npj | digital medicine

Article

Published in partnership with Seoul National University Bundang Hospital

https://doi.org/10.1038/s41746-024-01287-2

Feasibility of snapshot testing using wearable sensors to detect cardiorespiratory illness (COVID infection in India)

Olivia K. Botonis[1], Jonathan Mendley[1], Shreya Aalla[1], Nicole C. Veit[1,2], Michael Fanton[1,3], JongYoon Lee[4], Vikrant Tripathi[5], Venkatesh Pandi[6], Akash Khobragade[7], Sunil Chaudhary[8], Amitav Chaudhuri[9], Vaidyanathan Narayanan[10], Shuai Xu[4], Hyoyoung Jeong[11,12], John A. Rogers[2,11,13,14] & Arun Jayaraman[1,3]

The COVID-19 pandemic has challenged the current paradigm of clinical and community-based disease detection. We present a multimodal wearable sensor system paired with a two-minute, movement-based activity sequence that successfully captures a snapshot of physiological data (including cardiac, respiratory, temperature, and percent oxygen saturation). We conducted a large, multi-site trial of this technology across India from June 2021 to April 2022 amidst the COVID-19 pandemic (Clinical trial registry name: International Validation of Wearable Sensor to Monitor COVID-19 Like Signs and Symptoms; NCT05334680; initial release: 04/15/2022). An Extreme Gradient Boosting algorithm was trained to discriminate between COVID-19 infected individuals ($n = 295$) and COVID-19 negative healthy controls ($n = 172$) and achieved an F1-Score of 0.80 (95% CI = [0.79, 0.81]). SHAP values were mapped to visualize feature importance and directionality, yielding engineered features from core temperature, cough, and lung sounds as highly important. The results demonstrated potential for data-driven wearable sensor technology for remote preliminary screening, highlighting a fundamental pivot from continuous to snapshot monitoring of cardiorespiratory illnesses.

Nearly four years after the World Health Organization's (WHO) first recognized the emergence of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), a novel coronavirus causing COVID-19, individuals around the world are still impacted by its devastating consequences[1]. As the virus transitions into an endemic phase[2], at which point the infection rate will persist indefinitely[3,4], consistent monitoring of cases and disease transmission rates is vital to prevent further outbreaks[5,6]. Vaccines have demonstrated effectiveness in protecting against severe disease, hospitalization, and death[7], however breakthrough infection cases are inevitable due to acquired immunity waning over time[4]. Immunity is further undermined by viral mutations, as observed in vaccinated individuals contracting Omicron subvariants of COVID-19[8]. Variants of the SARS-CoV-2 (Severe Acute Respiratory Syndrome-Coronavirus-2) virus continue to emerge in countries around the world[9] and are becoming increasingly resistant to neutralization[10]. Further, because the signs and symptoms of COVID-19 can affect several organ systems, clinical

[1]Max Nader Lab for Rehabilitation Technologies and Outcomes Research, Shirley Ryan AbilityLab, Chicago, IL, USA. [2]Department of Biomedical Engineering, Northwestern University, Evanston, IL, USA. [3]Department of Physical Medicine and Rehabilitation, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. [4]Sibel Health, Niles, IL, USA. [5]Clinfinite Solutions, Hyderabad, Telangana, India. [6]Induss Hospital, Hyderabad, Telangana, India. [7]Grant Medical College and Sir Jamshedjee Jeejeebhoy Group of Hospitals, Mumbai, Maharashtra, India. [8]Lifepoint Multispecialty Hospital, Pune, Maharashtra, India. [9]Timetooth Technologies Pvt Ltd, Noida, Uttar Pradesh, India. [10]Bionic Yantra, Bengaluru, Karnataka, India. [11]Center for Bio-Integrated Electronics, Northwestern University, Evanston, IL, USA. [12]Department of Electrical and Computer Engineering, University of California Davis, Davis, CA, USA. [13]Department of Mechanical Engineering, Northwestern University, Evanston, IL, USA. [14]Department of Materials Science and Engineering, Northwestern University, Evanston, IL, USA.
e-mail: a-jayaraman@northwestern.edu

presentation is often heterogeneous and severity of symptoms can vary greatly[2,11]. Variable presentations of disease states may hinder current and future efforts to control transmission if infections are not detected early on.

Large-scale, high volume clinical testing can effectively reduce the spread of disease and promote earlier, targeted intervention[12]. To facilitate and boost testing, advancements can be made in the rapidness, accuracy, and scalability of testing options. The gold standard for molecular testing (nasopharyngeal swab (NPS) reverse transcription-quantitative polymerase chain reaction (RT-qPCR)[13]) requires specialized instrumentation and is unable to provide immediate time-sensitive results[14–16]. Comparatively, rapid antigen tests (RAT) offer a portable[14] and easily scalable[17] option for prompt point-of-care screening[18–20] but compensate with lower detection rates (overall pooled sensitivity of 68.4%) and reliance on additional resources to confirm a diagnosis[17,21,22]. The resulting impacts of false negative tests can be drastic. Besides allowing infected people to spread disease, they can depress testing rates by eroding trust in health care systems[23] and disincentivizing potential close contacts from getting tested when they might have otherwise done so[24].

Although both testing options have been well integrated into some societies, inequitable access is an on-going problem in developing countries. Despite low- and lower-middle-income countries making up 76.3% of the global population, only 36.9% of all tests across the world were used in these countries[25]. The WHO's Access COVID-19 Tools Accelerator (ACT-A) set its testing target for 1 test per 1000 people per day[26], but this target was underachieved predominantly by low- and lower-middle-income countries in the past 12 months as of January 6, 2023[25]. These regions have also particularly struggled with related plastic waste management and environmental contamination[27]. RT-PCR testing alone has created an alarming quantity of plastic waste in the past four years[28], much of which is non-biodegradable[29], creating a large and lasting environmental impact.

Consequently, there has been a remarkable upsurge in research to develop novel approaches for COVID-19 testing, some of which address concerns related to accessibility, implementation, and reliance on single-use materials. Notably, wearable sensing technology has been proposed as a viable alternative. Advancements in materials engineering have permitted low-profile, cost-effective sensing platforms conducive to large-scale use[30]. Triaxial accelerometer-based sensors have demonstrated potential to analyze changes in physiological patterns, specifically cough[31,32]. Additionally, wearable sensors interfacing with the skin can capture targeted clinical measures at high levels of accuracy (e.g., heart rate, respiration rate, body temperature, and percent oxygen saturation) as well as provide insight into the quality of health over time (e.g., sleep quality[33], physical activity[34,35], and disease diagnosis and treatment[36]). The ability of consumer-grade wearables, including smartwatches[37–39] and smart rings[40,41], to detect and monitor probable COVID-19 infection and other cardiorespiratory illnesses has been previously investigated. These systems often operate on continuous data which can provide important insights related to exposure tracking and long-term effects of illness[39,42–45]. While promising, single-user consumer technologies are not ubiquitous in all societies[46] and can limit usability in point-of-care clinical testing. Furthermore, the robustness of continuous data collection is still unclear, as continuous collection is known to generate large sums of noisy, heterogeneous data[47,48] and is affected by wear-time compliance[49].

To address these challenges, our team previously investigated the use of shareable wearable sensor systems to measure physiological symptoms of illness within a rapid "snapshot" of data[50]. Our novel methodology paired clinically meaningful physiological signal features with an activity-based movement sequence spanning less than two minutes to quantify the probability of a cardiorespiratory infection. We have since adopted a more advanced, comprehensive, and compact sensor platform, capable of capturing multimodal measurements including physical activity, cardiorespiratory function, and percent oxygen saturation. It also measures dermal temperature, giving it a distinct advantage over other commercial wearable devices[40]. The goal of the present study was to evaluate the feasibility of collecting "snapshot" data with this updated sensor. We demonstrated this technology's potential for detecting cardiorespiratory illnesses, in this case COVID-19, via real-world deployment of this sensor in India during the COVID-19 pandemic.

The technology shown in this work has the potential to be implemented as a rapid and reusable screening tool, proactively warning users presenting with cardiorespiratory symptoms of their infection. Shared diagnostic tools may effectively reduce reliance on single-use testing materials, offering an alternative option for preliminary screening at scale in the case of future pandemics or for long-term community-based monitoring of disease states. As an intended outcome of this work, we hope to provide supporting evidence for future applications of multimodal wearable sensing systems to address gaps in the implementation of current diagnostic and monitoring technologies, specifically using only a snapshot of physiologically relevant data. Innovative applications could become impactful in addressing global health disparities, especially in remote and under-resourced communities.

Results

Metadata

Data collection spanned across four different site locations over approximately nine months (June 2021 to April 2022) including 85 days of active data collection. We intended to recruit 400 participants and successfully enrolled 532 individuals in the study. However, some samples were dropped at the initial signal segmentation stage due to poor signal quality (defined by noise that obstructed identification of individual activities) or incorrect performance of the activity sequence (both confirmed by visually inspecting the signal), or download failure. Furthermore, we only considered individuals in the COVID cohort who were within 14 days of a positive RT-qPCR (reverse transcription-quantitative polymerase chain reactions) diagnosis (to remove confounding longitudinal effects of illness). The remaining 467 subjects (i.e., 295 COVID and 172 NON-COVID) were used in model development and testing. Each subject contributed one data sample to the final feature matrix. The corresponding metadata for those samples is reported in Table 1. Table 2 shows the percentages of feature availability for each sensor modality.

Hyperparameters

The best hyperparameters resulting from the internal cross-validation search are summarized over all model splits in Supplementary Table 1. We provide here the average and standard deviation of each; maximum tree depth = 6.07 (Std. 2.36); minimum child weight = 4.39 (Std. 2.23); learning rate = 0.26 (Std. 0.17); gamma = 2.0 (Std. 1.43). Trends on the effect of hyperparameter value on F1-Score can be visualized in Supplementary Fig. 1.

Model performance

The results of the trained XGBoost algorithm to classify COVID from NON-COVID individuals in the held-out test set are shown in Fig. 1. The average confusion matrix for a given test set and the distributions of the F1-score, recall, and precision are shown for the 100 model iterations in Fig. 1a. The mean F1-score was 0.80 (95% CI = [0.79, 0.81]), with the best performing model run achieving an F1-Score of 0.88 (see decision tree structure in Supplementary Fig. 2). The predicted probability of COVID-19 can be visualized in the distributions in Fig. 1b. Predicted values near one have higher model certainty for being classified as COVID, and lower values near zero have more certainty for being labeled as NON-COVID. As shown, true positive and true negative predictions were classified with higher certainty than false positive and false negative classifications, which were distributed more centrally around predictive values of 0.5, i.e., low certainty. Figure 1c displays the model performance (AUC) of the full set of physiological signals compared to each physiological signal used independently, used in identical model pipelines, and evaluated on the same test splits. A full report of the pairwise ROC curve comparisons for statistically significant differences (using methodology as reported in DeLong et al.[51]) is given in Supplementary Table 2. Important to note, the model using all physiological

Table 1 | Metadata table

		Covid + (n = 295)	Covid - (n = 172)
Age	Percent Available	100.0%	97.09%
	Mean ± Std	42.45 ± 12.44	42.56 ± 15.44
	Median	42.0	42.0
	Range	(18.0, 76.0)	(19.0, 73.0)
Sex	Percent Available	100.0%	98.84%
	Male	190	98
	Female	105	72
Days Since Positive Test	Percent Available	100.0%	—
	Mean ± Std	4.48 ± 3.6	—
	Median	3.0	—
	Range	(0.0, 14.0)	—
Height	Percent Available	99.32%	95.35%
	Mean ± Std	1.64 ± 0.08	1.66 ± 0.07
	Median	1.65	1.7
	Range	(1.42, 1.83)	(1.45, 1.85)
Weight	Percent Available	99.32%	95.35%
	Mean ± Std	66.95 ± 8.89	63.59 ± 10.89
	Median	68.0	67.5
	Range	(39.0, 89.0)	(39.0, 85.0)
BMI	Percent Available	99.32%	95.35%
	Mean ± Std	25.0 ± 3.06	22.83 ± 2.78
	Median	24.69	23.8
	Range	(15.62, 34.24)	(15.62, 28.4)
Symptoms	Percent Available	100.0%	—
	Fever	73.22%	—
	Cough	48.14%	—
	Body Aches	42.03%	—
	Sore Throat	35.25%	—
	Headache	33.9%	—
	Fatigue	22.71%	—
	Nasal Congestion	15.59%	—
	Nausea/Vomiting	11.19%	—
	Loss Taste/Smell	9.83%	—
	Diarrhea	9.15%	—
	Rash	8.14%	—
	Difficulty Breathing	5.08%	—
	Loss Appetite	0.0%	—
Comorbidities	Percent Available	100.0%	95.35%
	None	263	142
	Hypertension	17	9
	Diabetes	12	9
	Asthma	3	4
	Joint pain	2	4
	Tuberculosis	2	0
	COPD	1	0
	Hypothyroidism	1	0
	OCD	1	0

Applicable metadata available for subjects used in model development and testing, including age (in years), sex, number of days since positive test, symptoms, and comorbidities. *COPD* Chronic obstructive pulmonary disease. *OCD* Obsessive-compulsive disorder.

Table 2 | Feature availability

Sensor modality	% data available (standard deviation)	Total count of features per modality
Heart Rate	76.2% (6.3)	57
Respiration	91.5% (1.6)	32
Gait	92.8% (0.1)	36
Lung Sounds	91.9% (1.1)	30
Cough	84.4% (0.0)	52
Temperature	88.1% (6.4)	12
SpO2	65.0% (23.1)	54

Percentage of features available from 467 samples and the count of how many features were extracted per sensor modality. The standard deviation represents the variability between percent of features available within each modality.

features has statistically significant higher ROC than models using any physiological signals alone.

Feature Importance

XGBoost feature importance and SHAP (SHapley Additive exPlanation) values were obtained to examine the feature importance during training and model classification, respectively (Fig. 2). XGBoost assigns higher feature importance to features which heavily contributed to the construction of model decision trees during training. SHAP values are calculated by considering all possible combinations of features and how each particular feature contributes to the prediction[52]. Visually, SHAP values can elucidate the relative value of features (i.e., high or low values) that the model considered important when classifying each sample as COVID or NON-COVID. In both cases, the ranking was obtained by calculating the absolute mean of each feature's importance value across the 100 validation tests. All sensor modalities had at least one feature considered important, with the top three most important features stemming from temperature, cough, and lung sounds. Notably, features from the photo plethysmography (PPG) data were found to not be critical to model performance, likely due to the inconsistent availability of this signal in the extracted data. Supplementary Table 3 shows the percentage of data available for the 467 samples used in each of the feature deemed as important by the model.

Discussion

This study was a first step in determining if physiological changes of cardiorespiratory infection can be detected by a commercial-grade wearable sensor using only a snapshot of data. Real-world deployment of this technology while detecting probable COVID-19 in India advances our team's preliminary research[50], demonstrating the feasibility and potential clinical implications of translating snapshot data into a rapid screening tool for cardiorespiratory illnesses. Our results demonstrated the primary model encompassing all sensor modalities and activities performed more accurately than any single sensor modality feature set used alone. Thus, multimodal sensor fusion across a variety of movement-based activities may be beneficial to observing broad physiological manifestations of cardiorespiratory illness, which in turn can be quantified and classified by advanced machine-learning algorithms. In this proof-of-concept implementation, we were able to use XGBoost to classify a diverse set of individuals with or without COVID-19 infection with an average F1-score of 0.80 (recall, precision = 80%) and an AUC of 0.81.

We calculated SHAP values to reveal the underlying significance of physiological features during model classification. With the exception of percent oxygen saturation, all sensing modalities appeared among the most important ranked features during both training and classification. The top three features were chest temperature range, the spread (IQR) of cough accelerometer signals, and the regularity of the lung sounds (entropy) in the frequency domain. Specifically, greater change in temperature, lower cough signal variability, and less regularity in the lung sound frequency domain

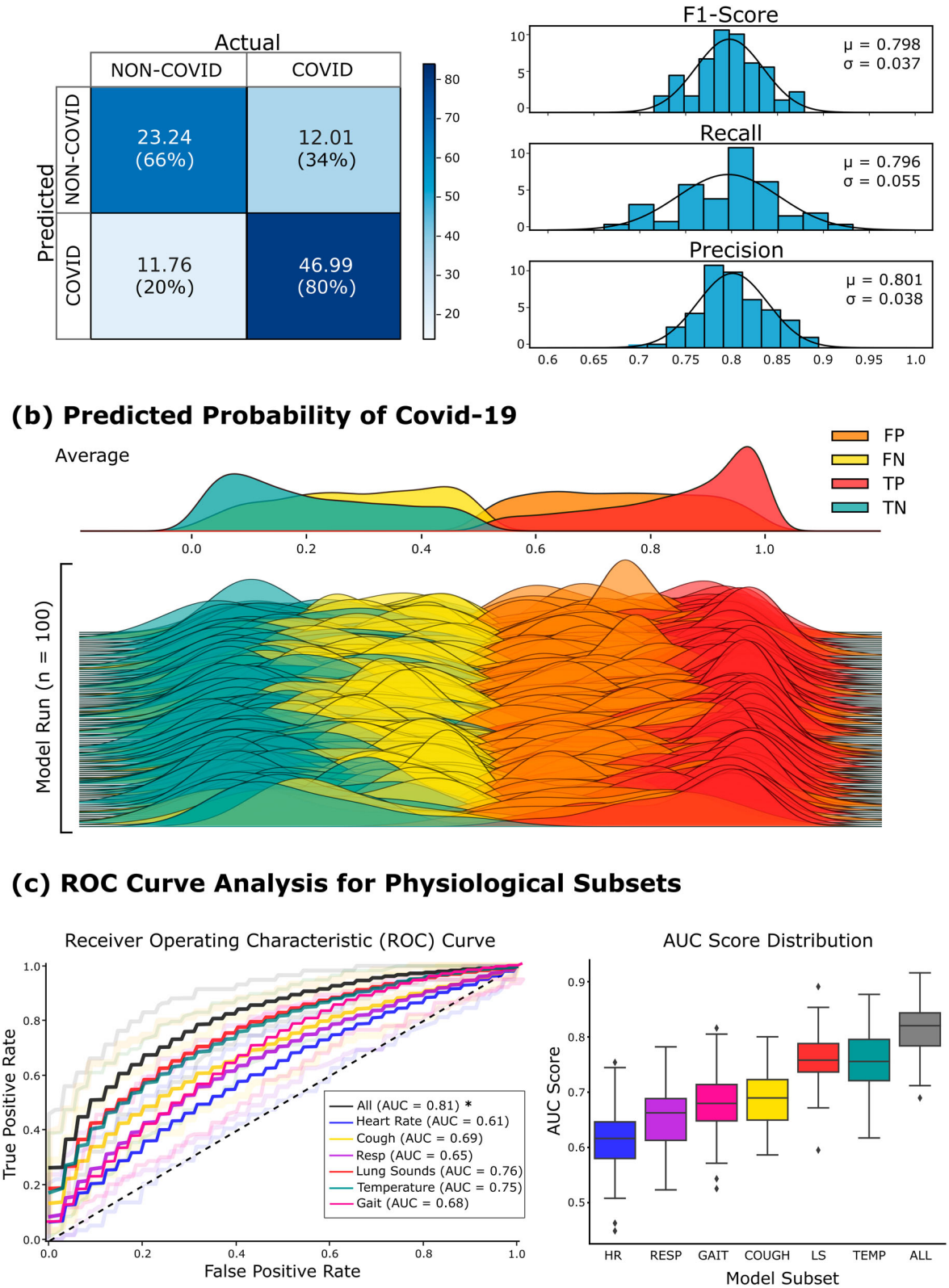

Fig. 1 | XGBoost to Classify COVID and NON-COVID. a Average confusion matrix over the 100 model iterations, as well as the mean and standard deviation of performance metrics (f1-score, recall, and precision). **b** Predicted probability confidence of COVID, where values near 1 correspond to high confidence of COVID classification, are color mapped to the model predictions; FP = False positive; FN = False negative; TP = True positive; FN = False negative. **c** ROC curves displayed for subsets of physiological signal features. The label "All" corresponds to the full model compromising all physiological sensor features. Faded lines represent the standard deviation of model performance. * indicates that this model was statistically significantly different from all other models using DeLong's method. Boxplot shows the interquartile range (IQR) of the AUC scores obtained from the cross validation, with whiskers extending to 1.5 times the IQR.

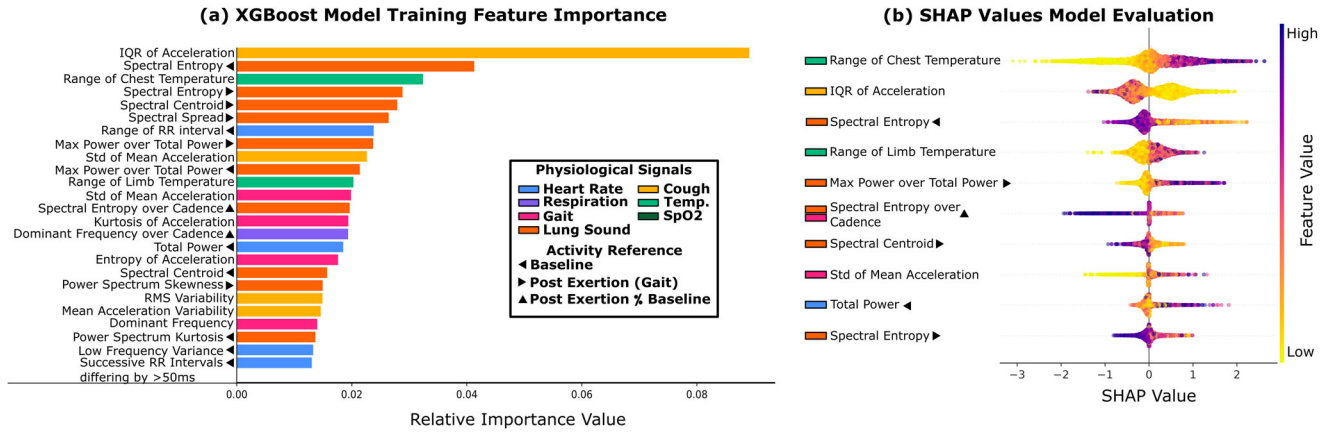

Fig. 2 | Feature Importance Separating COVID and NON-COVID. a Average feature importance across the 100 model runs, ranked in descending order based on the XGBoost decision tree model during training. Larger values on the x-axis correspond to a higher feature importance. **b** Top 10 features defined by SHAP values in descending order for model prediction. The SHAP values represent the marginal contribution of the feature for the model prediction. Larger SHAP values along the x-axis push the prediction towards a classification of the positive case (i.e., COVID). Smaller SHAP values rather push the prediction towards a classification of the negative case (i.e., NON-COVID). Each dot represents a single data sample from the 100 test sets, with the color corresponding to feature value (i.e., dark purple dots for the top feature corresponds to a high temperature range). In both **(a)** and **(b)**, features are color-coded to highlight the different sensor modalities and arrows represent the period in the activity segmentation of those features. Supplementary Table 3 shows the percentage of data availability for each of these features.

pushed the models towards a positive prediction of COVID. A possible reason for the trend identified in temperature is that infections can alter the natural fluctuations in temperature[53] that might occur during and after a short aerobic exercise[54]. Immune responses in infected individuals increase metabolic rate, making these responses energetically costly[55] and limiting the body's ability to regulate temperature efficiently during periods of acute illness. Additionally, Mason et al. (2022) reported the inclusion of temperature as a feature increasing their AUC score by 4.9% when using continuous monitoring of Oura rings for COVID detection[40]. It may also be understood why COVID-19 positive cases show less variability while coughing. In our dataset, nearly 50% of participants reported cough as a symptom. Thus, in this group, it is more likely that natural cough was triggered by central pattern generators which elicit rhythmic vibrations, versus in individuals forcing a cough[56]. Regarding the power spectral entropy of lung sounds, it can be noted that lower spectral entropy is associated with a more irregular distribution of power across observed frequencies, as opposed to higher values which are characteristic of a flat or consistent distribution of power[57]. High-frequency lung sound activity was observed in individuals with respiratory illness as compared to healthy individuals with less high frequency respiratory content (thus resulting in a flat signal), which is consistent with literature on this topic[58]. In all SHAP value cases, results should be cautiously inferred as the trends observed to support classification does not necessarily mean the prediction was correct.

To further evaluate the effectiveness of the model, we re-ran it using only the top 10 features identified by SHAP values, which yielded a superior AUC (0.85) and a similar F1 score of 0.81 compared to the full feature set model. This underscores the importance of feature selection in improving model efficacy and suggests that a well-chosen set of features can robustly classify respiratory illnesses, potentially simplifying the model without sacrificing accuracy. This information provides valuable insights for future studies, especially in exploring these features across a more diverse population of cardiorespiratory illnesses.

As COVID-19 continues to evolve, rapid testing methods are necessary to reduce transmission rates, especially in institutions with high-risk individuals and/or high transmission rates that require repeatable and proactive monitoring. Previous research has shown the practicality of using wearable sensors in this context. Radin et al. (2020) tracked infection via consumer wearable technology in real-time at the population level while predicting future cases[59]. Work from Mishra et al. (2020) as well as from Bogu and Snyder (2021) similarly demonstrated the use of smartwatch data to detect COVID-19 before the onset of symptoms[60,61]. In both instances, models relied on continuous, individual level changes in heart rate and step count. While longitudinal data is useful for tracking alterations in physiology, it often requires several days of recorded baseline activity on a personal device, which can be limiting in certain applications. Thus, we designed our activity sequence to capture a quick snapshot of physiological data during a brief baseline activity (i.e., seated rest) and after a period of exertion (i.e., fast walking) to elicit observable physiological change and extract innovative features to use for model training. Our study design's novelty can be observed when compared to Walter et al. (2024) which achieved a sensitivity of 0.47 using continuous data and the given features of the ANNE^{TM} One sensor system for early detection of COVID in home environments[62]. Our relatively greater performance emphasizes the potential for rapid, snapshot testing and engineered, physiologically relevant features. Similar to the comparisons made in longitudinal data, we normalized several features by individual-level baseline measures to account for relative changes in signal characteristics. By doing so, we can illuminate some level of individuality within a general use model. While we evaluated participants at a single time point, the same sensor approach could be applied to track cardiorespiratory illness at scale over time, simply by repeating the sequence at the desired number of time points per individual to render new model outcomes and insights.

When using this technology in practice, it is important to consider the consequential tradeoff between prioritizing sensitivity or specificity in testing for the presence of COVID-19. As sensitivity and specificity are inversely proportional[63], one must evaluate the consequences of false positives compared to those of false negatives when designing or choosing tests. These considerations may change as the disease spreads over time. At the beginning of the COVID-19 outbreak, testing methods with a higher sensitivity at the cost of lower specificity were more desirable to encourage isolation and prevent the spread of disease[64,65]. False negatives were a greater risk from a public health perspective, whereas false positives would merely lead to unnecessary testing and quarantine[66]. As COVID-19 reaches an endemic stage, false positive results have become a proportionally greater concern[67]. The costs of quarantining, including further isolation, mental anguish, and accidental viral exposure have become more significant. False positives amongst healthcare workers especially create additional burden on already strained healthcare systems[68]. Thus, it has become acceptable to utilize testing tools that sacrifice sensitivity but are less invasive and encourage testing. In the context of wearable technology, noninvasive screening tools could promote even greater testing rates and reduce transmission. For the scope of this manuscript, we treated false positives and false

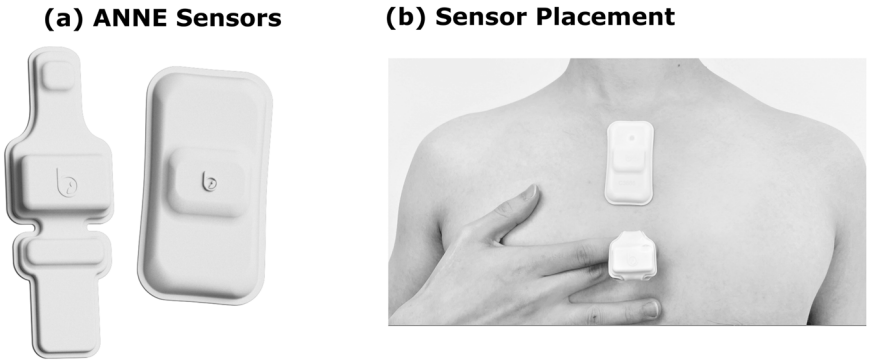

Fig. 3 | ANNE™ One Sensor System and Application. The ANNE™ One Sensor System is composed of two individually packaged wireless sensors as illustrated in **(a)**. The anatomical placement of each sensor, the limb sensor (left) and chest sensor (right), is demonstrated in **(b)**. Soft, flexible sensors are adhered to the skin with hydrogel adhesives. The form factor of the limb sensor reinforces the photodiode and LED to the index finger.

negatives equally using the balanced F1-score. In future iterations of this technology, the practical consequences of each should be considered. Perhaps the algorithm could even be re-trained to favor one over the other, depending upon the context of sensor use and the user preference.

Although the methodological approach is promising, there are several limitations in diagnostic reliability. First, the current iteration of this methodology does not directly measure viral material and cannot offer a definitive diagnosis. Second, a major limitation to the study is that we cannot confirm the level to which our model can separate COVID-19 from similar cardiorespiratory illnesses, due to a lack of metadata available to the clinical team and challenges in recruiting individuals with other symptoms or cardiorespiratory illnesses. Future work should evaluate large datasets differentiating between other types of cardiorespiratory illnesses, such as COVID-19 versus influenza, and should include presentations of illness that are more representative of what is observed in the community (i.e., asymptomatic infected individuals).

Furthermore, despite balancing positive and negative cases in the training data, the model still had slight preference in classifying samples as COVID-19 positive. This may be due to the internal cross-validation method retaining the original ratio of more positive cases in the test group. Thus, the internal hyperparameter values selected may have been influenced by classifying correctly for this ratio. This method is standard for practical applications of machine-learned models, which learn to classify positive cases with respect to the true prevalence in the community, however our dataset was not representative of the true community prevalence.

COVID-19 is currently transitioning into endemic stability, but despite this shift, widespread infection and associated complications persist. Therefore, researchers must continue developing and enhancing environmentally friendly, large scale screening tools, including wearable sensors. The expansion and adoption of this technology remains crucial, especially in light of warnings that pandemics are likely to occur more frequently in the future[69,70]. Our study demonstrated the potential of wearable sensors, particularly when using snapshot sensor data. If appropriately improved upon, this technology will facilitate early screening and community monitoring in inevitable future pandemics. In this proof-of-concept study, we trained on data collected during different viral dominance periods, however it is unknown whether this was an advantage to the model or perhaps a limiting factor in the model's ability to learn various strains. Transfer learning may be promising for future iterations of this work.

In summary, we have demonstrated the potential use of wearable sensing technology to screen for and monitor cardiorespiratory illnesses, specifically COVID-19. Current conventional testing methods have drawbacks concerning accessibility, implementation, and reliance on single-use materials. The alternative methodology proposed here is easily implementable for use in remote and under-resourced communities and could make early diagnostic testing more accessible. Rapid, reusable, and easy-to-scale diagnostic technologies may additionally lower environmental impact. As COVID-19 continues to evolve and global pandemics become increasingly common, innovative alternatives for testing are necessary to limit the spread of disease around the world. We see this work as providing evidence for the feasibility of wearable sensor systems in screening for and monitoring cardiorespiratory illnesses, especially as this form of technology continues to advance. Furthermore, future iterations of this work should attempt to include individuals infected with other cardiorespiratory disease in accordance with the prevalence of the infection of interest to further demonstrate the potential of this technology.

Methods

Sensing device

The ANNE™ One sensor system (Sibel Health; Niles, IL, USA) is a United States Food and Drug Administration (FDA) cleared, clinical-grade sensing platform capable of monitoring both full-body motion signals and physiological measures[30]. The system consists of two soft, flexible sensors: one anatomically positioned at the chest to measure tri-axial acceleration, electrocardiography (ECG), heart rate, respiratory rate, and proximal skin temperature, and the other positioned on the finger to measure photo plethysmography (PPG) for SpO2 and distal skin temperature (Fig. 3). The sensors are time-synchronized and connect to a tablet via Bluetooth for guided use and data storage. Acceleration was collected at 200 Hz in the direction of the x- and y- axes and 1600 Hz in the direction of the z-axis (sagittal plane). ECG, PPG, and skin temperature were recorded at 512 Hz, 128 Hz, and 1 Hz, respectively. The resulting sensor data was pre-processed with respect to the physiological signal of interest (see Feature Generation below).

Participants

This research was made possible in collaboration between Shirley Ryan AbilityLab (USA) and Bionic Yantra (India) through the United States-India Science and Technology Endowment Fund (USISTEF), supporting research for innovative solutions to address challenges posed by COVID-19. According to the FIND SARS-CoV-2 test tracker, India has been amongst the list of countries short of reaching the ACT-A testing target[25]. Thus, this collaboration was realized as an opportunity to implement novel testing technology in a region which may practically benefit from its use.

Characteristics of participants can be found in Table 1 (see Results). Individuals between 18 and 85 years of age were recruited from a sample of convenience across heterogeneous test sites as either experiencing (COVID) or not experiencing (NON-COVID) COVID-19 infection, confirmed via a RAT and/or RT-qPCR test. The COVID cohort consisted of COVID-19 positive individuals at COVID-19 inpatient or outpatient facilities at the time of data recording. The NON-COVID cohort consisted of members of the community who did not test positive for COVID-19. Individuals with implanted pacemakers or defibrillators and individuals pregnant at the time of consent and study involvement were excluded from this study.

All participants provided written and/or verbal consent prior to their participation in this research study. The study was approved by the Institutional Ethics Committee (IEC) of the Indian Institute of Public Health-Delhi (IIPH-D) and the S2J Independent Ethics Committee (S2J IEC). The

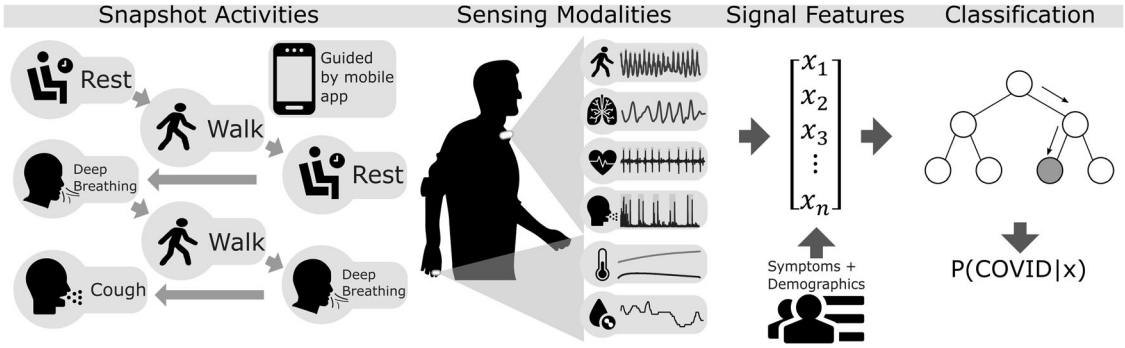

Fig. 4 | Overview of Rapid Snapshot Approach to Classify Probability of COVID-19. A short and structured sequence of movement-based activities are performed while wearing the ANNE™ ONE sensor system. Periods of rest and deep breathing are captured prior to and after gait to observe both baseline measurements and elicited responses to increased exertion. Raw sensor signals are translated into physiological measurements, including gait, respiratory rate, heart rate, cough, core, and peripheral temperature, and SPO2. Time series and frequency domain features of each sensing modality are extracted and transferred as input to a machine learning classifier to predict probability of COVID-19 infection.

trial was conducted as per the Indian Council of Medical Research Guidelines for Biomedical Research on Human subjects, including standards of the Declaration of Helsinki (Brazil, 2013) and other applicable guidelines. This study is a registered clinical trial with the NIH initially released April 15, 2022 under the registry name: "International Validation of Wearable Sensor to Monitor COVID-19 Like Signs and Symptoms" and identifier: NCT05334680.

Data collection

Data was collected in an uncontrolled, real-world clinical environment across multiple sites in India. The primary clinical research team trained external clinical staff on the data collection protocol including sensor application, the activity sequence, and data transfer. Initial data quality checks and routine site visits were conducted to verify protocol adherence and data integrity. Institutionally required personal protective equipment (PPE) was worn during all study procedures. Participants were asked to remain masked throughout the session. Sensor locations on the skin were cleaned with a medical wipe and sanitized before and after every sensor application. Sensors were systematically rotated to allow a complete sanitation cycle and were adhered to the skin as shown in Fig. 3. Clinical staff placed one sensor on the suprasternal notch and one on the tip of the index finger. Sensors placed on the chest were adhered using medical adhesive.

An overview of the data collection protocol can be visualized in Fig. 4. The activity sequence was designed to leverage changes in physiology from baseline measurements to measurements during recovery from exertion, emulating stress or walking tests that are commonly used to evaluate cardiorespiratory function[71,72]. Thus, the activity sequence aims to 1) address the capability of a rapid protocol to elicit measurable physiological change that can be picked up by a commercial-grade wearable sensor, and 2) determine if the physiological features captured are sensitive enough to detect differences in individuals with and without cardiorespiratory illness.

Under supervision by clinical staff, all participants began seated and were asked to refrain from talking during a short and simple sequence of standardized activities. Normal breathing was collected while seated before and immediately after 30 s of walking, in which participants were instructed to walk at a fast yet safe pace. Participants were then instructed to take five voluntary deep breaths while seated, followed by another 30 s period of fast yet safe walking, and then five voluntary deep breaths while seated. Participants then performed five consecutive voluntary coughs while seated. The entire activity sequence is approximately two minutes long. Between each activity, three consecutive hand-taps were performed onto the chest sensor as a reference for activity segmentation. Each participant performed one trial of the activity sequence.

Feature generation

Real time bio-signals were streamed from the sensors via Bluetooth to the ANNE™ Sync application. Data from the sensor system was directly uploaded to a cloud-based server for use in the feature extraction pipeline. Oxygen saturation index (SpO2), central body temperature, and peripheral body temperature were directly used as measured by the ANNE™ system's proprietary software (Sibel Health, Inc., Niles, IL, USA), including only measurements with the given signal quality index above 0. To access increased signal information from other data streams, all other physiological measures were custom derived from the raw sensor stream using methods in literature, as detailed in the sections below. The acceleration signal was used to segment the sequence into the activities described above using a custom code that required visual inspection and confirmation of starting and end points of each activity. The time points at which various physiological signals were analyzed and extracted relative to the movement-based activities can be visualized in Fig. 5. The design of the activity sequence permits accounting for physiological differences before and after exertion, thus many of the features were normalized to measures at baseline (i.e., resting) to capture this. A summary of the features input to the model across all sensing modalities can be found in Supplementary Table 4.

Estimation Of R-R Intervals

R-R intervals were extracted from the raw ECG time series signal using the Pan-Tompkins algorithm, which applies a series of filters to reduce noise and accentuate the peak of the R wave in the QRS complex[73]. In cases of inadequate detection by the Pan-Tompkins algorithm alone and after visual confirmation, a maximal overlap discrete wavelet transform (symlet wavelet filter, MATLAB R2022B) was applied to confirm the available output of the Pan-Tompkins algorithm and impute missing time points. In the case this alternative could not detect reliable QRS complexes, the heart rate signal was dropped for that sample. The final array of detected R peaks was passed into the hrv-analysis Python module to extract features related to heart rate variability[74]. Certain features that are not recommended for durations at or below 30 seconds[75] were ignored from the full set of hrv-analysis features. To additionally measure the change in heart rate, the post-exertion measurement was divided by the baseline measurement, thus normalizing the change amongst subjects.

Respiration Rate

To measure respiration rates, the angular motion due to breathing was reconstructed by tracking the rotation of the gravity vector in the accelerometer signal[76]. Tri-axial acceleration signals were down sampled (200 Hz) and filtered (2nd order Butterworth low-pass filter at 1 Hz). The

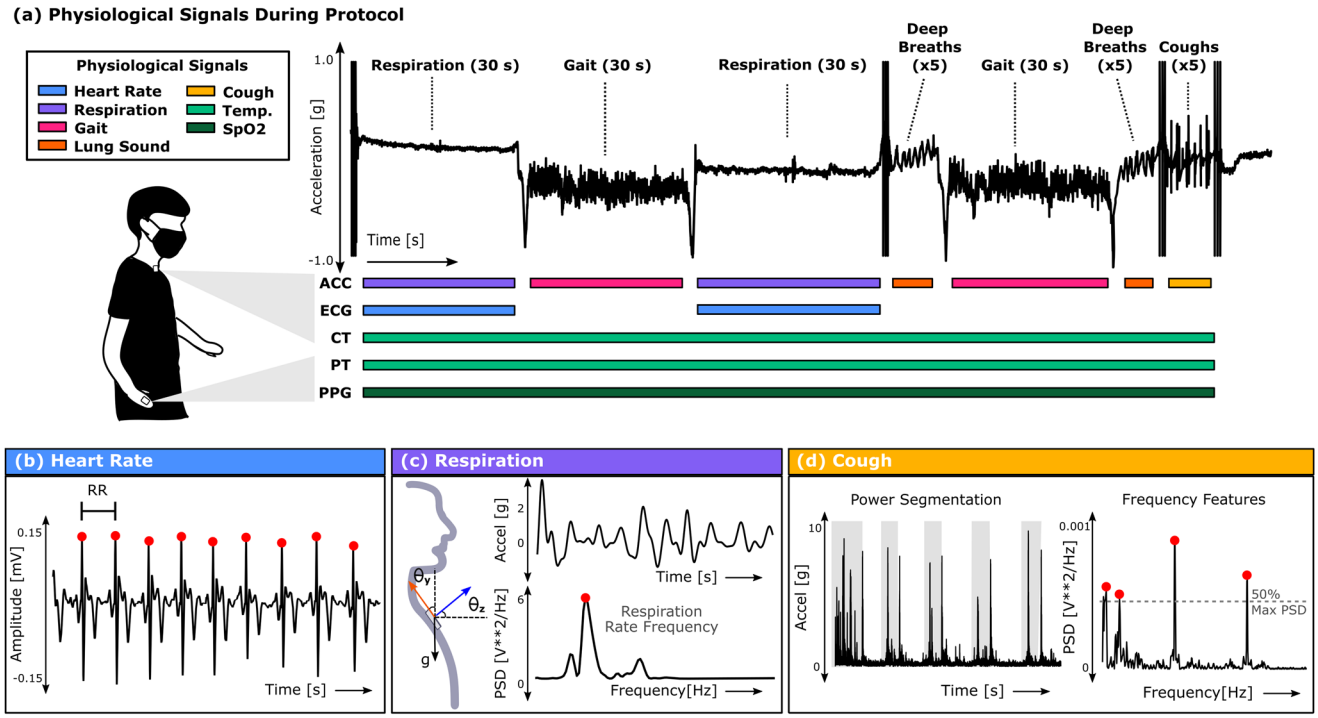

Fig. 5 | Extraction of Physiological Signals and Features from Rapid Snapshot Protocol. a A complete sequence of the rapid snapshot activity protocol is shown for the time series acceleration signal (in black). Directly below, colored bars representing physiological signals are vertically aligned in the time series domain for time points of various activity-wise extraction and are horizontally stacked with respect to the sensing modality from which they were extracted; ACC acceleration, ECG electrocardiography; CT central temperature, PT peripheral temperature, PPG photoplethysmography. **b**–**d** present supporting visual diagrams of the signal processing and/or feature extraction domains for heart rate, respiration, and cough, respectively. (**c**) also shows the chest sensor placement and the gravity vector with respect to the participant. Feature extraction processes for all physiological signals are detailed in their respective sections.

normalized vector of acceleration at each time point, a_t, was used to calculate the axis of rotation, r_t, between two consecutive measurements:

$$r_t = a_t \times a_{t-1} \tag{1}$$

Each axis of rotation was weighted using a Hamming window function. The resulting rotation angle, φ_t, calculated as

$$\varphi_t = \sin^{-1}((a_t \times r_t) \times a_t) \tag{2}$$

was filtered (8th-order Butterworth band-pass filter at 0.1 and 0.8 Hz) and differentiated to get the angular rate. The power spectral density of the angular rate was estimated using Welch's method, and the respiration rate in BPM (breaths per minute) was taken as the dominant signal frequency (see Fig. 5C). Frequency values were evaluated within physiologically relevant bounds; 10 to 32 BPM for baseline respiration[77,78] and 10 to 60 BPM for respiration after walking[79]. Frequency values with a power of at least 50% of the dominant signal frequency were included in calculations for the maximum signal power, average frequency, and sum of the signal power. To measure the difference in respiration rate before and after walking, we calculated the ratio of values—dividing the post-exertion measurement by the baseline measurement.

Cough signal properties

Cough accelerometer signals contain high-frequency oscillating peaks due to varying motions of the chest during a cough: breathing in, compression of muscles, and quick expulsion of air[80]. Each axis of the tri-axial accelerometer data was interpolated to the sampling frequency of the z-axis (1600 Hz) and filtered (5th-order, high-pass Butterworth filter at 40 Hz). The normalized acceleration vector at each time point was calculated. Welch's method was applied to the entire signal duration to segment individual coughs using a custom sliding window technique (0.2 s with a 50% overlap) based on a threshold relative to the mean signal power (see Fig. 5D). To compare cough data between subjects, cough signals were normalized via a linear scaling method based on the accelerometer signal range during seated baseline respiration. Features were extracted from both the time and frequency domains of each segmented cough signal. The mean and standard deviation of feature values across all detected coughs for a given individual (expected count of 5) were used in the final model.

Estimating lung sounds

High frequency lung sounds (i.e., crackles, wheezes) often captured via lung auscultation can provide important digital biomarkers of respiratory illness and severity[81]. Vibratory movements of the lung have been correlated to diagnosis of COVID-19 infection[82,83]. We therefore used the upper observable range of the frequency domain for our sensing device to characterize high frequency content as observable during the deep breathing activity. Accelerometer data was filtered (8th-order, high-pass Butterworth filter at 100 Hz) and the power spectral density of frequency was calculated via Welch's method. Information including the dominant frequency, statistical moments of the power spectrum, and the number of detected frequency peaks (at least 50% of the dominant frequency power) were included as features to the model.

Estimation Of Walking Cadence

To estimate the subject walking cadence, the L2-norm of the acceleration was calculated and the fast Fourier transform was applied to find the dominant frequency between a low and high frequency bound, 0.7 Hz and 3.5 Hz, respectively. The dominant frequency was considered to be the stepping frequency, i.e., cadence. Statistical moments and entropy of the power density spectrum were extracted as input to the model. To quantify post-gait activity relative to the exertion level of gait activity, mean heart rate, frequency and power of the respiration rate, and lung sound features during deep breathing were normalized by the dominant gait frequency. Gait

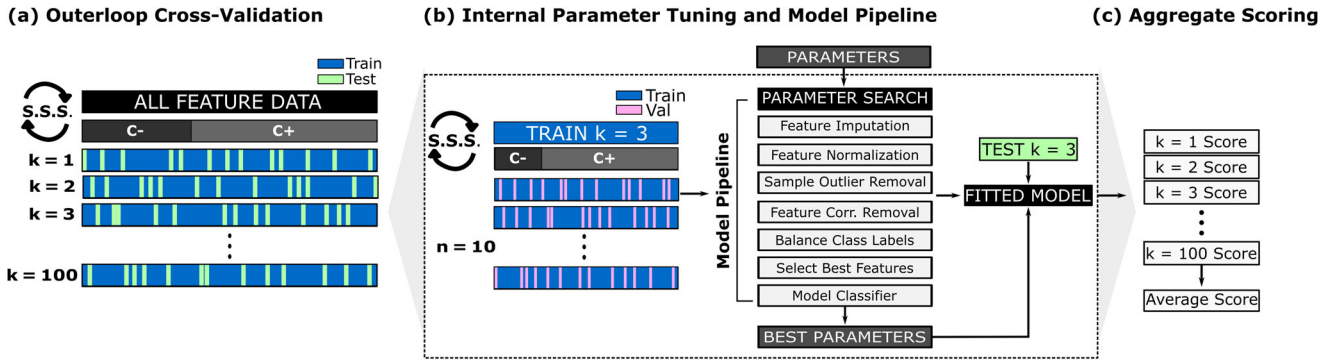

Fig. 6 | Machine Learning Cross-Validation Pipeline. This figure provides a visual overview of the model pipeline architecture. **a** The entire feature dataset was randomly split into train-test sets (k = 1:100, 80–20 split) using a stratified shuffle split method (S.S.S.). The original ratio of COVID (C +) to NON-COVID (C-) samples is maintained within each train and test set. **b** An example of the internal modeling structure occurring for each outer loop split. Training data from a single split (i.e., k = 3) was input to an internal hyperparameter tuning process, also employing S.S.S. cross-validation for $n = 10$ folds. Preprocessing steps, feature selection, and the modeling classifier steps are shown vertically in gray. Note that in the model pipeline, the step "Balance Class Labels" handles imbalance in the training data using Synthetic Minority Over-sampling Technique (SMOTE). The best parameters and training data were used to fit a retrained model operating on the same modeling steps. **c** The final metrics are reported as an average over all 100 outer loop splits.

features were also compared between the first and second walking pass to determine any effects of fatigue on the walking dynamics.

Model architecture and evaluation

All sensor features and select metadata information (age, sex) were provided as input to a machine-learning pipeline in Python. The pipeline architecture was designed to maximally utilize all data samples while also implementing external validation via split-sample validation. An overview of the pipeline is visually represented in Fig. 6. Training, validation, and testing splits prevented sample-wise data leakage across trained and evaluated partitions of the dataset and evaluated robustness of the model to generalize to unseen data. Additionally, the robustness of the model was validated with an additional cross-validation method described in Supplementary Table 5.

Stratified Shuffle Split technique was used to split the data into training and testing subsets (k = 100 folds) and preserve the ratio of COVID to NON-COVID in our sample[84]. Training data was further partitioned into training and validation subsets within a nested loop cross-validation procedure (Stratified Shuffle Split method, k = 10 folds) with the purpose of tuning hyperparameters. Optimal values of the internal model parameters were determined within a randomized search process[84,85], evaluating each internal fold on 100 randomly selected combinations of model hyperparameters. Within this nested optimization and training procedure, the feature data were provided as input to a series of standardized pipeline processes:

- Imputation of missing feature values using k-Nearest Neighbors technique[84,86] (reasons for missing data mentioned in Results and percentage of feature availability per sensor modality is presented in Table 2).
- Normalization of features using RobustScaler method, robust to outliers[84].
- Removal of outliers via Isolation Forest algorithm[84,87].
- Removal of highly correlated features (method = 'Pearson', threshold = 0.95)[88] to reduce feature dimensionality.
- Handling of imbalanced training data via SMOTE, known as Synthetic Minority Over-sampling Technique[89,90].
- Feature selection via SelectKBest ($n = 50$) which comprehensively compares feature values against the target variable[84] to reduce feature dimensionality.

As the final step in the pipeline, we used an extreme gradient boosting decision tree algorithm (XGBoost) in a binary supervised learning task to classify COVID from NON-COVID cases[91]. We defined a test positivity cut-off of 0.5 for the model's output probability, as this is the standard default in binary classification tasks. Model parameters specific to gradient tree algorithms known to affect model fit were lent to the hyperparameter optimization process: maximum tree depth, minimum child weight, learning rate, and gamma (Supplementary Table 1). The hyperparameters selected were those which resulted in the highest average harmonic F1-score across the ten-fold cross-validation. The best performing model from the internal cross-validation procedure was fit to the corresponding training data using the selected hyperparameters and evaluated upon the held-out test set to observe model generalization errors. For each of the train-test split iterations, the model was independently evaluated on metrics of recall, precision, F1-score, and Area under the Receiver Operating Characteristic Curve (AUC). The distribution of performance scores, optimized hyperparameter values, and feature importance across the 100 train-test splits are reported.

Data availability

Sensor data that support the findings of this study may be made available to an investigator upon request for academic, research, and non-commercial use.

Code availability

The code used to process and analyze the findings of this publication may be made available to an investigator upon request for academic, research, and non-commercial use.

Received: 24 January 2024; Accepted: 7 October 2024;
Published online: 19 October 2024

References

1. World Health Organization. *Situation Report - 1*. (2020).
2. Biancolella, M., Colona, V.L., Mehrian-Shai, R. et al. COVID-19 2022 update: transition of the pandemic to the endemic phase. *Hum. Genet*. **16**, 19 (2022).
3. Spencer, S. E. W. & Luban, J. *Is COVID-19 Reaching the Endemic Stage? UMass Chan Virologist Jeremy Luban Weighs In*. https://www.umassmed.edu/news/news-archives/2022/08/is-covid-19-reaching-the-endemic-stage-umass-chan-virologist-jeremy-luban-weighs-in/ (2022).
4. Locklear, M. *For COVID-19, Endemic Stage Could Be Two Years Away*. https://medicine.yale.edu/profile/caroline_zeiss/ (2022).
5. Are, E. B., Song, Y., Stockdale, J. E., Tupper, P. & Colijn, C. COVID-19 endgame: From pandemic to endemic? Vaccination, reopening and

evolution in low- and high-vaccinated populations. *J. Theor. Biol.* **559**, 111368 (2023).
6. Katzourakis, A. COVID-19: endemic doesn't mean harmless. *Nature* **601**, 485 (2022).
7. Lin, D.-Y. et al. Effectiveness of Covid-19 Vaccines over a 9-Month Period in North Carolina. *N. Engl. J. Med.* **386**, 933–941 (2022).
8. Andrews, N. et al. Covid-19 Vaccine Effectiveness against the Omicron (B.1.1.529) Variant. *N. Engl. J. Med.* **386**, 1532–1546 (2022).
9. Centers for Disease Control and Prevention. *CDC Announces Negative COVID-19 Test Requirement from Air Passengers Entering the United States from the People's Republic of China*. (2022).
10. Wang, Q. et al. Alarming antibody evasion properties of rising SARS-CoV-2 BQ and XBB subvariants. *Cell* https://doi.org/10.1016/j.cell.2022.12.018 (2023).
11. Menni, C. et al. Symptom prevalence, duration, and risk of hospital admission in individuals infected with SARS-CoV-2 during periods of omicron and delta variant dominance: a prospective observational study from the ZOE COVID Study. *Lancet* **399**, 1618–1624 (2022).
12. Ma, Y. et al. Long-term consequences of COVID-19 at 6 months and above: a systematic review and meta-analysis. *Int J. Environ. Res Public Health* **19**, 6865 (2022).
13. Zhou, Y. & OLeary, T. J. Relative sensitivity of anterior nares and nasopharyngeal swabs for initial detection of SARS-CoV-2 in ambulatory patients: Rapid review and meta-Analysis. *PloS one* **16**, e0254559 (2021).
14. Filchakova, O. et al. Review of COVID-19 testing and diagnostic methods. *Talanta* **244**, 123409 (2022).
15. Dutta, D. et al. COVID-19 Diagnosis: A Comprehensive Review of the RT-qPCR Method for Detection of SARS-CoV-2 https://doi.org/10.3390/diagnostics12061503. (2022).
16. Sharma, S., Shrivastava, S., Kausley, S. B., Rai, B. & Pandit, A. B. Coronavirus: a comparative analysis of detection technologies in the wake of emerging variants. *Infection* https://doi.org/10.1007/s15010-022-01819-6 (2022).
17. Reynard, C. et al. COVID-19 rapid diagnostics: practice review. *Emerg. Med J.* **39**, 70–76 (2022).
18. Mercer, T. R. & Salit, M. Testing at scale during the COVID-19 pandemic. *Nat. Rev. Genet* **22**, 415–426 (2021).
19. Ghasemi, S., Harmooshi, N. N. & Rahim, F. Diagnostic utility of antigen detection rapid diagnostic tests for Covid-19: a systematic review and meta-analysis. *Diagn. Pathol.* **17**, 36 (2022).
20. World Health Organization. *Antigen-Detection in the Diagnosis of SARS-CoV-2 Infection*. (2021).
21. Khandker, S. S., Hashim, N. H. H. N., Deris, Z. Z., Shueb, R. H. & Islam, M. A. Diagnostic accuracy of rapid antigen test kits for detecting SARS-CoV-2: A systematic review and meta-analysis of 17,171 suspected COVID-19 patients. *J. Clin. Med.* **10**, 3493 (2021).
22. Wells, C. R. et al. Comparative analyses of eighteen rapid antigen tests and RT-PCR for COVID-19 quarantine and surveillance-based isolation. *Commun. Med.* **2**, 84 (2022).
23. Cousins, S. Bangladesh's COVID-19 testing criticised. *Lancet (London, England)* **396**, 591 (2020).
24. Fetzer, T. *Measuring the Epidemiological Impact of a False Negative: Evidence from a Natural Experiment*, (2021).
25. FIND. SARS-COV-2 TEST TRACKER. *FIND Diagnosis for all* https://www.finddx.org/covid-19/test-tracker/ (2022).
26. World Health Organization. ACTaccelerator Access to COVID-19 Tools. https://www.act-a.org/diagnostics.
27. Peng, Y., Wu, P., Schartup, A. T. & Zhang, Y. Plastic waste release caused by COVID-19 and its fate in the global ocean. *Proc. Natl. Acad. Sci.USA* **118**, e2111530118 (2021).
28. Celis, J. E. et al. Plastic residues produced with confirmatory testing for COVID-19: Classification, quantification, fate, and impacts on human health. *Sci. Total Environ.* **760**, 144167 (2021).
29. Aragaw, T. A. & Mekonnen, B. A. Understanding disposable plastics effects generated from the PCR testing labs during the COVID-19 pandemic. *J. Hazard. Mater. Adv.* **7**, 100126 (2022).
30. Lee, K. et al. Mechano-acoustic sensing of physiological processes and body motions via a soft wireless device placed at the suprasternal notch. *Nat. Biomed. Eng.* **4**, 148–158 (2020).
31. Vyas, R. & Doddabasappla, K. FFT spectrum spread with machine learning (ML) analysis of triaxial acceleration from shirt pocket and torso for sensing coughs while walking. *IEEE Sens Lett.* **6**, 1–4 (2022).
32. Doddabasappla, K. & Vyas, R. Spectral summation with machine learning analysis of tri-axial acceleration from multiple wearable points on human body for better cough detection. *IEEE Sens. Lett.* **5**, 1–4 (2021).
33. Chen, P.-W. et al. Sleep monitoring during acute stroke rehabilitation: toward automated measurement using multimodal wireless sensors. *Sensors* **22**, 6190 (2022).
34. Ferguson, T. et al. Effectiveness of wearable activity trackers to increase physical activity and improve health: a systematic review of systematic reviews and meta-analyses. *Lancet Digit Health* **4**, e615–e626 (2022).
35. Shei, R.-J., Holder, I. G., Oumsang, A. S., Paris, B. A. & Paris, H. L. Wearable activity trackers–advanced technology or advanced marketing? *Eur. J. Appl. Physiol.* **122**, 1975–1990 (2022).
36. Iqbal, S. M. A., Mahgoub, I., Du, E., Leavitt, M. A. & Asghar, W. Advances in healthcare wearable devices. *npj Flex. Electron.* **5**, 9 (2021).
37. Gadaleta, M. et al. Passive detection of COVID-19 with wearable sensors and explainable machine learning algorithms. *NPJ Digit Med* **4**, 166 (2021).
38. Shandhi, M. M. H. et al. A method for intelligent allocation of diagnostic testing by leveraging data from commercial wearable devices: a case study on COVID-19. *NPJ Digit Med.* **5**, 130 (2022).
39. Mitratza, M. et al. The performance of wearable sensors in the detection of SARS-CoV-2 infection: a systematic review. *The Lancet. Digital health* **4**, e370–e383 (2022).
40. Mason, A. E. et al. Detection of COVID-19 using multimodal data from a wearable device: results from the first TemPredict Study. *Sci. Rep.* **12**, 3463 (2022).
41. Phipps, J. et al. Early adverse physiological event detection using commercial wearables: challenges and opportunities. *NPJ Digit Med.* **7**, 136 (2024).
42. Miller, D. J. et al. Analyzing changes in respiratory rate to predict the risk of COVID-19 infection. https://doi.org/10.1371/journal.pone.0243693 (2020).
43. Natarajan, A., Su, H.-W. & Heneghan, C. Assessment of physiological signs associated with COVID-19 measured using wearable devices https://doi.org/10.1038/s41746-020-00363-7. (2020).
44. Wu, R. et al. Feasibility of using a smartwatch to intensively monitor patients with chronic obstructive pulmonary disease: Prospective cohort study. *JMIR Mhealth Uhealth* **6**, e10046 (2018).
45. Jiang, W. et al. A Wearable Tele-Health System towards Monitoring COVID-19 and Chronic Diseases. *IEEE Rev. Biomed. Eng.* **15**, 61–84 (2022).
46. Holko, M. et al. Wearable fitness tracker use in federally qualified health center patients: strategies to improve the health of all of us using digital health devices. *NPJ Digit. Med.* **5**, 53 (2022).
47. Redmond, S. J. et al. What does big data mean for wearable sensor systems? Contribution of the IMIA Wearable Sensors in Healthcare WG. *Yearb. Med. Inf.* **9**, 135–142 (2014).
48. Cho, S., Ensari, I., Weng, C., Kahn, M. G. & Natarajan, K. Factors affecting the quality of person-generated wearable device data and associated challenges: Rapid systematic review. *JMIR mHealth and uHealth* **9**, e20738 (2021).
49. Faust, L. et al. Exploring compliance: Observations from a large scale fitbit study. in *Proceedings - 2017 2nd International Workshop on*

Social Sensing, SocialSens 2017 (part of CPS Week) 55–60. https://doi.org/10.1145/3055601.3055608 (Association for Computing Machinery, Inc, New York, New York, USA, 2017).
50. Lonini, L. et al. Rapid screening of physiological changes associated with COVID-19 using soft-wearables and structured activities: a pilot study. *IEEE J. Transl. Eng. Health Med.* **9**, 4900311 (2021).
51. DeLong, E., DeLong, D. & Clarke-Pearson, D. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* **44**, 837–845 (1988).
52. Lundberg, S. & Lee, S. A Unified Approach to Interpreting Model Predictions. *Cornell University* (2017).
53. Gassen, J., Nowak, T. J., Henderson, A. D. & Muehlenbein, M. P. Dynamics of temperature change during experimental respiratory virus challenge: Relationships with symptoms, stress hormones, and inflammation. *Brain Behav. Immun.* **99**, 157–165 (2022).
54. Neves, E. B. et al. Different responses of the skin temperature to physical exercise: Systematic review. in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* vols 2015-November (2015).
55. Muehlenbein, M. P., Hirschtick, J. L., Bonner, J. Z. & Swartz, A. M. Toward quantifying the usage costs of human immunity: Altered metabolic rates and hormone levels during acute immune activation in men. *Am. J. Hum. Biol.* **22**, 546–556 (2010).
56. Haji, A., Kimura, S. & Ohi, Y. A model of the central regulatory system for cough reflex. *Biol. Pharm. Bull.* **36**, 501–508 (2013).
57. Yu, X., Mei, Z., Chen, C. & Chen, W. Ranking power spectra: A proof of concept. *Entropy* **21**, (2019).
58. Kasim, N. et al. A comparison of the power of breathing sounds signals acquired with a smart stethoscope from a cohort of COVID-19 patients at peak disease, and pre-discharge from the hospital. *Biomed. Signal Process Control* **78**, 103920 (2022).
59. Radin, J. M., Wineinger, N. E., Topol, E. J. & Steinhubl, S. R. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. *Lancet Digit Health* **2**, e85–e93 (2020).
60. Mishra, T. et al. Early Detection Of COVID-19 Using A Smartwatch. *medRxiv* 2020.07.06.20147512 https://doi.org/10.1101/2020.07.06.20147512 (2020).
61. Bogu, G. K. & Snyder, M. P. Deep learning-based detection of COVID-19 using wearables data. *medRxiv* 2021.01.08.21249474 (2021).
62. Walter, J. R. et al. Use of artificial intelligence to develop predictive algorithms of cough and PCR-confirmed COVID-19 infections based on inputs from clinical-grade wearable sensors. *Sci. Rep.* **14**, 8072 (2024).
63. Parikh, R., Mathai, A., Parikh, S., Chandra Sekhar, G. & Thomas, R. *Understanding and Using Sensitivity, SpeciP City and Predictive Values*.
64. Pokhrel, P., Hu, C. & Mao, H. Detecting the coronavirus (CoVID-19). *ACS sensors* **5**, 2283–2296 (2020).
65. Bisoffi, Z. et al. Sensitivity, specificity and predictive values of molecular and serological tests for COVID-19: A longitudinal study in emergency room. *Diagnostics* **10**, 669 (2020).
66. Woloshin, S., Patel, N. & Kesselheim, A. S. False Negative Tests for SARS-CoV-2 Infection — Challenges and Implications. *N. Engl. J. Med.* **383**, 669 (2020).
67. Healy, B., Khan, A., Metezai, H., Blyth, I. & Asad, H. The impact of false positive COVID-19 results in an area of low prevalence. *Clin. Med., J. R. Coll. Physicians Lond.* **21**, e54–e56 (2021).
68. Surkova, E., Nikolayevskyy, V. & Drobniewski, F. False-positive COVID-19 results: hidden problems and costs. *Lancet Respir. Med* **8**, 1167–1168 (2020).
69. Penn, M. Statistics Say Large Pandemics Are More Likely Than We Thought. *Duke Global Health Institute* (2021).
70. Haileamlak, A. Pandemics will be more frequent. *Ethiop. J. Health Sci.* **32**, 228 (2022).
71. Noonan, V. & Dean, E. Submaximal exercise testing: Clinical application and interpretation. *Physical therapy* **80**, 782–807 (2000).
72. Altini, M. et al. Cardiorespiratory fitness estimation using wearable sensors: Laboratory and free-living analysis of context-specific submaximal heart rates. *J. Appl Physiol.* **120**, 1082–1096 (2016).
73. Pan, J. & Tompkins, W. J. A Real-Time QRS Detection Algorithm. *IEEE Trans. Biomed. Eng.* **BME-32**, 230–236 (1985).
74. Champseix, R., Ribiere, L. & Couedic, C. L. A Python Package for Heart Rate Variability Analysis and Signal Preprocessing. *J Open Res Softw* **9**, (2021).
75. Kim, J. W., Seok, H. S. & Shin, H. Is Ultra-Short-Term Heart Rate Variability Valid in Non-static Conditions? *Front Physiol.* **12**, 596060 (2021).
76. Bates, A., Ling, M. J., Mann, J. & Arvind, D. K. Respiratory rate and flow waveform estimation from tri-axial accelerometer data. in *2010 International Conference on Body Sensor Networks, BSN 2010* 144–150. https://doi.org/10.1109/BSN.2010.50 (2010).
77. Chatterjee, N. A. et al. Admission respiratory status predicts mortality in COVID-19. *Influenza Other Respir. Viruses* **15**, 569–572 (2021).
78. Natarajan, A., Su, H.-W., Heneghan, C., Blunt, L. & Niehaus, L. Measurement of respiratory rate using wearable devices and applications to COVID-19 detection. *NPJ Digit Med.* **4**, 136 (2021).
79. Your lungs and exercise. *Breathe (Sheffield, England)*, **12**, 97–100 (2016).
80. Doddabasappla, K. & Vyas, R. Statistical and Machine Learning-Based Recognition of Coughing Events Using Triaxial Accelerometer Sensor Data from Multiple Wearable Points. *IEEE Sens. Lett.* **5**, (2021).
81. Andrès, E., Gass, R., Charloux, A., Brandt, C. & Hentzler, A. Respiratory sound analysis in the era of evidence-based medicine and the world of medicine 2.0. *J. Med Life* **11**, 89–106 (2018).
82. Al Ismail, M., Deshmukh, S. & Singh, R. Detection of covid-19 through the analysis of vocal fold oscillations. in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* vols 2021-June 1035–1039 (Institute of Electrical and Electronics Engineers Inc., 2021).
83. Brown, C. et al. Exploring Automatic Diagnosis of COVID-19 from Crowd-sourced Respiratory Sound Data. **11**, (2020).
84. Pedregosa, F. et al. *Scikit-Learn: Machine Learning in Python*. *Journal of Machine Learning Research* vol. 12 (2011).
85. Bergstra, J. & Bengio, Y. Random search for hyper-parameter optimization. *Journal of Machine Learning Research* **13**, (2012).
86. Kuhn, M. & Johnson, K. *Applied Predictive Modeling*. *Applied Predictive Modeling*. https://doi.org/10.1007/978-1-4614-6849-3 (2013).
87. Liu, F. T., Ting, K. M. & Zhou, Z. H. Isolation-based anomaly detection. *ACM Trans Knowl Discov Data* **6**, (2012).
88. Galli, S. Feature-engine: A Python package for feature engineering for machine learning. *J. Open Source Softw.* **6**, 3642 (2021).
89. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. *SMOTE: Synthetic Minority Over-Sampling Technique*. *Journal of Artificial Intelligence Research* vol. 16 (2002).
90. Lemaître, G., Nogueira, F. & Aridas, C. K. *Imbalanced-Learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning*. *Journal of Machine Learning Research* vol. 18 (2017).
91. Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* vols 13-17-August-2016 785–794 (Association for Computing Machinery, 2016).

Acknowledgements

This study was funded by The United States-India Science & Technology Endowment Fund (USISTEF), COVID-19 Ignition Grant reference no. USISTEF/COVID-II/104/2020, and the Max Nader Lab for Rehabilitation Technologies and Outcomes Research at the Shirley Ryan AbilityLab, Chicago, IL. The work of M.F. was supported by the NIH under Grant T32

HD07418. We'd also like to thank Chaithanya Krishna Mummidisetty, Luca Lonini, Nicholas Shawen, Chandra Jayaraman, Sung Shin, and Kiana Mohammadian.

Author contributions

O.K.B. is considered first author. Conception, design, and study direction: O.K.B., J.M., S.A., N.C.V., M.F., V.N., A.C., J.A.R., A.J.; Data acquisition: O.K.B., J.M., S.A., N.C.V, V.T., V.P., A.K., S.C.; Data analysis: O.K.B., J.M., S.A., N.C.V., M.F., S.X., J.L., H.J., J.A.R.; Manuscript writing and review: O.K.B., J.M., S.A., N.C.V., M.F., J.L., V.T., V.P., A.K., S.C., A.C., V.N., S.X., H.J., J.A.R., A.J. All authors have read and approved the manuscript.

Competing interests

S.X. and J.L. have stock in and are employees of Sibel Health, which develops the technology used for this study. J.A.R. is a co-founder of Sibel Health. An international patent (SYSTEMS AND METHODS FOR RAPIDLY SCREENING FOR SIGNS AND SYMPTOMS OF DISORDERS, application number PCT/US2021/042846) has been filed by Shirley Ryan AbilityLab covering the methodology presented in this work, specifically the use of a processor or processing element and sensor(s) to detect physiological and movement changes associated with a disorder from predetermined scripted activities. O.K.B., M.F., and A.J. are listed as inventors on this patent. Authors J.M., S.A., N.C.V., V.T., V.P., A.K., S.C., A.C., V.N., H.J. declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41746-024-01287-2.

Correspondence and requests for materials should be addressed to Arun Jayaraman.