited for unstructured data like images and text; however, they demand large amounts of data and can be computationally intensive. The choice of the model in the current work depended on factors like the dataset size (which in our case was small), prediction task, problem complexity, and interpretability. While a Gradient Boosting algorithm (XGBoost [**84**]) was selected due to its ability to model complex relationships, handle missing data, and its robust performance [**85**], we also used a random forest model for comparison given the limited number of studies to directly compare our results with. We now describe the steps taken in our machine learning model development pipeline in order to predict the inflammation response score. First, data were split into train and test sets, taking into account the participants' balanced number of visits pre and post-inflammation. Five randomly selected participants were kept in the test set, and the remaining participants (n=50) were included in the train set (Figure 3.6 A). Next, we applied filter methods to the train set to reduce features and improve model performance (Figure 3.6 B). We removed duplicate features, features with a variance lower than 0.01 across the dataset (poor predictive potential since these features are nearly invariant across inflammation states), and highly correlated features (correlation threshold, $r > 0.8$). Finally, we employed hyperparameter tuning to optimize our machine learning model's performance. Specifically, we used GridSearchCV from the Scikit-learn library to systematically search through a predefined hyperparameter grid that included "max_depth", "learning_rate", "n_estimators", and "alpha". GridSearchCV performed multi-dimensional tuning, exploring various combinations of these hyperparameters, rather than tuning one hyperparameter at a time. GridSearchCV essentially performed an exhaustive search, trying every possible combination of hyperparameters defined in our grid to identify the combination that resulted in the least mean squared error (MSE) of the predicted IRS. The grid search approach was implemented with a 3-fold cross-validation to reduce the risk of overfitting (Figure 3.6 C). A 3-fold cross-validation when tuning the hyperparameters means that the train data is split into 3 approximately equal parts. The grid search algorithm iteratively uses each fold as a testing set while the remaining two folds are used as the training set. This process is repeated three times, with each fold serving as the testing set exactly once. After each iteration, the model's performance is evaluated based on the MSE values obtained from the prediction on the test set. Table 3.4 presents the hyperparameters, the ranges, the final tuned value and the step size during the tuning process.

Table 3.4. Tuned hyperparameters

Hyperparameter	Range	Meaning	Tuned value	Step size
Alpha	0.5 to 0.6	Controls the regularization term in the model	0.5	0.1
learning_rate	0.01 to 1	Determines the step size at each boosting iteration, influencing the impact of each tree on the final prediction	0.1	5 specific values (0.01, 0.1, 0.5, 0.8, 1)
max_depth	50 to 150	Limits the depth of each decision tree in the ensemble, preventing overcomplexity and reducing overfitting	50	10
n_estimators	50 to 100	Specifies the number of boosting rounds, representing the number of decision trees in the ensemble	50	10

3.3.7. Machine learning model training and evaluation

We trained and tested models using a leave-one-subject-out strategy (Figure 3.4). This approach involved training on the combined train data (n=50 participants for training which corresponds to 153 visits) and the test participants' data for 5 iterations (n=5 participants for testing, 16 visits), excluding one participant during each iteration. Specifically, we use the 50 (train) + 4 (test) participants data to train the model and we test its performance on the left out participant from the test set and we repeat that with each test participant. To assess model performance, mean and standard deviation of performance metrics (across the 5 test participants) are reported.

We developed 9 models to assess the predictive potential of different feature sources. For the BR, HR, MV signals, we developed 2 models based on features extracted from the 3 minutes of activity and the first minute of recovery, respectively. The model based on HRV features was only generated for the recovery period since HRV analysis is mainly relevant during rest/recovery [**55, 59, 62, 61**]. Additionally, we trained a model based on the features derived from the combined signals and a random forest model using HRV features during recovery for comparison.

Figure 3.4. Leave-one-subject-out training strategy

We evaluated the regression model's performance using the MSE values of the predicted IRS (Figure 3.6 E), then, we derived a binary outcome from the predicted values by setting a threshold n, classifying participants as no inflammation (IRS <= n) or high inflammation (IRS > n) (Figure 3.5). A threshold of n=0 is justified in the scatter plot of the predicted vs true IRS values in Figure 3.10. This binary outcome enabled the evaluation of model performance via classification-related metrics (Figure 3.6 G), including accuracy, sensitivity, specificity, precision, and AUC, which is the primary metric we used to evaluate the classification performance.
The effect of varying thresholds in the range [-0.3, 0.3] is explored.

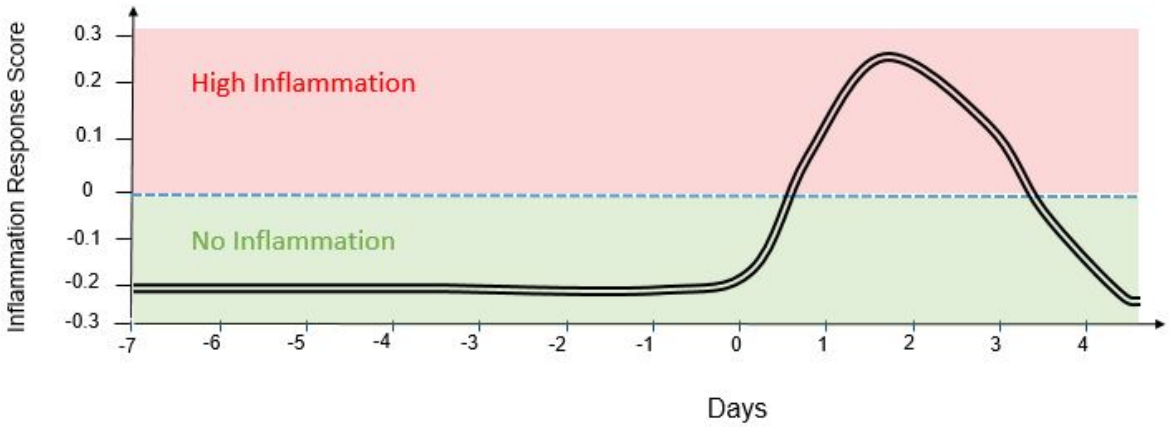

Figure 3.5. Representation of the binary conversion process.

An IRS score above the threshold 0 indicates inflammation, while a score equal to or below 0 indicates a healthy participant. Idealized inflammation response curve shown.

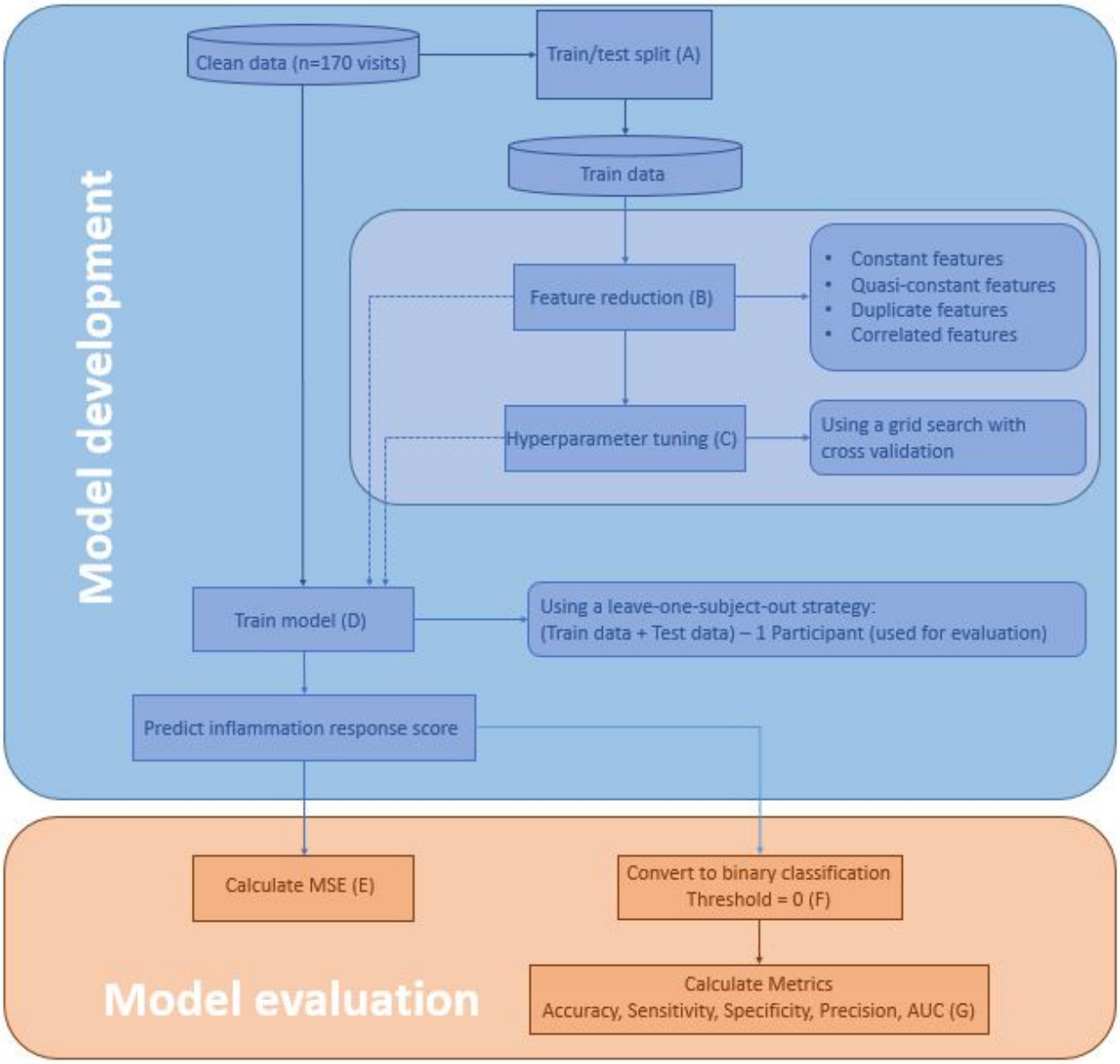

Figure 3.6. Model development and evaluation

Steps taken during the model development and evaluation phase are shown.

3.4. Results

3.4.1. Overview

The XGBoost models were able to predict the IRS based on features derived from the wearable sensors (Table 3.5). Models trained on the activity period (3 minutes of stair activity) achieved a maximum AUC of 0.66, the highest one was trained on HR features only. Models trained on the first minute of recovery achieved AUC scores of 0.43, 0.60, 0.69

and 0.73 for the BR, HR, MV, and HRV feature sets respectively. The highest performing model (recovery, HRV, AUC = 0.73) achieved a sensitivity of 0.7, a specificity of 0.77, and a precision of 0.75. The random forest model that leverages HRV features during recovery (HRV-RF) demonstrated lower performance compared to the XGBoost model.

Table 3.5. Model results according to the source of the features during activity vs recovery

Period	Features	MSE	Sensitivity	Specificity	Precision	AUC
Combined		0.08 (0.06)	0.38 (0.41)	0.92 (0.14)	0.67 (0.47)	0.65 (0.25)
Activity	BR	0.14 (0.12)	0.00 (0.00)	0.72 (0.27)	0.00 (0.00)	0.36 (0.13)
	HR	0.07 (0.07)	0.47 (0.45)	0.85 (0.20)	0.67 (0.24)	**0.66 (0.23)**
	MV	0.11 (0.09)	0.10 (0.20)	0.90 (0.20)	0.50 (0.50)	0.50 (0.16)
Recovery	BR	0.14 (0.16)	0.00 (0.00)	0.84 (0.20)	0.00 (0.00)	0.43 (0.10)
	HR	0.10 (0.11)	0.40 (0.49)	0.80 (0.24)	0.62 (0.12)	0.60 (0.12)
	MV	0.10 (0.09)	0.50 (0.45)	0.88 (0.14)	0.62 (0.41)	0.69 (0.26)
	HRV	0.13 (0.14)	0.70 (0.40)	0.77 (0.29)	0.75 (0.25)	**0.73 (0.16)**
	HRV-RF	0.14 (0.19)	0.40 (0.49)	0.83 (0.21)	0.50 (0.41)	0.62 (0.23)

Abbreviations: BR (breathing rate), HR (heart rate),
MSE (mean squared error), MV (minute ventilation),
HRV (heart rate variability), AUC (area under the curve).
Highest AUC per period (activity, recovery) shown in bold font.
Mean (Standard deviation) across the the 5 test participants shown.

3.4.2. A closer look at the best-performing model

Further details related to the model with the strongest performance (recovery period, HRV, AUC = 0.73) are now presented. The model comprises 8 features (see Table 3.6 and Figure 3.7 for a description of each feature and relative feature importance, respectively). Feature importance is a measure that quantifies the contribution of each input feature to the predictive performance of the model. It helps us understand which features have the most significant influence on the model's predictions. XGBoost measures feature importance by computing the gain achieved from each feature across all trees in the model. The gain represents the improvement in accuracy or reduction in loss resulting from a particular feature when used in a split decision within the trees. Features contributing to higher gain values are considered more important.

$$\text{Feature Importance} = \frac{\text{Sum of Gains for All Splits on a Feature}}{\text{Total Sum of Gains in the Model}} \tag{3.4.1}$$

Here, the gain is calculated for each feature as it is used in each decision tree for splitting. The importance score for a feature is then determined by aggregating the gains across all trees, normalizing these values by the total sum of gains in the model.

Table 3.6. Most important features

Feature name	Feature importance
lf_hf_ratio	0.174
std_hr	0.172
peaks	0.167
n_above_mean	0.142
skewness	0.123
range_nni	0.095
kurtosis	0.067
mean_nni	0.059

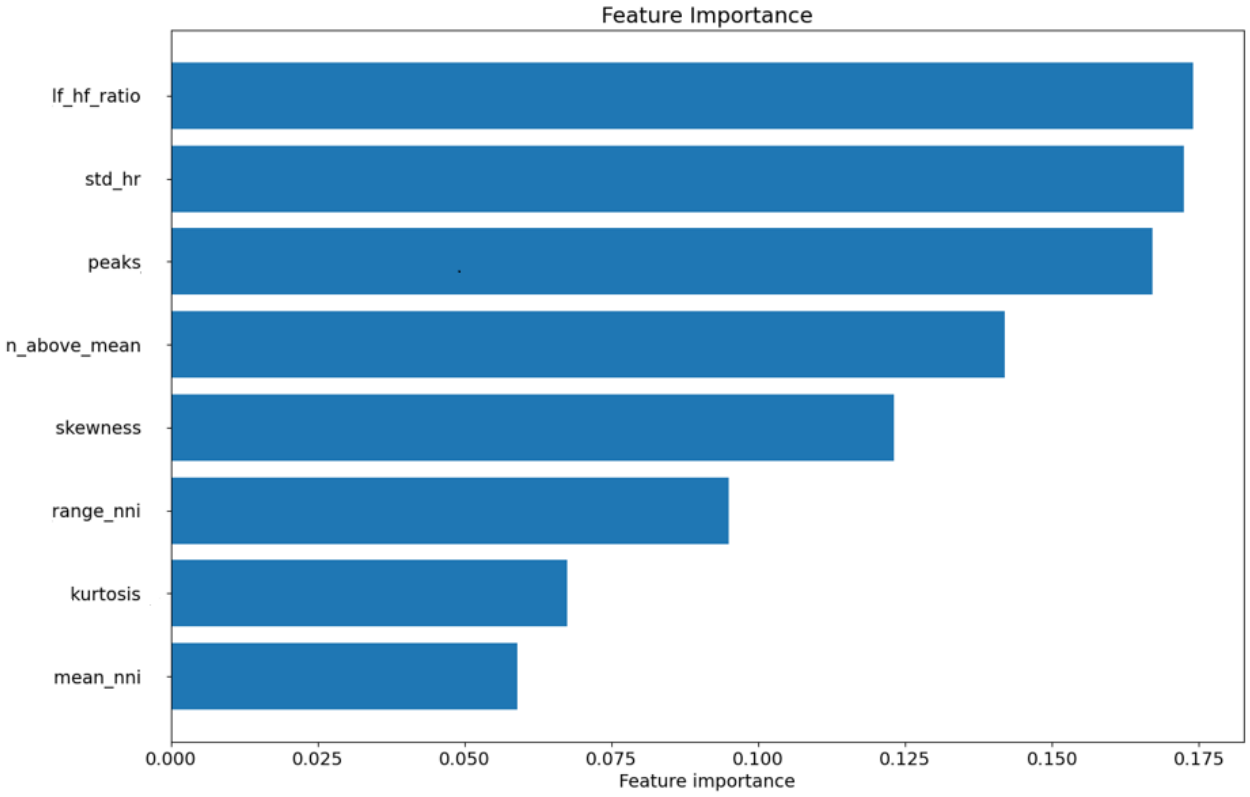

Figure 3.7. Feature importance of the best performing model

The most important feature is the low-frequency to high-frequency ratio with an importance of 0.174. The second and third most important features are the heart rate's standard deviation and the number of RRI peaks with an importance of 0.172 and 0.167, respectively.

A confusion matrix is presented (Figure 3.8).

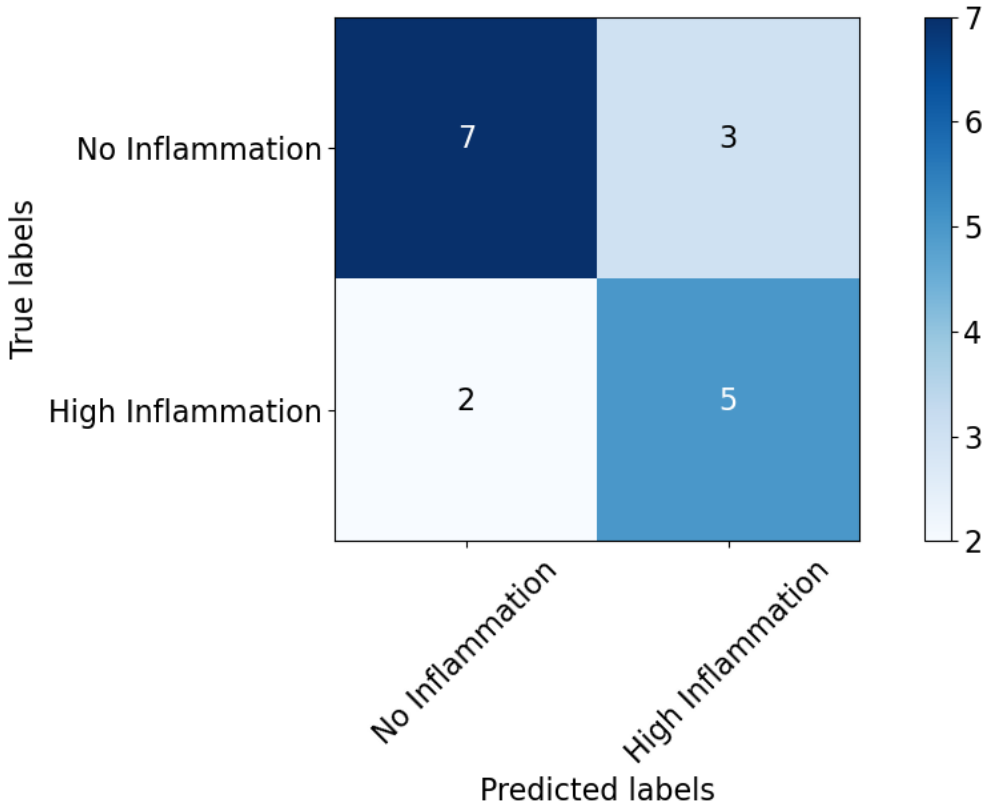

Figure 3.8. Confusion matrix of the best-performing model.

The confusion matrix provided here is a tabular representation used to evaluate the performance of our model. It consists of two classes: "no inflammation" and "high inflammation". The rows of the matrix represent the true labels, while the columns represent the predicted labels.

3.4.3. Effect of binary conversion threshold on model performance

For the best model, the binary threshold was varied over the interval [-0.3, 0.3] in steps of 0.1 which allows to obtain the ROC/AUC curve (Figure 3.9) (note that we cannot generate such curve automatically using Scikit Learn's pre-existing methods since this is a regression task that was later converted to a classification task). Increasing thresholds led to maximizing specificity while decreasing thresholds maximized sensitivity (Table 3.7).

Table 3.7. Effect of the threshold on sensitivity and specificity values

Threshold	Sensitivity	Specificity	True positives (visits)	True negatives (visits)
-0.3	1.00	0.00	16	0
-0.2	1.00	0.30	7	2
-0.1	1.00	0.57	7	5
0	0.70	0.77	5	7
0.1	0.00	0.83	0	11
0.2	0.00	0.83	0	11
0.3	0.00	1.00	0	15

Figure 3.9 is the corresponding ROC/AUC curve.

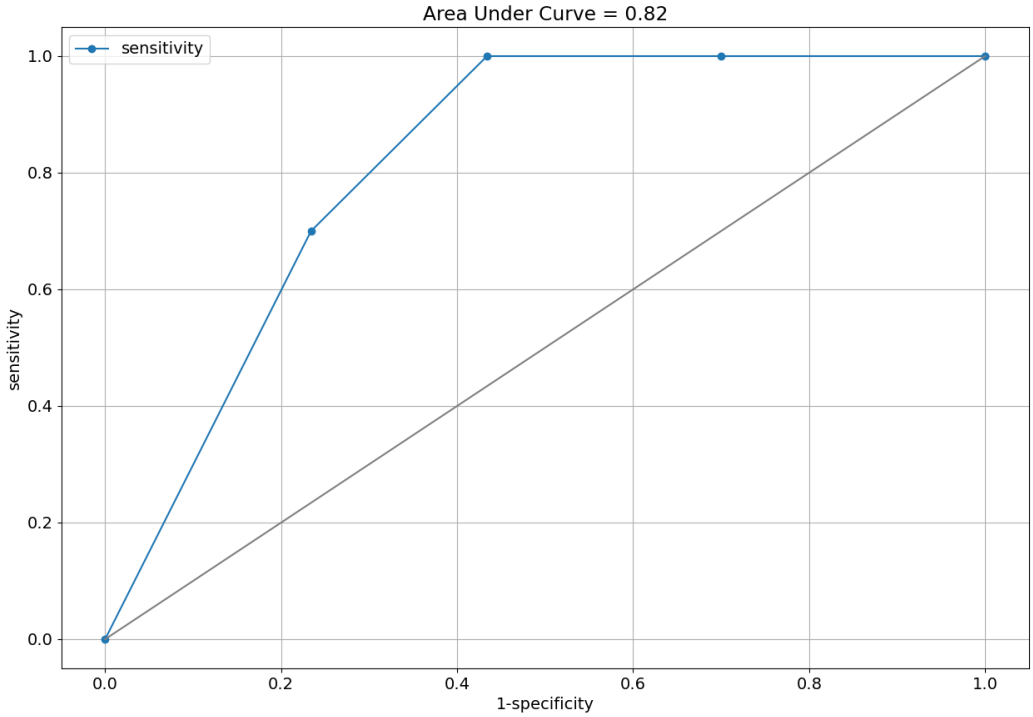

Figure 3.9. ROC/AUC curve of the best performing model

A scatter plot of the predicted vs the true IRS values is shown in Figure 3.10.

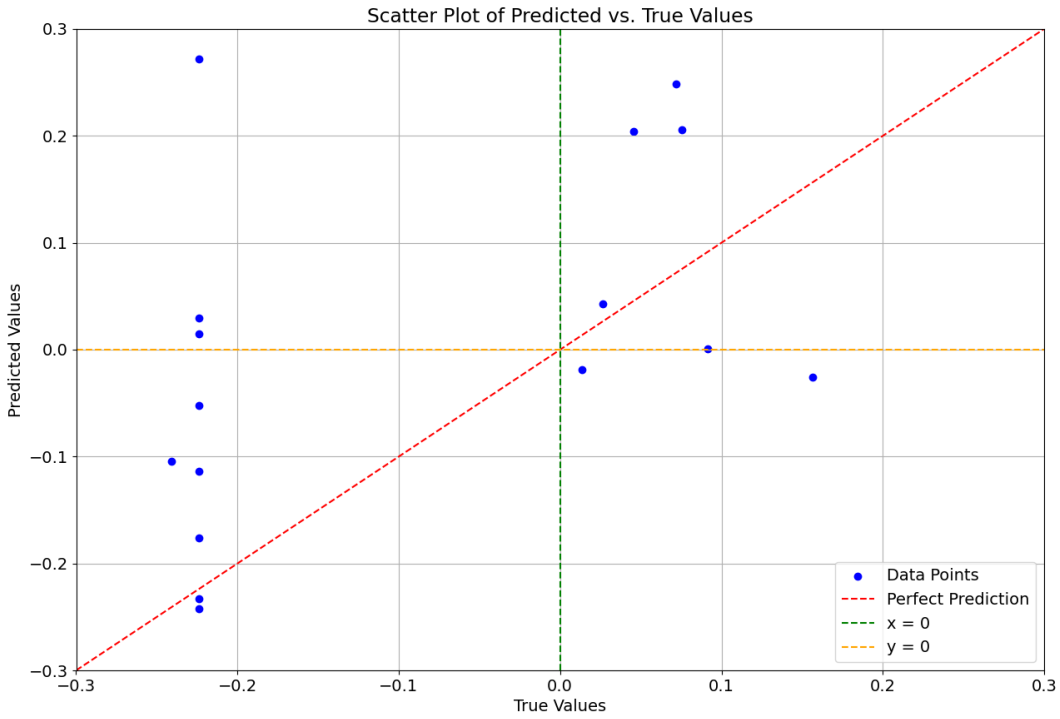

Figure 3.10. Scatter plot of the predicted vs true IRS values

This scatter plot shows the relationship between predicted and true values. Each point on the plot represents a predicted value. The plot includes a dashed diagonal line representing perfect prediction (where predicted equals true). In this scatter plot, we can geometrically see that the threshold of 0 sets apart the classes "no inflammation" and "high inflammation". The top-left as well as the bottom-right sides of the plot delimited by the dashed lines (x=0 and y=0) contain the misclassified predictions.

3.5. Discussion

3.5.1. Summary

This study aimed to assess the potential of wearables for early detection of VRTIs using machine learning algorithms and data collected during a controlled physical exercise (3-minute stair test). The model was able to predict the inflammation level (IRS) using features derived from the wearable device. We were able to obtain the highest predictive ability using HRV features during the first minute of recovery after the stair-stepping task.

3.5.2. Study design

The study design encompassed a structured timeline and a series of carefully planned visits and procedures. Even though moderate symptoms associated with viral replication are anticipated after the LAIV inoculation, we expected the participants' immune systems to have a similar response as if exposed to mild influenza. The choice of the number of days pre-inoculation is justified by the fact that we wanted to establish a general baseline by capturing the most truthful trend within the participants' daily lifestyle since a person's physiological activity is not the same during all days of the week. Additionally, the choice of exactly 5 days post-inoculation is due to the fact that we wanted to capture all physiological states (no inflammation, high inflammation), especially the peak of inflammation, which is estimated to happen during the first four days after inoculation [**42**]. Additionally, the use of the Astroskin/Hexoskin biosensor is backed up by its ability to ensure a non-invasive and convenient solution for collecting real-time data, allowing for a comprehensive assessment of participants' respiratory system throughout the study.

3.5.3. Effect of features on model performance

In line with our second objective, we demonstrated that features derived from different data sources have different predictive abilities. The model utilizing BR features during the activity period exhibited limited performance. Similarly, the MV during the activity period also displayed relatively lower performance metrics. However, the model trained on HR features during the activity period performed better than the other features in the same period. In contrast, models focused on the recovery period based on BR and MV features showed improvement compared to their activity-based counterparts. Notably, the MV features during the recovery period exhibited higher AUC and specificity, suggesting its potential relevance in predicting respiratory tract infections during recovery periods. Moreover, the model trained on HRV during the first minute of recovery revealed highest sensitivity, precision, and AUC values among the analyzed features. Finally, the model using the combined features derived from the HR, BR, MV, and RRI signals demonstrated poor performance metrics. This could be explained by the high number of features compared to the low number of visits which deepens the complexity of the relation between the different data points. These results demonstrate that HRV features are particularly important in enhancing the model's predictive performance. These findings are aligned with previous research [**55, 59, 62, 61**], reporting that HRV may act as a physiomarker of cardiac wellness during post-exercise periods.

3.5.4. Effect of binary conversion threshold on model performance

Our study design allows setting a binary conversion threshold to the IRS. An additional exploratory objective was to assess the effect of varying the threshold level on model performance. This analysis confirms that the most balanced tradeoff between high sensitivity and specificity is attained when setting a threshold of 0. Other threshold values result in either higher sensitivity or specificity. This approach enables flexibility in adjusting the sensitivity and specificity of the model and it becomes particularly significant when considering the real-world implementation of our models in different scenarios and settings.

To ensure the most effective utilization of our model, we recommend setting the threshold value in accordance with local public-health goals. The choice of the threshold should align with the specific objectives of the given context, taking into account the desired outcome and priorities.

In scenarios where our utmost focus lies in curbing the spread of infections. It becomes essential to prioritize sensitivity by selecting a lower threshold. This strategic decision aims to capture as many true positive cases as possible while reducing the probability of encountering false negatives. Early detection of individuals with VRTIs allows for the implementation of necessary measures that effectively deter any chance of further transmission. On the contrary, certain situations require avoiding unnecessary quarantines or disruption in infected individuals' lives. In such circumstances, adopting a higher threshold that prioritizes specificity proves advantageous. By doing so, we aim to minimize the number of individuals erroneously perceived as infected, thus preventing unnecessary isolation or quarantine measures.

3.5.5. Comparison with previous studies

Our results are not directly comparable to existing studies due to the uniqueness of our approach. It must be regarded that variations in approaches, methodologies, features and objectives between the studies have an impact and should be taken into consideration. Most studies relied on patient-reported diagnoses rather than quantitative measurements of inflammatory markers to predict VRTIs. In the context of VRTIs caused by COVID-19, Quer et al. [**24**] reported an AUC of 0.72 when relying on resting heart rate, sleep, and activity metrics. Mason et al. [**39**] and Conroy et al. [**40**] achieved AUCs of 0.82 (both studies) with sensitivity of 82 and 62% and specificity of 63 and 88%, respectively. Our model shows decreased performance compared to these studies (AUC of 0.73); however, our results are achieved through objective markers extracted from wearable devices. Moreover, our "diagnoses" (no inflammation vs high inflammation) are likely obtained more quickly, prior to the onset of symptoms, making our approach more useful in public-health settings.

3.5.6. Limitations

Several limitations are worth noting. The 3-minute stair test was performed under controlled laboratory conditions. Potential participants could reproduce this task in real-world settings, but the quality of task execution may affect the performance of our algorithms. Ideally, strenuous physical activity (and recovery) would be automatically detected using the wearable devices and algorithms would be robust enough to perform under these types of uncontrolled activities. This makes further investigations imperative for validating these findings in diverse settings that account for representing real-world conditions. In addition, missing and noisy data led to a reduction in the dataset size. The utilization of consumer-grade wearable devices (smaller and more convenient for every-day use compared to the Astroskin that comes with a head band and requires the user to put a relatively large device measuring the acceleration in the correct orientation) for data collection, while not eliminating the possibility of recording missing and noisy data, would allow for more extensive data collection to be conducted. Furthermore, the group of participants in our study was limited since we specifically focused on healthy adults. It is important to recognize that extending our conclusions to groups like the elderly or vulnerable populations may require additional investigation. These groups may exhibit variations in their physiological responses to infections and unique profiles of biomarker changes. Consequently, future research should aim to replicate and adapt our methodology to these diverse populations, ensuring broader applicability and effectiveness of our early detection approach. Finally, it is essential to recognize the limitations of our study concerning the specificity of detectable inflammations. Our emphasis on respiratory inflammation and specific biomarkers associated with this context may not encompass the full spectrum of inflammatory responses within the body. Inflammatory processes can vary significantly depending on underlying causes and affected systems. While our model displays promise in detecting respiratory tract infections, its specificity may be more confined to this particular inflammatory context. Researchers should exercise caution when applying our findings to broader inflammatory conditions, considering the need for context-specific models in such cases.

3.6. Conclusions

The findings validate the effectiveness of wearable biosensors and machine-learning techniques for early detection of VRTIs. By leveraging HRV features, the model exhibited notable accuracy in predicting the inflammation levels due to the onset of respiratory infections. This highlights the importance of incorporating quantitative measurements and advanced data analysis approaches in healthcare settings. The study's outcomes have significant implications for public health, as early detection can enable timely intervention and mitigation of future outbreaks. Ultimately, we aim to implement the developed model on

a user-friendly platform accessible to the general public. This would enable early detection of VRTIs and empower individuals to monitor their physiological signals and potentially identify signs of inflammation, allowing for proactive measures to be taken.

3.7. Conflicts of interest

P.C.D. has previously worked for Carré Technologies, Inc. (makers of the Hexoskin/Astroskin devices) and currently acts as a consultant for the company.

Chapter 4

Conclusions

The primary goals of this study were to assess the potential of wearable biosensors and machine learning algorithms for early detection of VRTIs and to identify the most predictive features of inflammation related to such infections. Our results demonstrate the feasibility and effectiveness of our approach in predicting the level of inflammation using HRV features derived from wearable devices during a controlled physical exercise, specifically a 3-minute stair test followed by 2 minutes of recovery. Wearable technologies enable continuous monitoring outside of clinical settings without requiring active patient participation, which makes self-monitoring easier and may lessen the burden on the public health system because early detection allows for prompt intervention and outbreak mitigation in the future.

Notably, our study design incorporated a structured timeline and meticulously planned visits and procedures. By selecting a pre-inoculation period of seven days, we aimed to establish a comprehensive baseline capturing participants' physiological activity within their daily lifestyle, considering that activity levels may vary throughout the week. The choice of the post-inoculation period of exactly five days allowed us to capture the different physiological states, including the peak of inflammation that occurs within the first four days after inoculation. By utilizing objective and quantitative data, we aimed to overcome the reliance on patient-reported symptoms, enabling a more accurate assessment of VRTI progression.

Regarding the analysis of predictive indicators, our findings revealed varying performance of features derived from different data sources. Features related to HRV during the first minute of recovery after the stair test exhibited the highest potential in predicting the level of inflammation, demonstrating the importance of HRV in enhancing the model's performance.

Furthermore, the choice of the binary conversion threshold for normalized IRS allowed us to explore the impact of different threshold levels on model performance. We found that setting a threshold of 0 yielded the most balanced tradeoff between high sensitivity and specificity. This flexible approach enables the adjustment of sensitivity and specificity based on the specific objectives and priorities of local public health goals and different real-world

implementation scenarios.

When comparing our study to existing research, it is important to acknowledge the uniqueness of our approach. While previous studies relied on patient-reported diagnoses or subjective measurements, our study employed objective markers automatically extracted from wearable devices along with biomarkers which are established indicators of inflammation during its early stages. Although our models showed slightly lower performance compared to studies utilizing different methodologies, our approach has the advantage of providing early detection prior to symptom onset, making it particularly valuable in public health settings.

Our study has several limitations worth noting. Firstly, we conducted the 3-minute stair test under controlled laboratory conditions, which may not accurately reflect its real-world performance, potentially impacting the algorithm's effectiveness. Secondly, challenges related to missing and noisy data prompted us to reduce the dataset size. The utilization of consumer-grade wearable devices for data collection could mitigate these issues, enabling more comprehensive data acquisition. Regarding the applicability of our findings, our research primarily focused on young healthy individuals. It is crucial to acknowledge that physiological responses and biomarker profiles may vary across groups, such as the elderly or vulnerable populations when establishing early detection of respiratory tract infections. To strengthen the reliability of our results, future studies should consider diverse sample sizes. Furthermore, it's important to acknowledge that our study's focus on respiratory inflammation and specific biomarkers may limit its specificity to this particular inflammatory context. Inflammatory processes vary widely based on causes and affected systems. Therefore, while our model shows promise in detecting respiratory tract infections, caution is needed when applying our findings to broader inflammatory conditions, necessitating context-specific models in such cases.

Looking ahead, our future work will focus on implementing the developed model on a user-friendly platform accessible to the general public. This would empower individuals to monitor their physiological signals and proactively identify signs of inflammation, facilitating early detection of viral respiratory tract infections. Additionally, further research should explore the validation of our findings in diverse populations and real-world settings. Through ongoing advancements in wearable technology and machine learning, we can continue to enhance our ability to detect and manage VRTIs, ultimately reducing their impact on public health.

References

[1] Magdalena Tuczy ´nska, Maja Matthews-Kozanecka, and Ewa Baum. Accessibility to non-covid health services in the world during the covid-19 pandemic: Review. *Frontiers in Public Health*, 9:760795, 2021.

[2] David M. Morens, Jeffery K. Taubenberger, and Anthony S. Fauci. Predominant role of bacterial pneumonia as a cause of death in pandemic influenza: Implications for pandemic influenza preparedness. *J Infect Dis*, 7:962–970, 2008.

[3] Christopher C Blyth, Steve A R Webb, Jen Kok, Dominic E Dwyer, Sebastiaan J van Hal, Hong Foo, Andrew N Ginn, Alison M Kesson, Ian Seppelt, and Jonathan R Iredell. The impact of bacterial and viral co-infection in severe influenza. *Influenza Other Respir Viruses*, 7:168–176, 2013.

[4] Nirav S. Shah, Jared A. Greenberg, Moira C. McNulty, Kevin S. Gregg, IV James Riddell, Julie E. Mangino, Devin M. Weber, Courtney L. Hebert, Natalie S. Marzec, Michelle A. Barron, Fredy Chaparro-Rojas, Alejandro Restrepo, Vagish Hemmige, Kunatum Prasidthrathsint, Sandra Cobb, Loreen Herwaldt, Vanessa Raabe, Christopher R. Cannavino, Andrea Green Hines, Sara H. Bares, Philip B. Antiporta, Tonya Scardina, Ursula Patel, Gail Reid, Parvin Mohazabnia, Suresh Kachhdiya, Binh-Minh Le, Connie J. Park, Belinda Ostrowsky, Ari Robicsek, Becky A. Smith, Jeanmarie Schied, Micah M. Bhatti, Stockton Mayer, Monica Sikka, Ivette Murphy-Aguilu, Priti Patwari, Shira R. Abeles, Francesca J. Torriani, Zainab Abbas, Sophie Toya, Katherine Doktor, Anindita Chakrabarti, Susanne Doblecki-Lewis, David J. Looney, and Michael Z. David. Bacterial and viral co-infections complicating severe influenza: Incidence and impact among 507 u.s. patients, 2013–14. *Journal of Clinical Virology*, 80:12–19, 2016.

[5] Balas M, Vasiliu D, Austria G, and Felfeli T. The impact of the covid-19 pandemic on wait-times for ophthalmic surgery in ontario, canada: A population-based study. *Clin Ophthalmol*, 17:1823–1831, 2023.

[6] Government of Canada. Covid-19 epidemiology update: Current situation. `https://health-infobase.canada.ca/covid-19/current-situation.html#graphHospVentICU`, 2023. Accessed on 2023-06-02.

[7] Adachi D, Johnson G, Draker R, Ayers M, Mazzulli T, Talbot PJ, and Tellier R. Comprehensive detection and identification of human coronaviruses, including the sars-associated coronavirus, with a single rt-pcr assay. *J Virol Methods*, 122:29–36, 2004.

[8] Setianingsih TY, Wiyatno A, Hartono TS, Hindawati E, Rosamarlina, Dewantari AK, Myint KS, Lisdawati V, and Safari D. Detection of multiple viral sequences in the respiratory tract samples of suspected middle east respiratory syndrome coronavirus patients in jakarta, indonesia 2015-2016. *Int J Infect Dis*, 86:102–107, 2019.

[9] Zhang W, Du RH, Li B, Zheng XS, Yang XL, Hu B, Wang YY, Xiao GF, Yan B, Shi ZL, and Zhou P. Molecular and serological investigation of 2019-ncov infected patients: implication of multiple shedding routes. *Emerging Microbes & Infections*, 9:386–389, 2020.

[10] James AS and Alawneh JI. Covid-19 infection diagnosis: Potential impact of isothermal amplification technology to reduce community transmission of sars-cov-2. *Diagnostics (Basel)*, 10:399, 2020.

[11] Guoguang Rong, Yuqiao Zheng, Yin Chen, Yanjun Zhang, Peixi Zhu, and Mohamad Sawan. Covid-19 diagnostic methods and detection techniques. *Encyclopedia of Sensors and Biosensors*, pages 17–32, 2023.

[12] Laura E. Lamb, Sarah N. Bartolone, Elijah Ward, and Michael B. Chancellor. Rapid detection of novel coronavirus/severe acute respiratory syndrome coronavirus 2 (sars-cov-2) by reverse transcription-loop-mediated isothermal amplification. *PLOS ONE*, 15:e0234682, 2020.

[13] WE Huang, B Lim, CC Hsu, D Xiong, W Wu, Y Yu, H Jia, Y Wang, Y Zeng, M Ji, H Chang, X Zhang, H Wang, and Z Cui. Rt-lamp for rapid diagnosis of coronavirus sars-cov-2. *Microb Biotechnol*, 13:950–961, 2020.

[14] X Zhu, X Wang, L Han, T Chen, L Wang, H Li, S Li, L He, X Fu, S Chen, M Xing, H Chen, and Y Wang. Multiplex reverse transcription loop-mediated isothermal amplification combined with nanoparticle-based lateral flow biosensor for the diagnosis of covid-19. *Biosens Bioelectron*, 166:112437, 2020.

[15] Weihua Yang, Xiaofei Dang, Qingxi Wang, Mingjie Xu, Qianqian Zhao, Yunying Zhou, Huailong Zhao, Li Wang, Yihui Xu, Jun Wang, Shuyi Han, Min Wang, Fenyan Pei, and Yunshan Wan. Rapid detection of sars-cov-2 using reverse transcription rt-lamp method. *MedRxiv*, 2020.

[16] S Lu, S Lin, H Zhang, L Liang, and S Shen. Methods of respiratory virus detection: Advances towards point-of-care for early intervention. *Micromachines (Basel)*, 6:697, 2021.

[17] Briony Hazelton, Timothy Gray, Jennifer Ho, V Mala Ratnamohan, Dominic E Dwyer, and Jen Kok. Detection of influenza a and b with the alere ™ i influenza a & b: a novel isothermal nucleic acid amplification assay. *Influenza Other Respir Viruses*, 9:151–154, 2015.

[18] Zahra Abbas, Mario Fiorino, Syed Muhammad Naqi, and Musarat Abbas. Covid-19 prediction infrastructure using deep learning. In *Proceedings of the 19th International Conference on Intelligent Environments (IE2023)*. IOS Press, 2023.

[19] A. Narin, C. Kaya, and Z. Pamuk. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. *Pattern Analysis and Applications*, 24,3:1207–1220, 2021.

[20] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015.

[21] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. Jamalipour Soufi. Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning. *Med Image Anal*, 65:101794, 2020.

[22] El Asnaoui K., Chawki Y., and Idri A. Automated methods for detection and classification pneumonia based on x-ray images using deep learning. *Artificial Intelligence and Blockchain for Future Cybersecurity Applications*, 90:257–284, 2021.

[23] Dina M. Ibrahim, Nada M. Elshennawy, and Amany M. Sarhan. Deep-chest: Multi-classification deep learning model for diagnosing covid-19, pneumonia, and lung cancer chest diseases. *Computers in Biology and Medicine*, 132:104348, 2021.

[24] G. Quer et al. Wearable sensor data and self-reported symptoms for covid-19 detection. *Nature Medicine*, 27:73–77, 2020.

[25] Tejaswini Mishra, Meng Wang, Ahmed A. Metwally, Gireesh K. Bogu, Andrew W. Brooks, Amir Bahmani, Arash Alavi, Alessandra Celli, Emily Higgs, Orit Dagan-Rosenfeld, Bethany Fay, Susan Kirkpatrick, Ryan Kellogg, Michelle Gibson, Tao Wang, Erika M. Hunting, Petra Mamic, Ariel B. Ganz,

Benjamin Rolnik, Xiao Li, and Michael P. Snyder. Pre-symptomatic detection of covid-19 from smartwatch data. *Nature Biomedical Engineering*, 4:1208–1220, 2020.

[26] Jayoung Kim, Alan S. Campbell, Berta Esteban-Fernández de Ávila, and Joseph Wang. Wearable biosensors for healthcare monitoring. *Nature Biotechnology*, 4:389–406, 2019.

[27] Hongyu Luo, Pierre-Alexandre Lee, Ieuan Clay, Martin Jaggi, and Valeria De Luca. Assessment of fatigue using wearable sensors: A pilot study. *Digit Biomark*, 4:59–72, 2020.

[28] C.M. Smith, S.N. Chillrud, D.W. Jack, et al. Laboratory validation of hexoskin biometric shirt at rest, submaximal exercise, and maximal exercise while riding a stationary bicycle. *Journal of Occupational and Environmental Medicine*, 61(4):e104–e111, 2019.

[29] Jesse Jayne Rutherford. Wearable technology. health-care solutions for a growing global population. *IEEE Engineering in Medicine and Biology Magazine*, 29(3):19–24, 2010.

[30] X Ding, D Clifton, N Ji, NH Lovell, P Bonato, W Chen, X Yu, Z Xue, T Xiang, X Long, K Xu, X Jiang, Q Wang, B Yin, G Feng, and YT Zhang. Wearable sensing and telehealth technology with potential applications in the coronavirus pandemic. *IEEE Reviews in Biomedical Engineering*, 14:48–70, 2021.

[31] MM Islam, S Mahmud, LJ Muhammad, MR Islam, S Nooruddin, and SI Ayon. Wearable technology to assist the patients infected with novel coronavirus (covid-19). *SN Computer Science*, 1(6):320, 2020.

[32] Shizuko Takahashi, Eisuke Nakazawa, Sakurako Ichinohe, Aru Akabayashi, and Akira Akabayashi. Wearable technology for monitoring respiratory rate and spo2 of covid-19 patients: A systematic review. *Diagnostics*, 12(10):2563, 2022.

[33] M Tayal, A Mukherjee, U Chauhan, M Uniyal, S Garg, A Singh, AS Bhadoria, and R Kant. Evaluation of remote monitoring device for monitoring vital parameters against reference standard: A diagnostic validation study for covid-19 preparedness. *Indian Journal of Community Medicine*, 45(2):235–239, 2020.

[34] W Jiang, S Majumder, S Kumar, S Subramaniam, X Li, R Khedri, T Mondal, M Abolghasemian, I Satia, and MJ Deen. A wearable tele-health system towards monitoring covid-19 and chronic diseases. *IEEE Reviews in Biomedical Engineering*, 15:61–84, 2022.

[35] Bret Nestor, Jaryd Hunter, Raghu Kainkaryam, Erik Drysdale, Jeffrey B Inglis, Allison Shapiro, Sujay Nagaraj, Marzyeh Ghassemi, Luca Foschini, and Anna Goldenberg. Machine learning covid-19 detection from wearables. *The Lancet Digital Health*, 5(4):182–184, 2023.

[36] M. Risch et al. Investigation of the use of a sensor bracelet for the presymptomatic detection of changes in physiological parameters related to covid-19: an interim analysis of a prospective cohort study (covigapp). *BMJ Open*, 12(6):e058274, 2022.

[37] A. Natarajan, H.W. Su, and C. Heneghan. Assessment of physiological signs associated with covid-19 measured using wearable devices. *NPJ Digital Medicine*, 3:156, 2020.

[38] A. Alavi et al. Real-time alerting system for covid-19 and other stress events using wearable data. *Nature Medicine*, 28:175–184, 2022.

[39] A.E. Mason et al. Detection of covid-19 using multimodal data from a wearable device: results from the first tempredict study. *Scientific Reports*, 12:3463, 2022.

[40] B. Conroy et al. Real-time infection prediction with wearable physiological monitoring and ai to aid military workforce readiness during covid-19. *Scientific Reports*, 12:3797, 2022.

[41] Richard Mayeux. Biomarkers: Potential uses and limitations. *NeuroRx*, 2:182–188., 2004.

[42] M. T. McClain, R. Henao, J. Williams, B. Nicholson, T. Veldman, L. Hudson, E. L. Tsalik, R. Lambkin-Williams, A. Gilbert, A. Mann, G. S. Ginsburg, and C. W. Woods. Differential evolution of peripheral

cytokine levels in symptomatic and asymptomatic responses to experimental influenza virus challenge. *Clinical and Experimental Immunology*, 183(3):441–451, 2015.

[43] J. Ten Oever, M.G. Netea, and B-J. Kullberg. Utility of immune response-derived biomarkers in the differential diagnosis of inflammatory disorders. *Journal of Infection*, 72:1–18, 2016.

[44] A.J. Kapasi, S. Dittrich, I.J. González, and T.C. Rodwell. Host biomarkers for distinguishing bacterial from non-bacterial causes of acute febrile illness: a comprehensive review. *PLOS ONE*, 11(8):e0160278, 2016.

[45] RT Jr Davey, R. Lynfield, D.E. Dwyer, et al. The association between serum biomarkers and disease outcome in influenza a (h1n1)pdm09 virus infection: results of two international observational cohort studies. *PLOS ONE*, 8(2):e57121, 2013.

[46] N. Lee, CK. Wong, PK. Chan, et al. Cytokine response patterns in severe pandemic 2009 h1n1 and seasonal influenza among hospitalized adults. *PLOS ONE*, 6(10):e26050, 2011.

[47] Y. Chi, Y. Zhu, T. Wen, et al. Cytokine and chemokine levels in patients infected with the novel avian influenza a (h7n9) virus in china. *Journal of Infectious Diseases*, 208(12):1962–1967, 2013.

[48] Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology. Heart rate variability - standards of measurement, physiological interpretation, and clinical use. *Circulation*, 93:1043–1065, 1996.

[49] DWP Williams, J. Keonig, L. Carnevali, et al. Heart rate variability and inflammation: a meta-analysis of human studies. *Brain, Behavior, and Immunity*, 80:219–226, 2019.

[50] A. Haarala, M. Kahonen, C. Eklund, J. Jylhava, T. Koskinen, L. Taittonen, et al. Heart rate variability is independently associated with c-reactive protein but not with serum amyloid a. the cardiovascular risk in young finns study. *European Journal of Clinical Investigation*, 41(9):951–958, 2011.

[51] Frederick Hasty, Guillermo García, COL Héctor Dávila, S. Howard Wittels, Stephanie Hendricks, and Stephanie Chong. Heart rate variability as a possible predictive marker for acute inflammatory response in covid-19 patients. *Military Medicine*, 86(1-2):e34–e38, 2021.

[52] M. Timothy, Paula S. McKinley, Teresa E. Seeman, Tse-Hwei Choo, and Seonjoo Lee. Heart rate variability predicts levels of inflammatory markers: Evidence for the vagal anti-inflammatory. *Brain, Behavior, and Immunity*, 49:94–100, 2015.

[53] Nikhil Singh, Kegan James Moneghetti, Jeffrey Wilcox Christle, David Hadley, Daniel Plews, and Victor Froelicher. Heart rate variability: An old metric with new meaning in the era of using mhealth technologies for health and exercise training guidance. part one: Physiology and methods. *Arrhythm Electrophysiol Rev*, 7(3):193–198, 2018.

[54] Denis Mongin, Clovis Chabert, Manuel Gomez Extremera, Olivier Hue, Delphine Sophie Courvoisier, Pedro Carpena, and Pedro Angel Bernaola Galvan. Decrease of heart rate variability during exercise: An index of cardiorespiratory fitness. *PLOS ONE*, 17(9):e0273981, 2022.

[55] D. A. Boullosa, E. S. Barros, S. del Rosso, F. Y. Nakamura, and A. S. Leicht. Reliability of heart rate measures during walking before and after running maximal efforts. *International Journal of Sports Medicine*, 35(12):999–1005, 2014.

[56] M. Buchheit and C. Gindre. Cardiac parasympathetic regulation: respective associations with cardiorespiratory fitness and training load. *American Journal of Physiology - Heart and Circulatory Physiology*, 291:H451–H458, 2006.

[57] CR Cole, EH Blackstone, FJ Pashkow, CE Snader, and MS Lauer. Heart rate recovery immediately after exercise as a predictor of mortality. *New England Journal of Medicine*, 341:1351–1357, 1999.

[58] TA Dewland, AS Androne, FA Lee, RJ Lampert, and SD Katz. Effect of acetylcholinesterase inhibition with pyridostigmine on cardiac parasympathetic function in sedentary adults and trained athletes. *American Journal of Physiology - Heart and Circulatory Physiology*, 293:H86–H92, 2007.

[59] PJ Kannankeril, FK Le, AH Kadish, and JJ Goldberger. Parasympathetic effects on heart rate recovery after exercise. *Journal of Investigative Medicine*, 52:394–401, 2004.

[60] H. Tsuji, MG Larson, FJ Jr Venditti, ES Manders, JC Evans, CL Feldman, and D Levy. Impact of reduced heart rate variability on risk for cardiac events. the framingham heart study. *Circulation*, 94:2850–2855, 1996.

[61] J. J. Goldberger, F. K. Le, M. Lahiri, P. J. Kannankeril, J. Ng, and A. H. Kadish. Assessment of parasympathetic reactivation after exercise. *American Journal of Physiology - Heart and Circulatory Physiology*, 290:H2446–H2452, 2006.

[62] J. Stanley, J. M. Peake, and M. Buchheit. Cardiac parasympathetic reactivation following exercise: implications for training prescription. *Sports Medicine*, 43:1259–1277, 2013.

[63] S. Carrasco, R. González, M. J. Gaitán, and O. Yáñez. Reproducibility of heart rate variability from short-term recordings during five manoeuvres in normal subjects. *Journal of Medical Engineering and Technology*, 27:241–248, 2003.

[64] M.Z. Shen, Y. Zhou, J.W. Ye, A.A.A. Al-Maskri, Y. Kang, et al. Recent advances and perspectives of nucleic acid detection for coronavirus. *Journal of Pharmaceutical Analysis*, 10(2):97–101, 2020.

[65] V.M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, et al. Detection of 2019 novel coronavirus (2019-ncov) by real-time rt-pcr. *Euro Surveillance*, 25(3):2000045, 2020.

[66] J.M. Radin, N.E. Wineinger, E.J. Topol, et al. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the usa: a population-based study. *The Lancet Digital Health*, 2:E85–E93, 2020.

[67] J.C. Emery, T.W. Russell, Y. Liu, et al. The contribution of asymptomatic sarscov-2 infections to transmission on the diamond princess cruise ship. *eLife*, 9:e58699, 2020.

[68] Amir Hadid, Emily G McDonald, Matthew P Cheng, Jesse Papenburg, Michael Libman, Phillipe C Dixon, and Dennis Jensen. The we sense study protocol: A controlled, longitudinal clinical trial on the use of wearable sensors for early detection and tracking of viral respiratory tract infections. *Contemporary Clinical Trials*, 128:107103, 2023.

[69] JS Peiris, WC Yu, CW Leung, et al. Re-emergence of fatal human influenza a subtype h5n1 disease. *The Lancet*, 363(9409):617–619, 2004.

[70] Y. Monteerarat, S. Sakabe, S. Ngamurulert, et al. Induction of tnfα in human macrophages by avian and human influenza viruses. *Archives of Virology*, 155(8):1273–1282, 2010.

[71] E. Harbour, M. Lasshofer, M. Genitrini, et al. Enhanced breathing pattern detection during running using wearable sensors. *Sensors (Basel)*, 21(16):5606, 2021.

[72] J. Montes, J.C. Young, R. Tandy, et al. Reliability and validation of the hexoskin wearable bio-collection device during walking conditions. *International Journal of Exercise Science*, 11(7):806–816, 2018.

[73] N.J. Carter and M.P. Curran. Live attenuated influenza vaccine (flumist®; fluenztm). *Drugs*, 71(12):1591–1622, 2011.

[74] Maria S Garcia-Gutierrez, Francisco Navarrete, Francisco Sala, Ani Gasparyan, Amaya Austrich-Olivares, and Jorge Manzanares. Biomarkers in psychiatry: Concept, definition, types and relevance to the clinical reality. *Frontiers in Psychiatry*, 11:432, 2020.

[75] Fang Li, Chun-Hao Chang, Yu-Chun Chung, et al. Development and validation of 3 min incremental step-in-place test for predicting maximal oxygen uptake in home settings: A submaximal exercise study

to assess cardiorespiratory fitness. *International Journal of Environmental Research and Public Health*, 18(20):10750, 2021.

[76] Walid Soliman Ismail Mahmoud Elsaid. Evaluating the validity and reliability of harvard step test to predict vo2max in terms of the step height according to the knee joint angle. *Theories and Applications the International Edition*, 1:126–132, 2011.

[77] N. T. V. Kieu, S. J. Jung, S. W. Shin, H. W. Jung, E. S. Jung, Y. H. Won, Y. G. Kim, and S. W. Chae. The validity of the ymca 3-minute step test for estimating maximal oxygen uptake in healthy korean and vietnamese adults. *J Lifestyle Med*, 10(1):21–29, 2020.

[78] O. Lee, S. Lee, M. Kang, J. Mun, and J. Chung. Prediction of maximal oxygen consumption using the young men's christian association-step test in korean adults. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 119(5):1245–1252, 2019.

[79] H. Lewthwaite and D. Jensen. Multidimensional breathlessness assessment during cardiopulmonary exercise testing in healthy adults. *European Journal of Applied Physiology*, 121(2):499–511, 2021.

[80] S. P. Blackie, M. S. Fairbarn, N. G. McElvaney, et al. Normal values and ranges for ventilation and breathing pattern at maximal exercise. *Chest*, 100:136–142, 1991.

[81] S. Foll, M. Maritsch, F. Spinola, et al. Flirt: a feature generation toolkit for wearable data. *Computer Methods and Programs in Biomedicine*, 212:106461, 2021.

[82] Hexoskin api documentation: Datatype resource. `https://api.hexoskin.com/docs/resource/datatype/`, 2023.

[83] numpy.trapz. `https://numpy.org/doc/stable/reference/generated/numpy.trapz.html`, 2022.

[84] Xgboost. `https://pypi.org/project/xgboost/`, 2023.

[85] Tianqi Chen and Carlos Guestrin. Xg boost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, 2016. ACM.