An Analysis of COVID-19 Recovery Duration: Smartwatch Activity and Self-Report

Z. Schäfer[1], T. Rolighed Thomsen[2], H.J. Schulz[1],F. Mølgaard Thayssen[3],
M. Christensen[3], L. Lindholt[4], K. Svendsen[5], A. Mathisen[1], K. Grønbæk[1]

***Abstract*— This study investigates the relationship between smartphone tracking data and self-reported COVID-19 experiences to better understand recovery patterns after COVID-19 infection in patients. We developed COronaVIden, an app that collects activity data through the iOS Health and Google Fit APIs, along with questionnaires on the timeline of infection, symptoms, and WHO-5 well-being scores. Of more than 5,000 participants (with 632 positive cases) who provided data between January 2019 and November 2021, 177 met our inclusion criteria. The average recovery duration based on activity data is 23.04±15.52 days, while self-perceived recovery duration average is 16.51±10.33 days. We developed two predictive models: In model 1 we used physical activity and demographic data, and in model 2 we used physical activity, demographic data, as well as additional post-infection symptoms that were made available to us by the users. We achieved a mean absolute error of 10.66 and 9.42 respectively, showing 29% and 37% improvement over the baseline, which is the median of recovery duration. The Shapley method was used to assess the contributions of features to the duration of recovery. Key findings indicate that being female, older and experiencing fatigue are correlated with longer recovery periods. Higher physical activity before infection was correlated with a shorter recovery time, possibly indicating better baseline health, although this was not significant in our population. Our findings demonstrate that wearables and smartphones can effectively track disease recovery patterns, offering valuable information for public health strategies, although more research is needed to account for potential confounding factors.**

I. INTRODUCTION

Since the World Health Organization declared the COVID-19 pandemic a public health emergency in January 2020, it has evolved into a significant and important public health challenge, marked by seasonal variations. While most infected individuals experience mild to moderate symptoms lasting 4-6 days without the need for hospitalization, around 5% develop severe illnesses requiring intensive medical care [3], [4]. Notably, the elderly and those with chronic conditions face a higher risk of a severe disease course. During the COVID-19 pandemic, smartphones and wearables have played an instrumental role in shaping our responses to the challenges posed by the virus. Notable examples include tracking apps designed for monitoring and managing the spread of COVID-19. Additionally, applications utilizing physiological responses through wearables, such as alterations in resting heart rate, step counts, and changes in sleep patterns, have been employed to collect valuable insights into the virus impact on individuals [1], [2]. In a study conducted in [1], an online detection algorithm, CuSum was developed for the real-time detection of early stages of COVID-19 symptoms. The authors show that in a substantial number of cases, they can detect the early onset of the disease based on changes in the sleep pattern, heart rate and daily steps. The results are used for an early diagnosis and therefore mitigating the spread of COVID-19 disease. In another study [2], the authors introduce a real-time smartwatch-based alerting system that can detect aberrant physiological and activity signals associated with early COVID-19 symptoms. The algorithm also identified signals resulting from other events such as stress, alcohol consumption, intense exercise, and travel, indicating the potential for a broader health monitoring beyond infectious diseases. The alerts in the system are triggered based on deviations from physiological and activity signals, allowing for early self-isolation, testing, and effective allocation of healthcare resources. The impact of activity levels during the UK's COVID-19 lockdown was highlighted in a study by [5]. Tracking activity data before, during, and after the lockdown revealed a notable decline in physical activity. Young people were more active before the lockdown but showed the least activity after restrictions were lifted. In contrast, those older than 65 increased their activity as soon as the lockdown was relaxed and remained more active throughout the lockdown period.

In [7], the impact of pre-pandemic physical activity on hospitalization in older adults was studied. Combining three large-scale randomized clinical trials, the study included 61,557 participants aged 45 years or older. Physical activity levels, calculated based on metabolic equivalent hours (MET), categorized participants as inactive, insufficiently active, or sufficiently active. The study found that sufficiently active individuals had a significant reduction in infection and hospitalization. A similar study by [8] also reported an association between pre-COVID physical activity and reduced COVID-19 severity.

The widespread adoption of smartphones and wearables enables real-time monitoring of disease trajectories through physiological activity changes. Research by [9] demonstrated prolonged physiological effects in COVID-19 patients, including elevated resting heart rates lasting 2-3 months, along with altered step counts and sleep patterns. Our study investigates how smartphone tracking data and self-reported

[1]Department of Computer Science, Aarhus University, Aarhus, Denmark
[2]Department of Chemistry and Biosciences, Aalborg University, Aalborg, Denmark
[3]Alexandra Institute, Aarhus, Denmark
[4]Region Midtjylland, Regionshospitalet Horsens, Denmark
[5]Department of Public Health, Aarhus University, Denmark

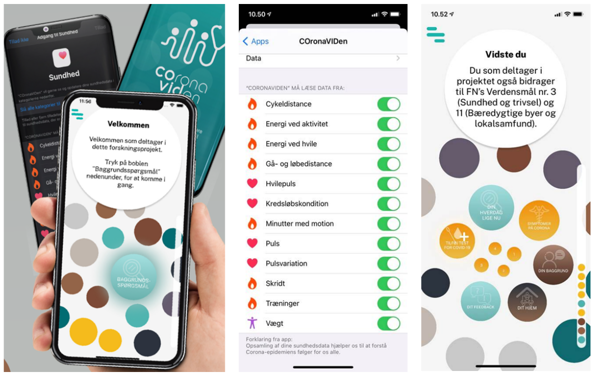

Fig. 1: **FIGURE 1.** The COronaVIDen App interface (in Danish) and list of data collection permissions.

experiences during COVID-19 infection can enhance understanding of the recovery process, with potential applications to future infectious diseases. In the following sections, we detail our app design for symptom collection, data gathering methods, and predictive models using Shapley values to forecast recovery duration based on physical activities, followed by key findings and study limitations. All analysis code is available as Jupyter Notebooks on Github.

II. Materials & Methods

A. The COranaVIDen app

We developed the COronaVIDen app, available on both iOS and Android platforms that was designed to gather data through the iOS Health and Google FIT APIs. Figure 1 illustrates screen captures from the app's user interface (UI).

The recruitment process started on July 14th, 2020, and continued until November 8th, 2021, which is the end of the recruitment process. Data collection occurred in real-time, 24/7. To put the collected data into context, participants were also asked for access to historic activity tracking data that was by default gathered and stored by their devices for the period of January 1st, 2019, until November 8th, 2021. Access to the historic activity tracking data recorded by the cell phones' operating systems (iOS or Android) was explicitly covered by the ethics board approval, made transparent to the participants, and the historic data was only accessed during the runtime of the study as part of the study.

The app includes various concise questionnaires that are visually represented as expanding bubbles, prompting users to answer relevant questions. In Figure 1, a column on the right side displays small ellipses indicating the number and types of questionnaires already answered by users. Some bubbles allow for multiple responses, covering events such as symptoms, test results, and recovery from a COVID-19 infection. All completed questionnaires are timestamped, allowing for comparison with automatically collected data from sensors on smartphones or wearables during the same period, typically with precision within an hour.

When a user opened the app for the first time, a consent form had to be filled in to allow data collection and activity tracking, including collection of historic data from January 1st, 2019. The participants had the possibility to withdraw their consent at any time. The data was anonymized and stored in a cloud database, then promptly transferred in real-time to a secure data environment managed and operated by the Alexandra Institute, a Danish Research and Technology Organization. In compliance with GDPR regulations, the Alexandra Institute serves as the Data Processor, while Aarhus University assumes the role of the Data Controller. Historic data collection commenced at the time of consent, marking the initiation of continuous 24/7 data collection for users. Initial questionnaires were prompted by the app shortly after installation and consent. Data collection utilized a message queue system and a NoSQL database to prevent data loss. Upon request, the data was transferred to a view-only SQL database, facilitating data analyses through Jupyter Notebook ML and Visual Analytics tools. The analysis of the data was conducted from a frozen view captured on 08.11.2021, hence data until and inclusive 8.11.2021 was analyzed.

B. Population & Data Collection

The project aimed for recruiting app users before having COVID-19 infection, and the data collection was performed "in the wild" without a specified cohort or any constraints. The public was actively encouraged to download the COronaVIDen app from either the Apple AppStore or Google Play. Recruitment efforts spanned various channels, including Facebook, LinkedIn, campaigns at libraries and companies, personal email outreach, and distribution through e-Boks (a trusted Nordic provider of secure platforms and secure digital mail) to all employees in the Capital Region of Denmark. This approach involved engaging with a dynamically expanding and inclusive population group, selected solely based on the criterion of downloading the COronaVIDen app and consistently contributing personal health data from January 1st, 2019, to November 8th, 2021.

As of November 8th, 2021, 4,083 users downloaded the app and maintained consent throughout the study period. In Table 1 we show the registered details of all users. However, not everyone was eligible for the presented study due to the data collection approach. Many users lacked sufficient activity tracking histories, others did not fully respond to certain questionnaire sections, and many did not experience a Covid-19 infection within the study period. Therefore, a set of inclusion criteria was applied to filter the population of app downloaders into a homogenous group with consistent data quality:

1) Only users who provided responses to all questions regarding potential COVID-19 infections, including recalled start and end dates, were included. Among 718 total positive cases, 632 cases had registered the recovery date and the onset of Covid date.
2) Users with adequate step data at least 90 days before infection onset and at least six hours of activity per day were selected. From 500 cases with activity data, 177 users remained which satisfied these criteria. This is also to provide sufficient historical context for the

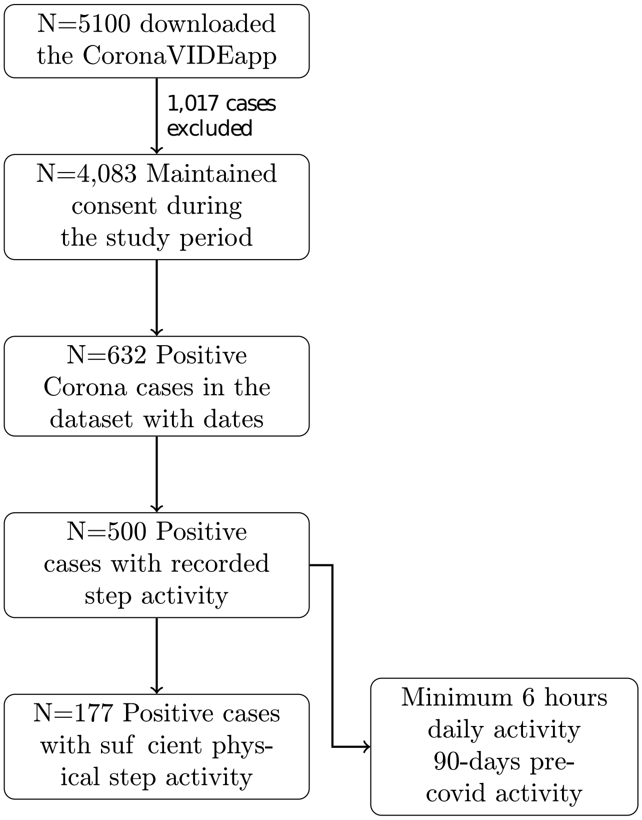

Fig. 2: Study selection flowchart showing the filtering of participants from initial app downloads to final analysis cohort.

change point algorithm. In Fig. 2 we show the registered information of the remaining population.

The filtering process is summarized in Figure. 2.

C. Ethics and Informed Consent

Users were asked in the app for written consent to collect information according to the agreed research purpose and in accordance with the Data Protection Act in Denmark (which complies with GDPR). The research project was submitted to the national Committee on Health Research Ethics. It was not considered as a health science research project and should therefore not be reported to the Committee on Health Research Ethics. The project was approved by the Aarhus University internal Ethics Board before being approved again for release by Apple and Google. Following data were collected: Information on registration, Google ID or Apple ID and data from questionnaires in the app such as age, statements, Corona tests, gender and family relationships, and activity data from smartphone and smartwatch through Apple HealthKit or Google Fit. The users gave their written consent through the mobile application. Additionally, information on locations was collected in the background. Collected information has not been shared with other than CoronaLytics project partners. None of the authors had access to information that could identify individual participants either during or after the study. The dataset was completely anonymized to protect participant privacy.

D. Estimating Recovery Duration

Initially, we aimed to estimate the recovery time for users who tested positive during the study using step count data. Among the 632 participants who tested positive, we selected users that have adequate step count information both before and after the infection. Specifically, we required a 90-day window before the onset of infection and at least 120 days after diagnosis, with a minimum of 6 hours of registered step activity per day. To address missing values, we employed interpolation techniques and normalized the data to represent the number of steps per hour. This leaves us with a total of 177 users who tested positive. We defined the recovery time as the time point when a notable change in step counts occurs following the onset of COVID-19. The assumption is that users will experience decreased activity levels following infection, eventually transitioning to a higher activity level, significant enough to be recognized as a change point or recovery state. This is achieved by employing a change point algorithm from the Python library rupture [10]. The library is for detection of change points in an offline setting for non-stationary signals. We used a dynamic programming algorithm for detecting the changes, where it computes the best partition for which the sum of errors is minimum. See [10] for more details on the library and algorithm. This estimated change point is then used in a predictive model versus the self-reported duration to gain insights into the differences in predictions.

E. Prediction of Recovery Time and Shapley Values

We employed all the features that we obtained from data to predict the reported and estimated recovery time for participants separately. The set of features used are as follows: Demographic information (gender, age, education level, health condition before infection), symptoms within the first two weeks after onset of COVID-19 (consists of 27 symptoms that users had to check if they were encountered, see appendix for the list of all symptoms), tracking information which is the average step counts in a week prior to the COVID-19 infection. We applied Extreme Gradient Boosting (XGBoost) algorithm [11] to predict the recovery times based on the features. This model is selected due to its robustness to mixed data types. To find the best set of parameters, a grid search approach with mean absolute error as the optimization function are used.

In the second step, we leverage Shapley values to understand how individual feature values contribute to the predicted outcomes and explain varying recovery times. Originally from game theory, Shapley values compute each player's contribution to a coalition's payout function. In machine learning, we apply this concept by approximating the Shapley value for each feature's contribution to the prediction, given all other observations.

The Shapley value framework is governed by four fundamental axioms that enable feature estimation. The efficiency axiom ensures that the total change in prediction attributed

Category	Total	Positives
n	4083	632
Female	2553 (62.5)	413 (65.3)
Male	730 (17.9)	84 (13.3)
Other	5(0.1)	0(0)
Unknown	795(19.5)	135(21.4)
Age		
15-29	522 (12.8)	73 (11.6)
30-39	388 (9.5)	82 (13.0)
40-49	767 (18.8)	133 (21.0)
50-59	915 (22.4)	139 (22.0)
60-69	572 (14.0)	59 (9.3)
70-79	109 (2.7)	10 (1.6)
80-89	7 (0.2)	1 (0.2)
Unknown	803 (19.7)	135 (21.4)
Yes	1132 (27.7)	162 (25.6)
No	2115 (51.8)	331 (52.4)
Unknown	836 (20.5)	139 (22.0)
Education		
Compulsory school	199 (4.9)	16 (2.5)
One or more short courses	18 (0.4)	3 (0.5)
Upper secondary certificate	192 (4.7)	30 (4.7)
VET certificates	404 (9.9)	60 (9.5)
Short higher education (2-3yrs)	333 (8.2)	50 (7.9)
Intermediate higher education	1274 (31.2)	211 (33.4)
Long higher education(>4yrs)	849 (20.8)	126 (19.9)
Unknown/missing	814 (19.9)	136 (21.5)
Existing Health Condition		
Musculoskeletal complications	432 (10.6)	58 (9.2)
Overweight	186 (4.6)	25 (4.0)
Psychiatric disease	114 (2.8)	22 (3.5)
Lung disease	98 (2.4)	13 (2.1)
Cardiovascular disease	93 (2.3)	8 (1.3)
Gastrointestinal complication	51 (1.2)	10 (1.6)
Diabetes	49 (1.2)	5 (0.8)
Neurological disease	48 (1.2)	5 (0.8)
Skin complications	41 (1.0)	1 (0.2)
Kidney or bladder disease	14 (0.3)	3 (0.5)
Unknown/missing	3336 (81.7)	537 (85.0)

TABLE I: Demographics and health conditions of study participants.

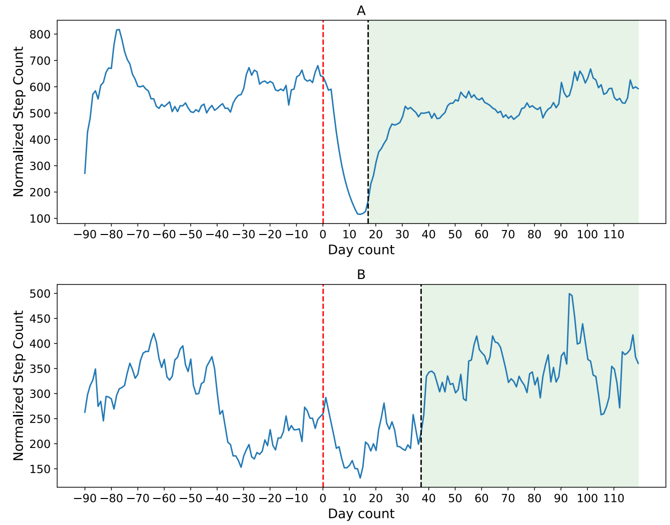

Fig. 3: Two examples of physical activity (step counts) before and after diagnosis of COVID-19 (red line dash). The black line indicates the estimated recovery time based on change point detection algorithm. Subject A is in category age of 30-39, with average daily steps of 645 pre-COVID-19 and subject B is in category age of 50-59 with pre-COVID-19 step activity of 245 steps. Both subjects are female.

to all features equals the difference between the model's prediction and the dataset's average prediction. Under the symmetry axiom, features with identical impacts across all possible combinations receive equal Shapley values. The dummy feature axiom states that features contributing nothing to predictions receive a Shapley value of zero. Finally, the additivity axiom specifies that in combined models, a feature's Shapley value is the sum of its values from individual models or subsets.

These axioms create a fair and consistent framework for attributing feature importance across different contexts. By approximating Shapley values for each feature, we can identify which features are most crucial in understanding patient recovery time. For detailed Shapley value approximation methods and mathematical formulations, we refer to [12] and its references.

III. Results

A. Recovery time distribution

In Fig. 3 we show step counts for two subjects along with the estimated recovery times. The onset of COVID-19 is indicated by zero and therefore the step counts are divided into after and before COVID-19 onset.

The second time point with dashed vertical line is the predicted change point, i.e., the time at which a significant change since COVID-19 onset in terms of step count activity is detected. In most subjects the immediate activity of patients after COVID-19 onset significantly reduces, and after an average of 23.04±15.52 days increases based on the estimated change points. This contrasts with 16.51±10.33 days based on self-reported recovery duration.

In our study, we removed subjects with a reported recovery duration longer than 58 days based on the interquartile range. The decision to omit such subjects was due to outliers and a misinterpretation of recovery time based on estimated change points. In Fig. 4, we provide an example in which the recovery estimate is 67 days since the onset of COVID-19, which significantly exceeds the patient's reported recovery of only 4 days. This inconsistency can significantly skew the overall analysis when we rely on recovery duration based on step counts. There are no significant changes in the activity of the step count after the onset of COVID-19 and the change detected at the end of the measurements is not informative. When we inspected other features of this subject, we noticed that the subject reported only one symptom, which is body ache. Therefore, we presumed that subjects with a high estimated recovery duration and a low number of reported symptoms should be among the category of patients with mild symptoms who had a quick recovery duration.

For participants with an estimated recovery of more than 58 days and less than three symptoms, the average reported recovery is 19.63 ± 3.0, with a median of 9. These subjects have an average of 1.72 ± 0.90 recorded symptoms. Therefore, to have an objective measure to remove outliers,

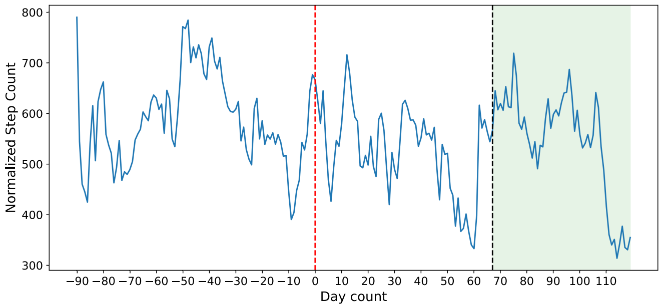

Fig. 4: An example of a subject with no significant immediate change in the step activity after COVID-19 diagnosis. The reported recovery registered for this patient is 4 days and total number of symptoms reported is 1. We excluded such subjects from the study since the estimated change point does not reflect their recovery duration.

we excluded subjects based on the interquartile range (11 subjects were excluded in this case), corresponding to 58 days duration.

B. Prediction of Recovery Time and Model Explainability

We employ two distinct sets of features to predict the recovery time and use Shapley values to interpret the results for individual cases. In the first model, we use the step count data (the average daily step counts a week prior to the onset of COVID-19) and demographic information (age, sex, and educational level) to predict the recovery duration. The dependent variable (recovery time) is defined in three distinct ways:

1) minimum between the estimated recovery time and reported recovery duration (MR)
2) reported recovery duration (RR)
3) estimated recovery durations based on the detected change points (ER)

In the case of MR, we set the estimated recovery duration to be the minimum between the reported and the estimated recovery durations for participants with a low number of reported symptoms and estimated recovery longer than 58 days. This way more subjects are included in the analysis. We therefore apply a prediction model for three different dependent variables, namely, RR, MR, and ER. By applying a predictive model on each setting, we aim to obtain a more reliable and accurate estimate and impact of different features on the length of recovery.

In the other setting for the features, we incorporate step count data and demographic information along with all symptomatic features reported within the first two weeks of COVID-19 diagnosis. In all cases, we fit a Gradient Boosting model and compute the baseline prediction which is the median of the samples. Table II shows the accuracy of the models for prediction of recovery duration (RR, ER, MR) along with the baseline prediction for using additional symptomatic features. Both models show substantial improvements over the baseline across all different defined recovery durations. Model 1 shows improvements ranging from 29% to 40.4% and model 2 shows improvements ranging from 32.2% to 41.5%.

Recovery Duration	**RR**	**ER**	**MR**
Model 1	10.66	11.44	7.75
Model 2	9.42	11.53	7.61
Baseline	15	17	13

TABLE II: Mean absolute errors (MAE) in days for model 1 and 2. The reported MAE in days are reported for the following dependent variables; RR: Reported Recovery duration, ER: Estimated Recovery, and MR indicates the minimum recovery. The baseline accuracy for each variable corresponds to the median of the corresponding variable.

IV. Discussion

The results in Figure 5 and Figure 6 present the impact of specific feature values on predicting the recovery duration, i.e., whether an increase in the feature values have a positive contribution on the prediction of recovery duration (longer) or a negative contribution (shorter). In Figure 5(a), despite variability in the distribution of Shapley values, high step counts tend to lean towards the negative Shapley estimates, implying a shorter recovery duration, which indicates its protective effect. In other words, higher physical activity prior to COVID-19 diagnosis seems to correlate with a shorter predicted recovery duration. As expected, most subjects with a higher age are associated with a positive Shapley value, indicating an increase in the duration of recovery. The Shapley values calculated for the estimated recovery duration in models 1 and 2 have a higher variability than the case for self-reported recovery data (i.e., Figure 6(a) versus Figure 6(b). This discrepancy can be attributed to several factors, with the primary limitation being outliers in the data and the misinterpretation of change points estimated by the change point detection algorithm. Another notable observation is the negative Shapley values associated with higher education levels and being male. Education level may be linked to occupational factors, which could influence individuals' recovery trajectories. It must be noted that due to collection and filtering of data, it can be the case that the population of male in general is in a better health condition than women. In numerous previous studies it is widely acknowledged that men face a higher risk of severe or fatal COVID-19 outcomes compared to women [13]–[15]. However, multiple longitudinal studies have shown women are at a higher risk of developing long COVID-19. In a study conducted by [16] which analyzed 10 different studies, being female was associated with a higher risk of developing long-COVID-19. For more details see [16] and the references therein. Other findings corresponding to the step counts in Figure 5(b,c) and Figure 6(b,c) reflect similar outcomes with a higher variability.

In Figure 6, Shapley value distribution presented for the symptomatic features in our model. In the case of reported recovery duration, we can see a clear correlation between long recovery durations and presence of symptoms such as fatigue, body ache and difficulty in movement. This finding

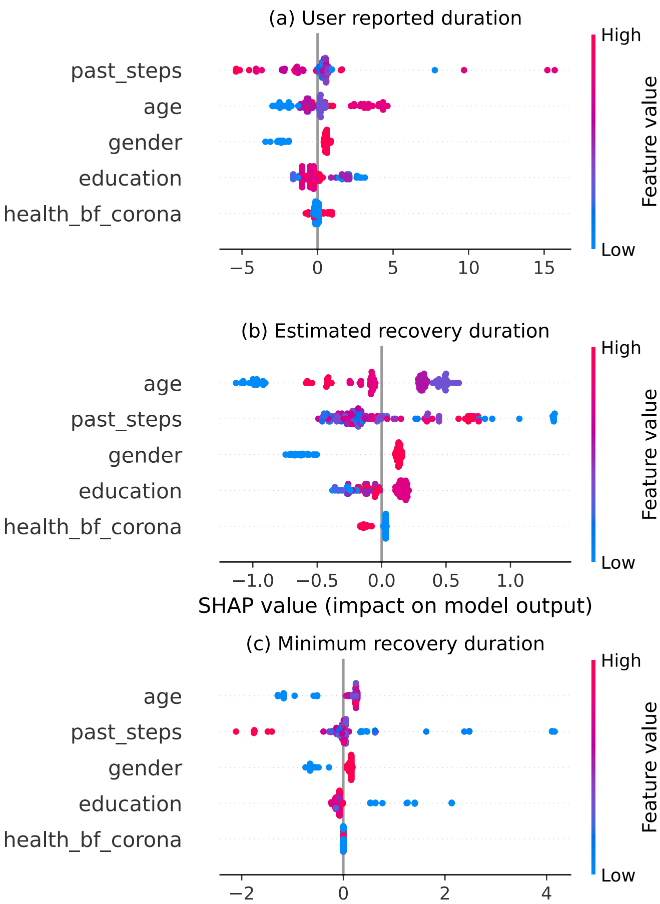

Fig. 5: Summary of feature importance using Shapley values. A positive value (shown in red) means that feature contributes to longer predicted recovery times. Conversely, a negative value (shown in blue) indicates that feature contributes to shorter predicted recovery times. The larger the absolute value, the stronger the feature's influence on the prediction.

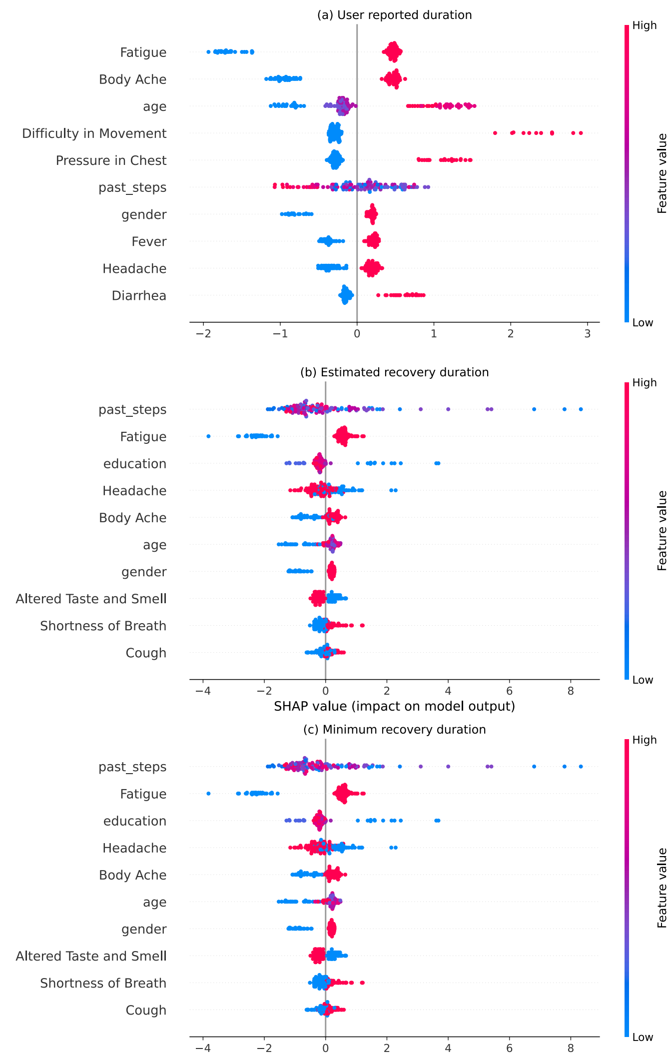

Fig. 6: An example of a subject with no significant immediate change in the step activity after COVID-19 diagnosis. The reported recovery registered for this patient is 4 days and total number of symptoms reported is 1. We excluded such subjects from the study since the estimated change point does not reflect their recovery duration.

aligns with previous research and may also be attributable to the phenomenon of long COVID-19, wherein individuals experience prolonged duration of reported fatigue, dyspnea, and joint pain [17]. In a recent study [18], persistent fatigue is observed to be one of the most common symptoms experienced by individuals with long-COVID-19, with 72% of individuals reporting it. The study suggests that such symptoms in patients with long COVID-19 may be associated with dysregulation in the nervous system, including the central, peripheral, and autonomic nervous systems.

To examine the effect of different features in predicting recovery duration, we analyze the individual SHAP distributions for the subjects presented in Figure3. In Figure7, we show the SHAP summary plot for two subjects and the difference in their SHAP values. Figure 7(a) and Figure7(b) correspond to two female subjects with self-reported recovery durations of 37 and 17 days, respectively. Our model predicts recovery durations of 22 and 13 days for these subjects. For the first subject (plot a), the model's prediction is significantly influenced by pre-COVID-19 physical inactivity. The presence of symptoms such as fatigue and cough also appear to increase the predicted recovery duration. For the second subject (plot b), the recovery prediction is mostly influenced by symptoms such as fatigue, body ache, and pressure in the chest. Unlike the first subject, pre-COVID-19 physical activity is not a significant factor for the second subject.

A. Limitations and Future Work

Utilizing unguided or "in the wild" data, such as data collected retrospectively, presents both opportunities and challenges in understanding population health challenges like COVID-19. While this approach allows for the analysis of a large dataset, it also introduces potential biases due to data quality issues, inconsistent tracking habits, and demographic

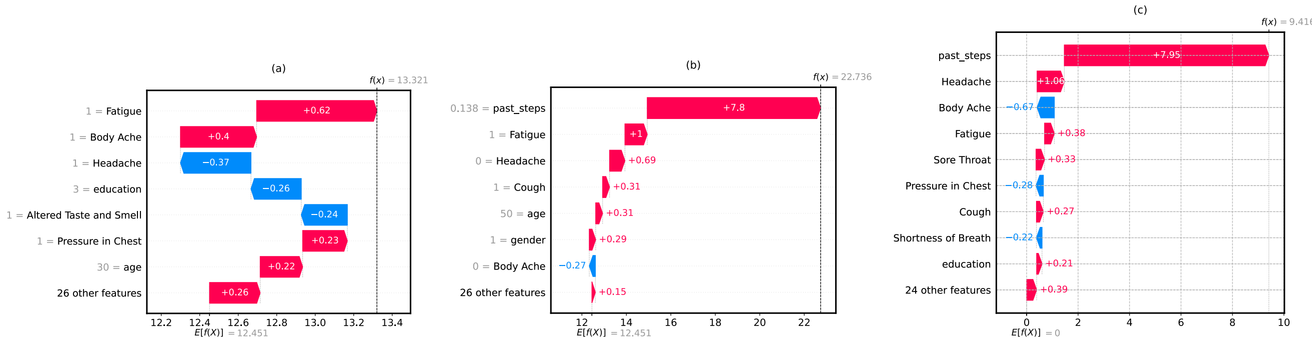

Fig. 7: Plots a,b compare SHAP values between two female patients with different recovery duration. The estimated recovery time for subject 1 and 2 are 37 and 17 days respectively. Plot c illustrates the differences in feature contributions that led to these varying predictions.

variations among users.

The study faced challenges with retrospectively collected physical activity data during COVID-19, particularly in distinguishing between reduced tracking and actual decreased activity. While 39 subjects were excluded due to missing data, research suggests that combining physical activity tracking with additional metrics like heart rate variability (HRV) could provide better insights into recovery patterns. Studies by [19] and [20] demonstrate that HRV and smartphone-based cardio-respiratory fitness measurements can effectively monitor health conditions and predict disease outcomes. Though consumer-based activity trackers offer promising opportunities for population health analysis, successful implementation requires addressing data quality, bias, and privacy concerns.

V. Conclusion

Our findings revealed that being female, older age, and experiencing fatigue were consistently associated with longer COVID-19 recovery times, while higher pre-infection physical activity generally indicated faster recovery, likely due to better overall health. The predictive models showed 35% improvement over baselines. Future research will focus on developing more comprehensive models incorporating additional physiological responses and biomarkers, while improving data sampling methods and expanding demographic diversity.

References

[1] T. Mishra, and others, "Pre-symptomatic detection of COVID-19 from smartwatch data," *Nature Biomedical Engineering*, vol. 4, no. 12, pp. 1208–1220, Dec. 2020.

[2] A. Alavi, and others, "Real-time alerting system for COVID-19 and other stress events using wearable data," *Nature Medicine*, vol. 28, no. 1, pp. 175–184, Jan. 2022.

[3] C. Wang, P. W. Horby, F. G. Hayden, and G. F. Gao, "A novel coronavirus outbreak of global health concern," *The Lancet*, vol. 395, no. 10223, pp. 470–473, 2020.

[4] World Health Organization, "WHO Coronavirus (COVID-19) Dashboard," 2024.

[5] H. McCarthy, H. W. W. Potts, and A. Fisher, "Physical activity behavior before, during, and after COVID-19 restrictions: Longitudinal smartphone-tracking study of adults in the United Kingdom," *Journal of Medical Internet Research*, vol. 23, no. 2, 2021.

[6] E. J. Thompson, D. M. Williams, A. J. Walker, R. E. Mitchell, C. L. Niedzwiedz, T. C. Yang, C. F. Huggins, A. S. F. Kwong, R. J. Silverwood, G. Di Gessa, and others, "Long COVID burden and risk factors in 10 UK longitudinal studies and electronic health records," *Nature Communications*, vol. 13, no. 1, p. 3528, Jun. 2022.

[7] D. Muñoz-Vergara, and others, "Prepandemic Physical Activity and Risk of COVID-19 Diagnosis and Hospitalization in Older Adults," *JAMA Network Open*, vol. 7, no. 2, p. e2355808, Feb. 2024.

[8] L. Malisoux, A. Backes, A. Fischer, G. Aguayo, M. Ollert, and G. Fagherazzi, "Associations between physical activity prior to infection and COVID-19 disease severity and symptoms: results from the prospective Predi-COVID cohort study," *BMJ Open*, vol. 12, no. 4, p. e057863, Apr. 2022.

[9] J. M. Radin, and others, "Assessment of Prolonged Physiological and Behavioral Changes Associated With COVID-19 Infection," *JAMA Network Open*, vol. 4, no. 7, p. e2115959, Jul. 2021.

[10] C. Truong, L. Oudre, and N. Vayatis, "Selective review of offline change point detection methods," *Signal Processing*, vol. 167, p. 107299, Feb. 2020.

[11] T. Chen, and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016.

[12] B. Rozemberczki, and others, "The shapley value in machine learning," *arXiv preprint arXiv:2202.05594*, 2022.

[13] C. Gebhard, V. Regitz-Zagrosek, H. K. Neuhauser, R. Morgan, and S. L. Klein, "Impact of sex and gender on COVID-19 outcomes in Europe," *Biology of Sex Differences*, vol. 11, no. 1, p. 29, Dec. 2020.

[14] J. M. Jin, and others, "Gender Differences in Patients With COVID-19: Focus on Severity and Mortality," *Frontiers in Public Health*, vol. 8, Apr. 2020.

[15] D. E. L. Promislow, "A Geroscience Perspective on COVID-19 Mortality," *The Journals of Gerontology: Series A*, vol. 75, no. 9, pp. e30–e33, Sep. 2020.

[16] E. J. Thompson, and others, "Long COVID burden and risk factors in 10 UK longitudinal studies and electronic health records," *Nature Communications*, vol. 13, no. 1, p. 3528, Jun. 2022.

[17] A. Carfì, R. Bernabei, and F. Landi, "Persistent Symptoms in Patients After Acute COVID-19," *JAMA*, vol. 324, no. 6, p. 603, Aug. 2020.

[18] N. J. Maffitt, M. Germann, A. M. E. Baker, M. R. Baker, S. N. Baker, and D. S. Soteropoulos, "Recovery of neurophysiological measures in post-COVID fatigue: a 12-month longitudinal follow-up study," *Scientific Reports*, vol. 14, no. 1, p. 8874, Apr. 2024.

[19] N. Bourdillon, S. Yazdani, L. Schmitt, and G. P. Millet, "Effects of COVID-19 lockdown on heart rate variability," *PLOS ONE*, vol. 15, no. 11, 2020.

[20] M. T. Eades, A. Tsanas, S. P. Juraschek, D. B. Kramer, E. Gervino, and K. J. Mukamal, "Smartphone-recorded physical activity for estimating cardiorespiratory fitness," *Scientific Reports*, vol. 11, no. 1, p. 14851, 2021.