Wearable Data Reveals distinct Characteristics of Individuals with Persistent Symptoms after a SARS-CoV-2 Infection

Katharina Ledebur[1,2,3†*], Marc Wiedermann[3†], Christian Puta[4,5,6], Stefan Thurner[1,2,7], Peter Klimek[1,2,8,9], Dirk Brockmann[3,10,11]

[1]Institute of the Science of Complex Systems, CeDAS, Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria
[2]Complexity Science Hub, Josefstädter Strasse 39, A-1080 Vienna, Austria
[3] Computational Epidemiology Group, Robert Koch Institute, 13353 Berlin, Germany
[4] Department of Sports Medicine and Health Promotion, Friedrich-Schiller-University Jena, Jena, Germany
[5] Center for Sepsis Control and Care (CSCC), Jena University Hospital/Friedrich-Schiller-University Jena, Jena, Germany
[6] Center for Interdisciplinary Prevention of Diseases Related to Professional Activities, Jena, Germany
[7]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA
[8]Supply Chain Intelligence Institute Austria, Josefstädter Strasse 39, A-1080 Vienna, Austria
[9]Division of Insurance Medicine, Department of Clinical Neuroscience, Karolinska Institutet, SE-17177 Stockholm, Sweden
[10] Center Synergy of Systems & Center for Interdisciplinary Digital Sciences, TUD Dresden University of Technology, 01069 Dresden, Germany
[11] Faculty of Biology, TUD Dresden University of Technology, 01069 Dresden, Germany
* ledebur@csh.ac.at
† Lead authors of the manuscript.

Abstract

Evidence on long-term health effects of SARS-CoV-2 infection remains scarce. Using a wearable-derived behavioral and physiological dataset (n = 20,815), we identified 137 individuals characterized by self-reported persistent fatigue and shortness of breath following a SARS-CoV-2 infection. Compared with symptom-free COVID-19 positive and negative controls, these individuals had higher resting heart rates (mean difference 2.37/1.49 bpm) and lower daily step counts (mean 3,030/2,909 steps fewer), even at least three weeks *prior* to SARS-CoV-2 infection. In addition, persistent COVID-19 symptoms were associated with a significant reduction in mean quality of life (WHO-5, EQ-5D), even *before* infection. Analysis of individual wearable time series suggests that individuals with persistent symptoms may have been more exposed to pre-existing health conditions/lower fitness levels prior to SARS-CoV-2 infection. Our study demonstrates the potential of wearable devices to track physiological and physical activity dynamics under natural conditions in the context of infectious and chronic diseases.

Introduction

SARS-CoV-2 infections exhibit a diverse range of long-term impacts on individual health. The WHO defines post-COVID-19 condition (PCC) as *the continuation or development of new symptoms 3 months after the initial SARS-CoV-2 infection, with these symptoms lasting for at least 2 months with no other explanation* [1]. PCC is a multi-systemic condition [2] that has led to increased demand for health care due to increased disease burden [3]. Studies on persistent symptoms post-SARS-CoV-2 infection vary

widely in design, including differences in symptoms studied, follow-up duration, and study populations [4–7]. In addition, challenges in data availability - such as a lack of adequate control cohorts, insufficient pre-infection health data, a focus on severe cases, and a reliance on subjective self-reports or infrequent measurements - have limited our understanding of persistent symptoms post-SARS-CoV-2 and those affected by it [5, 8–12].

Data from wearable devices such as smartwatches or fitness trackers, offer a novel approach to understanding the long-term impact of SARS-CoV-2 by providing objective measurements of heart rate, physical activity, step count, and sleep duration [13–15]. Combining this data with individual symptom reports not only enables the detection of critical events (e.g infection onset) but also also provides a deeper understanding of how consequences of a SARS-CoV-2 infection affect individual physiology and behavior [16–18]. These datasets provide a unique opportunity to establish pre-infection health baselines through continuous, non-invasive, and cost-effective measurements.

During the COVID-19 pandemic, several initiatives were launched to collect such data [18–20]. One such initiative is the Corona Data Donation project [13, 21–23] by the Robert Koch Institute, Germany's federal agency for public health research. Between April 2020 and December 2022, over 500,000 voluntary participants registered for the project, obtained and activated the Corona Data Donation Application (CDA), and over 120,000 submitted more than 600 days of their basic wearable data (such as resting heart rate (RHR) and step count) using a custom-made smartphone application. In addition to daily averages, wearable data is available with exceptional resolution, down to the order of seconds. This permits a deep analysis of wearable-derived time-series, e.g. short high-intensity activities, regularities, and potential disruptions under natural conditions. Moreover, the high temporal resolution of the data and its coverage of long periods of time permit a detailed analysis of wearable data signals in all phases of a SARS-CoV-2 infection, i.e. the acute phase (from the week of the positive SARS-CoV-2 test to four weeks after the test), the sub-acute phase (four weeks to 12 weeks after the test), and the post-acute phase (12 weeks after the test date) [24]. One can clearly distinguish between individuals exhibiting weak regulation responses following infection (Fig. 1a,b) from those with strong ones (Fig. 1c,d). Participants also completed monthly and weekly surveys about COVID-19 tests, symptoms as well as their subjective wellbeing and health-related Quality of Life (QoL).

We leveraged the unique combination of wearable data and survey data from the Corona Data Donation project to analyze the physical health of individuals reporting persistent symptoms following a SARS-CoV-2 infection. We focused on cohort-level averages of physiological metrics (specifically RHR) behavioral (step count) metrics. To make sense of the observed changes in wearable data, we then compared these objective measures with self-reported QoL (WHO-5 and EQ-5D) to show how signals in passively measured health data translate directly into subjectively experienced changes or limitations in daily life.

Our study aims to investigate resting heart rate, step count, subjective symptoms report and self reported quality of life to analyse possible interactions with pre-existing health conditions before infection, time-course of infection and post-infection phases. To this end, we first identified a cohort of individuals that reported a positive SARS-CoV-2 test ($COVID-19[+]$) along with their negative controls ($COVID-19[-]$). Within $COVID-19[+]$ we compared those that reported long-lasting symptoms ($COVID-19[+]PS$) with those that did not ($COVID-19[+]NS$) to characterize the PCC phenotype.

Results

Increased frequency of reports on prolonged symptoms after infection

We evaluated the relative frequency of the reported symptoms of shortness of breath and fatigue in the *full cohort* (Fig.2) in pre- (<0 weeks to the reported COVID-19 test), acute- (0-4 weeks), sub-acute (5-12 weeks), and post-phase (>12 weeks) [24]. During the acute phase of infection, fatigue was reported as a symptom by up to 69.76% (SE = 0.54%) of individuals in $COVID-19[+]$ and

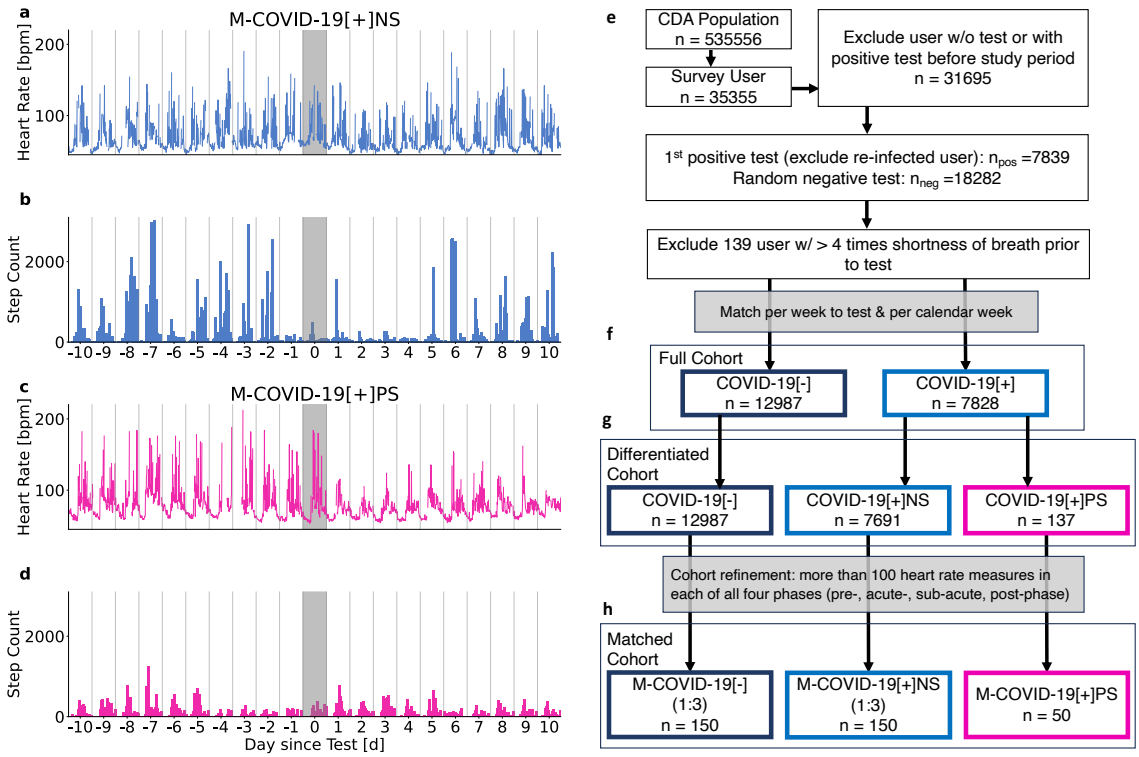


Fig 1. Time series of two representative individuals from the CDA population (a-d) and cohort diagram (e-h). Average daily heart rate [bpm] (a,c) and step counts (b,d) per 15 minutes relative to the day of the reported positive SARS-CoV-2 test (shaded grey area). Heart rate and step count for an individual without persistent symptoms 10 days before and after a reported positive test (a-b). Response in RHR and step count is more pronounced for an individual reporting persistent shortness of breath and fatigue (c-d). Both participants exhibited reduced step count a day prior to the test and in the following days, with the participant reporting persistent symptoms showing prolonged reduction. (e) Workflow of cohort creation. (f) Full cohort encompassing a positive cohort ($COVID-19[+]$) and negative control cohort ($COVID-19[-]$). (g) *Differentiated cohort* encompassing the persistent symptoms cohort ($COVID-19[+]PS$), the positive control cohort ($COVID-19[+]NS$), and the negative control cohort ($COVID-19[-]$). (h) *Matched cohort* consisting of the persistent symptoms match cohort ($M-COVID-19[+]PS$), the positive match cohort ($M-COVID-19[+]NS$) and the negative match cohort ($M-COVID-19[-]$).

by 19.53% (SE = 0.46%) of individuals in $COVID-19[-]$. Fatigue was reported significantly more often in $COVID-19[+]$ than in $COVID-19[-]$ for at least 6 weeks starting from the week of the reported SARS-CoV-2 test, Fig. 2a. Shortness of breath was reported by up to 13.55% (SE = 0.51%) and 2.46% (SE = 0.23%) by individuals in $COVID-19[+]$ and $COVID-19[-]$, respectively. It was increasingly reported for 15 weeks in $COVID-19[+]$ (Fig. 2b). The combination of both symptoms was reported by up to 12.2% (SE = 0.09%) and 1.84% (SE = 0.05%) by individuals in $COVID-19[+]$ and $COVID-19[-]$ and at increased frequency in $COVID-19[+]$ for a minimum of 11 weeks, see Fig. 2c.

Both, shortness of breath and fatigue, are potentially long lasting symptoms that could help identify individuals suffering from PCC [2, 25, 26]. Other assessed symptoms (cough, chills, loss of smell/taste, fever, diarrhea, limb ache and runny nose) did not significantly persist beyond the sub-acute phase when comparing $COVID-19[+]$ with $COVID-19[-]$ (Supplementary Figure 1).

Based on these findings, we defined a subset of $COVID-19[+]$ as those individuals who reported shortness of breath and fatigue on at least 5 occasions from the week of the reported positive SARS-CoV-2 test ($COVID-19[+]PS$). Individuals in $COVID-19[+]PS$ reported on average 6.2 (Std = 2.33, SE = 0.2) unique symptoms during the acute phase compared to 2.41 (Std = 2.39, SE = 0.01) symptoms reported by individuals with a positive SARS-CoV-2 test but no persistent symptoms

Table 1. Characterization of the study cohorts as defined in Fig. 1. For additional information on the reported health states per cohort, see Tab. S1.

	COVID-19 [+][1]	COVID-19 [-]	COVID-19 [+]NS[2]	COVID-19 [+]PS[3]	M-COVID-19 [-]	M-COVID-19 [+]NS[4]	M-COVID-19 [+]PS[5]
N	7,691	12,987	7,669	137	150	150	50
Vaccinated [%]	90.0***	88.3	90.0***	88.3	91.3	90.0	88.0
Not vaccinated [%]	0.4	0.5	0.4	0.7	0.0	0.0	2.0
Mean Age	52.4***	55.0	52.4***	52.7	53.6	53.6	53.6
(Std)	(11.9)	(12.0)	(11.9)	(11.7)	(10.5)	(10.5)	(10.5)
Female [%]	36.1**	34.1	35.9**	47.4**	50.0	50.0	50.0
Male [%]	54.4	54.7	54.7	40.1**	40.0	40.0	40.0
Unspecified [%]	9.8***	11.9	9.8***	11.7	10.0	10.0	10.0
Mean BMI	26.2	26.5	26.2	27.7*	26.9	25.5	27.4
(Std)	(4.7)	(5.1)	(4.7)	(5.7)	(6.2)	(4.0)	(4.6)

[1] $COVID-19[+]$ vs. $COVID-19[-]$; [2] $COVID-19[+]NS$ vs. $COVID-19[-]$; [3] $COVID-19[+]PS$ vs. $COVID-19[+]NS$; [4] $M-COVID-19[+]NS$ vs. $M-COVID-19[-]$; [5] $M-COVID-19[+]PS$ vs. $M-COVID-19[+]NS$. p-values: *p < 0.05, **p < 0.01, ***p < 0.001.

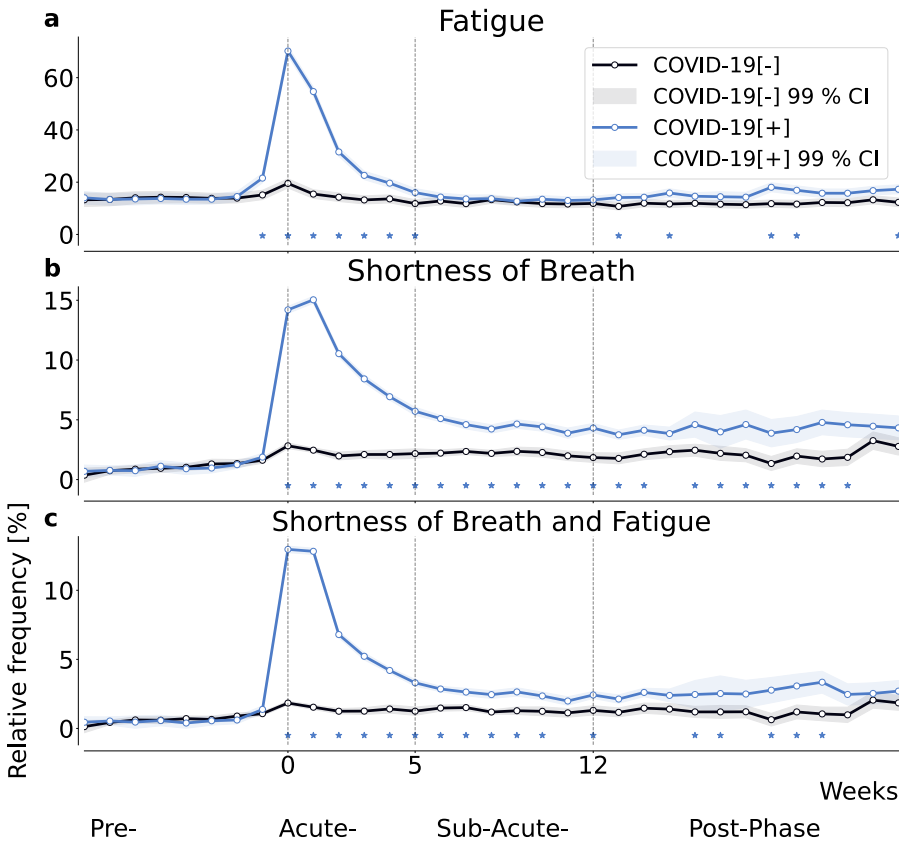


Fig 2. Relative frequency of self-reported fatigue (a), shortness of breath (b), and their combination (c), all relative to the week of the reported SARS-CoV-2 test for positive (P) individuals and a matched negative control (NC) cohort. Shading indicates the 99% confidence interval, i.e, 2.576 times the standard error of a binomial distribution. Asterisks indicate significant differences between the cohorts using a two-sided two proportion z-test with a significance level of 0.01.

($COVID-19[+]NS$) and 0.59 (Std = 1.14, SE = 0.00) symptoms reported by individuals in $COVID-19[-]$ (Supplementary Figure 2).

Supplementary Figure 3 illustrates relative frequencies of all symptoms in all phases, relative to the test date window for all cohorts, comprehensively.

Persistent symptoms correspond to more pronounced and prolonged RHR-changes

We analyzed wearable-derived data around the SARS-CoV-2 test (Fig.3) for the three age- and gender-matched (1:3) cohorts: $M-COVID-19[+]PS$, $M-COVID-19[+]NS$ and $M-COVID-19[-]$ (Fig.1h).

To assess differences in RHR, we computed the average RHR for each individual over the last seven days, obtaining one data point per day relative to the test date. We adjusted for seasonal differences by subtracting the daily mean RHR of the CDA population (see Supplementary Figure 4a for seasonal differences in RHR).

To standardize the data and ensure comparability between individuals, we calculated per-user z-scores by subtracting each user's average RHR and dividing by the standard deviation in the seven days prior to a reported SARS-CoV-2 test. This normalization centers all time series around zero, measured in units of standard deviation. We then averaged these Z-scores across all individuals within each cohort (see Methods for details).

We observed distinct RHR regulation patterns between the cohorts (Fig. 3 a). Individuals in $M-COVID-19[+]NS$ and $M-COVID-19[-]$ had comparable levels of mean RHR except during the acute phase.

The minimum Z-score during the acute phase for the $M-COVID-19[+]PS$ cohort was -0.86, compared to -0.45 for the $M-COVID-19[+]NS$ cohort, indicating that the average RHR in the $M-COVID-19[+]PS$ cohort was 0.86 standard deviations below the pre-phase mean, compared to 0.45 standard deviations for the $M-COVID-19[+]NS$ cohort. The range between the maximum and minimum deviation in the period 14 days before to 20 days after the test was more pronounced for the $M-COVID-19[+]PS$ cohort, with a factor difference of 1.3 in the corresponding z-scores. In addition, $M-COVID-19[+]PS$ individuals had a transient tachycardia followed by prolonged relative bradycardia, that did not return to baseline (z-score ≥ 0) until 18 days after infection. The relative bradycardia of $M-COVID-19[+]NS$ individuals persisted until 15 days after infection.

Notably, $M-COVID-19[+]PS$ individuals have on average an elevated RHR compared to the two control cohorts already *prior* to the reported SARS-CoV-2 test (inset in Fig.3a). Specifically, $M-COVID-19[+]PS$ individuals exhibited an average increase of 2.37 bpm and 1.49 (between 21 to 7 days to the test) compared to $M-COVID-19[+]NS$ and $M-COVID-19[-]$ individuals, respectively.

Physical activity profiles in acute and sub-acute infection regulation

To assess differences in physical activity patterns between individuals in the $M-COVID-19[+]PS$, $M-COVID-19[+]NS$, and $M-COVID-19[-]$ cohorts, we analyzed the average step count per user per day in the pre-, acute-, sub-acute and post-phase of the *match cohort* (Fig.3b). Seasonal differences in activity (i.e., higher and lower step count during summer and winter, respectively, see (Supplementary Figure 4b)), were accounted for by subtracting the average of all participants in the CDA population from the average steps per day, see Methods Section for details. Results are presented relative to the mean steps per day during pre-phase of the $M-COVID-19[-]$ cohort.

The respective median step count for $M-COVID-19[+]PS$-individuals was consistently below the median of both control cohorts in all four phases. Although the three cohorts are age- and sex-matched, the lower step counts in the $M-COVID-19[+]PS$ cohort may reflect reduced activity levels. The mean step count per day was also below the mean step count of the two control cohorts in all four phases. During the pre-phase, the mean step count of the $M-COVID-19[+]PS$ cohort was 3,030 steps less than that of the $M-COVID-19[+]NS$ cohort and 2,909 steps less than the $M-COVID-19[-]$ cohort. This indicates a generally lower level of activity compared to individuals from the $M-COVID-19[+]NS$ and $M-COVID-19[-]$ cohort. In addition, $M-COVID-19[+]PS$ individuals engaged at least three weeks prior to the infection in 1.12 and 0.99 days of high physical activity less than individuals from the $M-COVID-19[+]NS$ and $M-COVID-19[-]$ cohorts, respectively (Supplementary Figure 5).

Persistent symptoms relate to lower well-being and quality of life

To explore the relationship between wearable data characteristics and perceived well-being and QoL, we analyzed responses to an adapted version of the WHO-5 wellbeing index [27, 28] and a modified

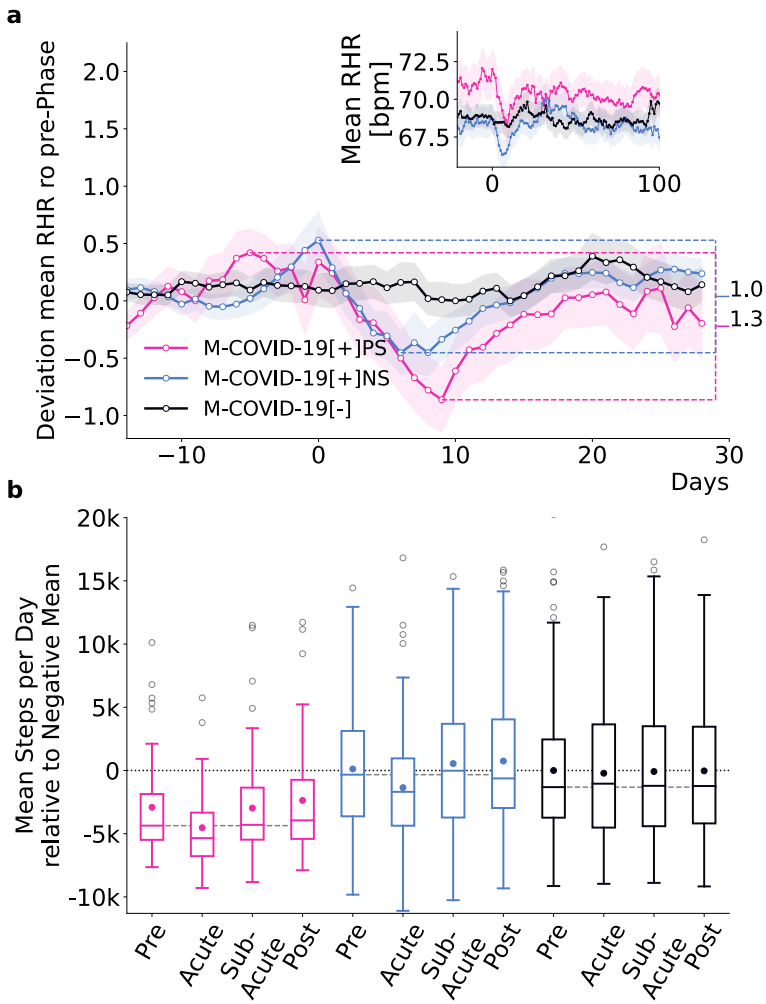


Fig 3. Wearable data analysis of the *match cohort*. (a) Z-transformed mean RHR (average of all 15-minute RHR measurements within the last seven days) relative to the seasonal mean RHR with respect to the mean and standard deviation up to seven days prior to the date of the reported test of all individuals in the $M-COVID-19[+]PS$ (pink), $M-COVID-19[+]NS$ (blue) and $M-COVID-19[-]$ (black) cohorts. The difference between the maximum and minimum z-transformed RHR within 14 to and 20 days after the date of the reported SARS-CoV-2 test was more pronounced (1.3 vs 1.0) and more prolonged for $M-COVID-19[+]PS$ than for $M-COVID-19[+]NS$. Shading indicates standard errors. The inset shows the average RHR relative to the SARS-CoV-2 test date. Already prior to the SARS-CoV-2 test, $M-COVID-19[+]PS$-individuals showed an increased RHR compared to $M-COVID-19[+]NS$ and $M-COVID-19[-]$. (b) Average steps per day relative to the mean of $M-COVID-19[-]$ during pre-phase (adjusted for seasonal variation) in all four phases for all individuals in all three cohorts. Boxes indicate quartiles, whiskers the range of the distribution of mean steps per day, scatter points outside the boxes mark the outliers (we do not show outliers $>$ 20k steps per day), and scatter points within the box mark the mean. The dashed grey line indicates the median of the mean steps per day during the pre-phase. Median values for $M-COVID-19[+]PS$ were consistently below the seasonal mean of the CDA population and below the median of the two control cohorts across all phases, indicating lower activity levels compared to the control cohorts. Mean number of steps per day for $M-COVID-19[+]PS$ were below the mean number of steps per day of the two control cohorts in all four phases. Likewise we found a reduction in the variance of the $M-COVID-19[+]PS$ compared to the two control cohorts.

EQ-5D health-related QoL survey (see SI for details) across the *differentiated cohort* (Fig. 1g), see Fig. 4. Due to the timing of the EQ-5D administration (post-2022), responses were not differentiated by pre- or post-SARS-CoV-2 test periods. By then, most participants had already reported a test (see Supplementary Figure 8).

Our version of the WHO-5 wellbeing index is constructed from five questions rated on a five-point scale (ranging from never (1) to always (5), with (3) is neutral; see SI for details). We examined the response for each individual question as well as the average response (Fig. 4 a, c-g). The overall distribution of wellbeing in $COVID-19[+]PS$ was skewed towards low values, (Fig. 4 a) with a mean value of 2.62 (Std = 0.55, SE = 0.05) for the $COVID-19[+]PS$ cohort and a mean of 3.22 (Std = 0.63, SE = 0.01) and 3.2 (Std = 0.69, SE = 0.01) for the $COVID-19[+]NS$ and $COVID-19[-]$ cohort, respectively. We also observed lower wellbeing scores for all individual outcomes (Fig. 4 c-g). Users in $COVID-19[+]PS$ reported particularly low scores for feeling energetic and rested upon waking up (Fig. 4d), possibly indicating activity and sleep issues. Statistically significant differences ($\alpha = 0.001$) between the $COVID-19[+]PS$ and the control cohorts were found in regards to all WHO-5 questions, as determined by the Mann-Whitney U test. Likewise, we found statistically significant differences between the reports of the two control cohorts except for the questions "In the past four weeks I felt fresh and rested when waking up" (d), "In the past four weeks I was calm and relaxed" (g), and the mean of wellbeing. We observed statistically significant differences between all three cohorts, nevertheless the significance level (p-value) was lower when comparing the $COVID-19[+]PS$ cohort to the control cohorts than for the intra-control cohort comparison.

In addition, we examined EQ-5D health-related QoL assessed by five questions answered on an ordinal scale ranging from the best ("I have no problems/No"; coded as 5) to the worst outcome ("I am not able to do that/extreme"; coded as 1) (Fig. 4b, h-l). The average response to the five questions, a measure of overall QoL, was systematically lower for individuals in $COVID-19[+]PS$ compared to the control cohorts (Fig. 4b) with a mean of 2.79 (Std = 0.62, SE = 0.06). Again, $COVID-19[+]NS$ and $COVID-19[-]$ show similar values across modified EQ-5D questions and overall with a mean of 3.55 (Std = 0.46, SE = 0.01) and 3.52 (Std = 0.52, SE = 0.01) for $COVID-19[+]NS$ and $COVID-19[-]$, respectively. Likewise, $COVID-19[+]PS$-individuals reported lower QoL across all five questions (Fig. 4h-l) compared to the control cohorts, particularly concerning problems moving around and pain and physical symptoms (Fig. 4 b and k), again possibly indicating the presence of comorbidities or pre-existing conditions.

The answers of the $COVID-19[+]PS$ cohort were significantly different from the control cohorts for all QoL questions ($\alpha = 0.001$). Notably, also the answers of the $COVID-19[+]NS$ compared to the $COVID-19[-]$ cohort were significantly different except for the the mean of QoL (b) and the question "Do you have problems moving around?" (h), and "Do you have problems when it comes to general activities (work, studying, housework, family- or leisure activities)?" (j). However, the significance level was again lower when comparing the $COVID-19[+]PS$ cohort to the two control cohorts than for the intra-control cohort comparison.

We also analyzed the responses to the wellbeing and QoL questions before (Supplementary Figure 6) and after (Supplementary Figure 7) the reported SARS-CoV-2 tests and found systematically lower values of wellbeing and QoL for the $COVID-19[+]PS$ cohort before as well as after the reported test compared to the $COVID-19[+]NS$ and $COVID-19[-]$ cohorts. This suggests that individuals with lower levels in wellbeing and QoL may be at higher risk of developing long-term symptoms, possibly due to pre-existing general physical or mental health problems that may be related to another underlying condition.

In addition, we performed simple logistic regression analysis to identify characteristics that might serve as predictors of the development of persistent symptoms. Given the imbalance between the $M-COVID-19[+]PS$ and $M-COVID-19[+]NS$ cohorts, we focused on the area under the precision-recall curve (AUPRC). Using 10-fold cross-validation with stratified sampling, the model achieved an

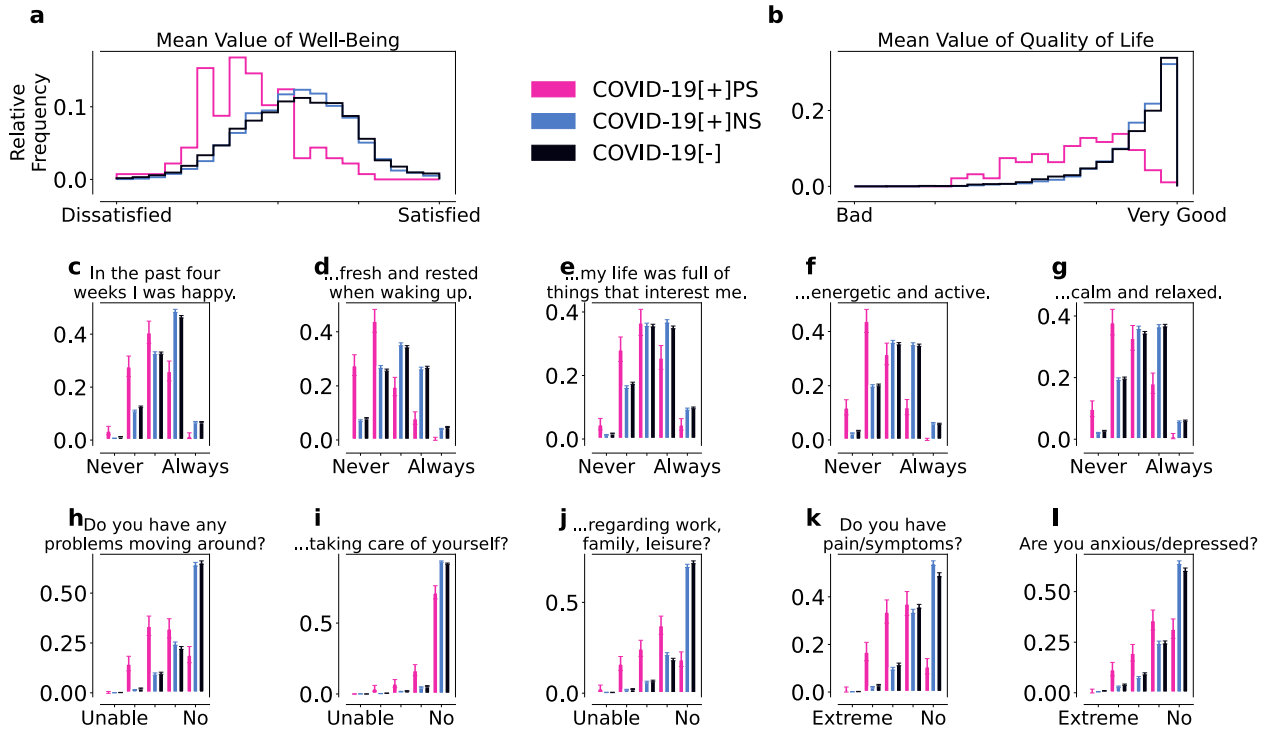


Fig 4. WHO-5 wellbeing (a,c-g) and modified EQ-5D/QoL (b,h-l) for the $COVID-19[+]PS$ (pink), $COVID-19[+]NS$ (blue), and $COVID-19[-]$ (black) cohorts. The individual WHO-5 and modified EQ-5D scores were both averaged to obtain the overall wellbeing (a) and QoL (b) scores, respectively. Overall, $COVID-19[+]PS$ individuals reported more issues with wellbeing (c-g) and QoL (h-l) than the control cohorts. Error bars indicate standard errors. The responses from all $COVID-19[+]PS$ cohort were significantly different ($\alpha < 0.001$) from the responses of the two control cohorts, as determined by the Mann-Whitney U test. All responses of the two control cohorts were significantly different as well except for the mean value of wellbeing (a), the mean value of QoL (b), and the questions "In the last four weeks I was calm and relaxed" (g) and "Do you have problems going around" (h).

AUPRC of 0.75, indicating moderate predictive performance. Feature importance analysis, based on mean coefficients and odds ratios across cross-validation folds, identified number of unique symptoms reported during the acute-phase of infection, chronic bronchitis, unspecified sex, and allergies as the most influential features, with mean odds ratios (OR) of 4.40, 1.65, 1.53, and 1.35, respectively. Mean pre-infection RHR also emerged as a notable predictor, with a mean OR of 1.31. These findings suggest that integrating wearable-derived data such as RHR with traditional clinical variables can enhance the identification of individuals at higher risk of developing persistent symptoms. This highlights the potential value of wearable technology in supplementing conventional approaches for early detection and management of PCC.

Table 2. Predictive features for developing long COVID based on logistic regression analysis. The table shows the mean coefficients, standard deviations, and mean odds ratios from 10-fold cross-validation.

Feature	Mean Coefficient	Std Coefficient	Mean Odds Ratio
Number of Symptoms	1.482	0.119	4.40
Chronic Bronchitis	0.498	0.089	1.65
Gender (Unspecified)	0.427	0.178	1.53
Allergies	0.298	0.135	1.35
Mean RHR Pre-Phase	0.273	0.149	1.31
Mental Health Condition	0.260	0.109	1.29
Age	0.142	0.152	1.15
Gender (Male)	0.058	0.157	1.06
Steps During Acute Phase	0.058	0.161	1.06
Asthma	0.053	0.103	1.05
Increased Blood Lipids/Cholesterol	-0.030	0.084	0.97
Other Conditions	-0.075	0.076	0.93
No Disease	-0.078	0.163	0.93
Hypertensive Disease	-0.226	0.132	0.80
Mean RHR Acute-Phase	-0.372	0.179	0.69
Gender (Female)	-0.485	0.123	0.62
Mean Steps per Day Pre-Infection	-0.750	0.143	0.47
WHO-5 Wellbeing Score	-1.038	0.115	0.35

Discussion

Our study comprehensively characterizes how persistent self-reported symptoms in acute and post-acute COVID-19 infection relate to changes in resting heart rate and physical activity measured by consumer grade wearable sensors. This observational study provides a descriptive analysis of the variation in heart rate, physical activity, and self-reported health status among different cohorts, and highlights patterns that may be linked to persistent symptoms following SARS-CoV-2 infection. Based on RHR, step count, and survey-based QoL, we identified patterns associated with a phenotype of persistent COVID-19 symptoms compared to COVID-19 positive and negative controls. Individuals with this phenotype are characterized by persistent (sub-acute and post-phase) shortness of breath and fatigue, and tended to have elevated RHR levels and less physical activity *prior* to infection, as well as, on average, lower well-being and health-related QoL scores compared with positive and negative controls.

We found notable differences in the prevalence of fatigue, shortness of breath, and their co-occurrence in SARS-CoV-2 positive and negative individuals. Both symptoms have been reported as typical, persistent symptoms of COVID-19 which are common in individuals with PCC[2, 5, 6, 29]. Indeed, according to a WHO case definition, fatigue and dyspnoea (shortness of breath) are the main symptoms of PCC [30]. Our analysis confirmed that these symptoms were significantly more frequent in the positive cohort than in the controls, with a marked persistence beyond the acute phase. Other symptoms did not show the same prolonged pattern, highlighting the specificity of these particular symptoms in relation to PCC.

A comparison of the wearable data (RHR and step count) revealed that during the acute phase (0-4 weeks after positive SARS-CoV-2 test) of COVID-19, individuals who reported persistent symptoms exhibited (1) higher RHR before infection, (2) on average more pronounced bradycardia up to

18 days after the positive SARS-CoV-2 test, and (3) lower physical activity before during and after infection compared to positive controls. Similar differences have been observed in previous studies when comparing SARS-CoV-2 positive and negative individuals [13, 14]. Individuals that later reported persistent symptoms already exhibited an elevated average RHR (mean increase of 2.37 bpm/1.49 bpm) compared to positive/negative control cohorts prior to their SARS-CoV-2 testing. In addition, these individuals were less likely to engage in high-intensity activities (i.e. regular physical activity and/or training) and were generally less active than the two control cohorts. Specifically, their daily step count was on average 3,030/2,909 steps less and they engaged in high activity 1.12/0.99 days less during the pre-phase than the positive/negative controls. Given the known inverse relationship between RHR and physical activity [31, 32], these patterns suggest that lower physical fitness levels, possibly due to pre-existing conditions, could increase susceptibility to developing long-term symptoms.

When comparing the $COVID-19[+]$ and $COVID-19[-]$ cohorts and the $COVID-19[+]NS$ and $COVID-19[-]$ cohorts, we found significantly higher reported vaccination rates in the positive cohorts, but no significant difference in the proportion reporting no vaccination. This suggests that the observed difference in vaccination rates may be due to reporting bias rather than actual differences in health status.

The positive cohorts ($COVID-19[+]$ and $COVID-19[+]NS$) were also generally younger than the negative cohort ($COVID-19[-]$). This age difference may lead to an underestimation of our results. There were also fewer individuals in the negative cohort who reported their gender, suggesting a potential reporting bias.

When comparing the $COVID-19[+]PS$ (persistent symptoms) and $COVID-19[+]NS$ (no persistent symptoms) cohorts, we observed a significantly higher proportion of women in the persistent symptom group, suggesting an association between female gender and the development of persistent symptoms, consistent with the existing literature [4, 33].

We observed a significantly higher BMI in the $COVID-19[+]PS$ cohort compared to the $COVID-19[+]NS$ cohort. While this suggests that individuals with persistent symptoms ($COVID-19[+]PS$) may have a higher BMI, it is important to note that the number of participants who provided both weight and height information was too small to reliably include BMI in the logistic regression analysis.

The $COVID-19[-]$ cohort reported fewer cases of "no disease in the past 12 months," indicating more comorbidities compared to the positive cohorts. They had higher rates of conditions such as hypertension, elevated cholesterol, chronic bronchitis, and mental health diagnoses. This suggests that the negative cohort may be less healthy overall, which may cause our results to represent a lower bound of the true effect, as the negative cohort may be less healthy overall.

When comparing the $COVID-19[+]PS$ cohort to the $COVID-19[+]NS$ cohort, the $COVID-19[+]PS$ group reported more allergies, asthma, chronic bronchitis, mental illness, and "other conditions". This suggests a higher burden of comorbidities among those with persistent symptoms. Similarly, in the matched cohorts ($M-COVID-19[+]PS$ vs. $M-COVID-19[+]NS$), the $M-COVID-19[+]PS$ cohort had higher rates of allergies, chronic bronchitis, and mental illness, supporting the notion that pre-existing health conditions may play a role in the development of persistent symptoms. This is in line with findings on the association of these conditions and the risk of developing PCC [34, 35].

It is important to note that our findings should not be interpreted in a way that increased RHR, lower step count, and pre-existing health conditions solely determine the manifestation of persistent symptoms. Instead they emphasize that individuals displaying these characteristics should be given extra attention and protection measures.

Our analysis of self-reported well-being and QoL showed significantly lower levels for individuals with persistent symptoms compared to the control cohorts, both before and after the reported SARS-CoV-2 test. This suggests the possibility of underlying physical or mental health problems, such as chronic conditions or allergies (see Supplemetary Table 1), that may predispose individuals to persistent symptoms. Specifically we find a mean difference of 0.6/0.58 in mean well-being score and a

mean difference of 0.76/0.73 between $COVID-19[+]PS$ and $COVID-19[+]NS/COVID-19[-]$ individuals. This is in line with recent findings, where individuals reporting persistent symptoms also reported poorer mental health in comparison to a control group [36], as well as a study reporting disturbed sleep as an important risk factor for PCC [37].

In addition to our descriptive study, we analyzed the importance of variables associated with the risk of developing persistent symptoms in a predictive model. We found that the number of symptoms during the acute infection phase, chronic bronchitis, unspecified sex, allergies, and mean pre-infection RHR were among the most influential factors. Wearable-derived data may thus complement previously identified risk factors for the development of persistent symptoms [38, 39].

While "unspecified gender" emerged as a risk factor, this likely reflects differences in reporting behavior rather than biological risk. Individuals who do not report gender may differ in unmeasured characteristics, such as privacy concerns or health-seeking behaviors, that affect symptom reporting or health outcomes. Thus, this finding should be interpreted with caution, as it may reflect data completeness issues or participant behavior rather than a direct risk for developing persistent symptoms.

The use of wearable data allows detailed examination of activity and heart rate patterns, providing insight not only into how SARS-CoV-2 infection may affect behavior and physiology in the long term, but also into the health status of individuals prior to contracting the disease. The data provides distinct advantages for evaluating the general health state of individuals holistically, continuous, under natural conditions, and outside of healthcare facilities [40, 41]. Wearables thus have the potential to improve patient-centered care by empowering individuals to monitor their health and manage their symptoms more effectively [42, 43]. In addition, data donation projects allow us to investigate individuals across a broad spectrum of SARS-CoV-2 infection severity, since participants enter the study prior to an infection. This approach enables us to establish a healthy state baseline, based on measurements prior to an infection, for individuals without biasing the study towards specific infection severities. In other words, participants are not selectively enrolled into the study based on their expected severity of subsequent SARS-CoV-2 infection symptoms or outcomes. This is of particular significance in the context of PCC research, since this condition is not exclusively linked to severe COVID-19 cases [44].

Several factors should be considered when interpreting our results. The Corona Data Donation Project had an overrepresentation of males and an underrepresentation of adolescents and the elderly (65+), who are at higher risk for long-term consequences of SARS-CoV-2 [13, 45]. The study population is likely to be more health-conscious than the general population because of the association of wearable device use with health-related habits [46]. We did not explicitly account for vaccination status, as most participants are partially vaccinated or did not report their status (see Table 1). This, along with other cohort-specific characteristics, highlights the need to interpret our findings in the context of our study population, which may not be fully generalizable to other groups. We also did not differentiate between variants of concern (VoC), which may trigger different immune responses and heart rate regulation. However, most participants reported their SARS-CoV-2 test date in 2022, when the Omicron variant (B.1.1.529) was dominant in Germany.

Keeping all these considerations in mind, it is reasonable to assume that our results under-represent the true disease burden of persistent symptoms and rather provide a lower bound for expected long-term imprints of an infection [36, 47–50]. We also did not control for wear time, but Supplementary Figure 9 shows similar distributions across cohorts.

The sample size of the study is limited by the requirement for sufficient wearable data (at least 100 heart rate measurements) in all four phases (pre, acute, sub-acute, and post) around a SARS-CoV-2 test. We also used strict criteria for defining persistent symptoms (at least five cases of shortness of breath and fatigue after the test), focusing on the most common symptoms observed in PCC [5–7, 29, 30]. While this focused approach inevitably introduces a degree of selection bias by targeting a specific subset of individuals with persistent symptoms, it is critical for accurately identifying those who are clearly affected and understanding their unique health profiles. We recognize that this targeted selection may limit the generalizability of our findings to broader populations. However, it allows for

a more precise examination of the physiological and behavioral characteristics unique to those with persistent symptoms, thereby increasing the relevance and specificity of our findings.

The two cohorts studied for symptom reporting are not matched by age and gender, only by calendar week and time relative to a SARS-CoV-2 test. The small differences in age (2.6 years) and gender (1.5%) distributions between these cohorts are unlikely to affect the results(Table 1). Previous studies have found gender differences in symptoms only after the reported SARS-CoV-2 test for symptoms not included in our study [5].

Further studies are needed on the use of wearables to monitor physiological parameters and activity levels over time in individuals with PCC. Particularly, focusing on other persistent symptoms apart from shortness of breath and fatigue can aid in recognizing different patterns of symptom progression and immune response triggers. Moreover, high-frequency heart rate and activity can potentially be relevant for studying postural orthostatic tachycardia syndrome (POTS), a potential consequence of an infection with SARS-CoV-2 and a likely criterion for identifying PCC [51, 52]. Further combination of survey and wearable data would provide a deeper understanding of the relationship between subjective symptom reports and objective wearable data, allowing physiological and behavioural measures to be properly translated into experienced alterations to patients' everyday lives. Further integration of wearable technology with telemedicine platforms facilitating the monitoring of symptoms in real-time and enabling remote consultation with healthcare professionals for individuals suffering from PCC can benefit patients and ease the burden on healthcare providers.

The unique advantage of wearable technologies is that continuous real-time data can be collected cheaply and on a large scale to provide fine-grained information on individual health, which is an ideal complement to traditional point-wise measurements in clinical research.

Methods

Data Characteristics

From April 12, 2020, to December 31, 2022, a total of 535,556 individuals downloaded and used the Corona-Datenspende App (CDA) [23], actively contributing a minimum of one essential data point each. The wearable data, i.e. RHR and step count, was submitted by linking the the app with consumer-grade smartwatches and fitness trackers. Participants could additionally consent to engage in periodic surveys on matters related to COVID-19, e.g., test results, weekly symptoms and monthly questions on QoL.

Weekly symptom reports were collected starting from October 21, 2021. By November 9, 2022, 35,355 individuals had participated in at least one survey. Participants could choose to report no symptoms or select one or more symptoms from a predefined list: Shortness of Breath, Fatigue, Headache, Sore Throat, Loss of Smell and Taste, Diarrhea, Runny Nose, Cough, Chills, Limb Ache, and Fever.

Comparisons were made between individuals who reported a positive SARS-CoV-2 test ($COVID-19[+] = 7{,}691$ individuals) and a negative control cohort ($COVID-19[-] = 12{,}987$ individuals) (see Methods and Fig.1e). Participants in the $COVID-19[+]$ cohort reported symptoms an average of 18.4 times (standard deviation (Std) = 11.2). Participants in the NC cohort submitted reports 11.1 times on average (Std = 11.4). The $COVID-19[+]$ cohort was further split in sub-cohorts, one with individuals that reported persistent symptoms ($COVID-19[+]PS = 137$ individuals) and the rest ($COVID-19[+]NS = 7669$ individuals), which together with the $COVID-19[-]$ cohort yielding in three cohorts, see Methods and Fig.1e) for details on cohort creation.

For these cohorts we analysed survey responses on wellbeing and QoL. The questions on WHO-5 well-being were distributed since October 2021 and rolled out together with the symptom surveys. 35,492 individuals submitted corresponding answers. Particularly, all 137 individuals reporting persistent symptoms, 7,609 individuals from the $COVID-19[+]NS$ cohort and 12,866 users from the $COVID-19[-]$ cohort answered the survey on well-being at least once. The WHO-5 wellbeing index

[27, 28] is answered on a six-point scale ranging from never(0) to always (5). For technical reasons the CDA uses a slightly altered version that utilizes a five-point scale ranging from 1 to 5 with 3 being neutral. The five questions ask: 'In the past four weeks I was happy and in a good mood.', '... I felt calm and relaxed.', '... I felt energetic and active.', '... I felt fresh and rested when waking up.', and '... my everyday life was full of things that interest me.'

The modified EQ-5D survey on QoL was distributed since April 2022 to 12,179 participants. 92 individuals with persistent symptoms, 2,647 from the $COVID{-}19[+]NS$ cohort and 12,866 individuals from the $COVID-19[-]$ cohort answered the five questions on QoL a least once. The survey asks to indicate what best describes the state of health today: 'Do you have any problems moving around?', 'Do you have any problems taking care of yourself, washing or dressing yourself?', 'Do you have problems when it comes to general activities (work, studying, housework, family- or leisure activities)?', 'Do you have pain/physical symptoms?', and 'Are you anxious or depressed?'. The answers are on an ordinal scale ranging from the best possible outcome ("I have no problems/No"; coded as 5) to the worst possible outcome ("I am not able to do that/extreme"; coded as 1). For three age- and gender matched cohorts (1:3) denoted *persistent symptoms match* ($M-COVID-19[+]PS = 50$), *positive match* ($M{-}COVID{-}19[+]NS = 150$) and *negative match* ($M{-}COVID{-}19[-] = 150$), we analyzed high frequency longitudinal data on RHR and step count (see Methods and Fig. 1e). Data on RHR and step count were collected in intervals as low as 60 seconds. We show four exemplary timeseries spanning three weeks each in Fig.1a-d.

Cohort creation

To study reported symptoms relative to a reported positive SARS-CoV-2 test ($COVID-19[+]PS$) and compared them with a negative control ($COVID-19[-]$) cohort we looked at 35,355 individuals who filled out the corresponding surveys. We excluded 3,660 individuals who reported a positive SARS-CoV-2 test in a previously distributed one-time questionnaire to avoid previous infections with SARS-CoV-2 in the population. We divided the population into two groups: those with at least one positive SARS-CoV-2 test and those with all negative tests. Both cohorts are limited to the period between October 21, 2021, and November 9, 2022. For individuals with positive tests, we used the date of their first positive test if multiple tests were positive. To avoid including individuals with re-infections, we excluded those reporting a test more than 28 days after the chosen test date, resulting in 7839 individuals. Among the 18,282 individuals reporting only negative tests, we randomly selected a test date. We also excluded users who reported shortness of breath more than four times before their test result, as a proxy for pre-existing respiratory conditions. To account for seasonal effects in symptom reporting, we adjusted the negative cohort to match the relative frequency of negative reports per calendar week and the temporal distance to the test week observed in the positive cohort. This yields a positive cohort ($COVID-19[+]NS$) of 7,691 individuals and a negative control ($COVID-19[-]$) of 12,987 individuals (full cohort, Fig. 1f).

For the analysis of WHO-5 and modified EQ-5D, we selected individuals from $COVID{-}19[+]$ who reported both shortness of breath and fatigue at least five times after their positive SARS-CoV-2 test. This subset, comprising 137 individuals, forms the persistent symptoms cohort ($COVID-19[+]PS$). We compared their survey responses with those of the remaining 7,669 individuals in the Positive Cohort (referred to as the positive control cohort ($COVID-19[+]NS$)) and the 12,987 individuals from the negative control ($COVID-19[-]$) cohort (differentiated cohort Fig. 1g).

To study differences and changes in physiological parameters (RHR) and behavior (step count) between the $COVID-19[+]NS$, $COVID-19[-]$, and $COVID-19[+]PS$ cohorts, we analyzed fine-grained wearable data collected by wearable devices. We standardized the data by aggregating high-frequency measurements into 15-minute intervals. Four phases relative to the reported SARS-CoV-2 test are defined: pre-phase (start of data collection until the week of the positive test), acute phase (zero to four weeks after the test), sub-acute phase (five to twelve weeks), and post-phase (twelve weeks or more after the positive test), aligning with recent definitions of COVID-19 disease stages [24]. After excluding individuals with less than 100 15-minute interval heart rate measurements across all phases, the persistent symptoms match ($M-COVID-19[+]PS$) cohort comprises 50 individuals. We created (1:3) age- and gender-matched positive match ($M-COVID-19[+]NS$) and negative

match ($M-COVID-19[-]$) cohorts, each with 150 individuals selected from the $COVID-19[+]NS$ and $COVID-19[-]$ cohorts, respectively (match cohort, Fig. 1h). We chose a 1:3 matching, common in medical research, to increase statistical power and to improve the generalizability of our findings by ensuring that the control group more accurately represents the population from which the cases and controls were drawn.

For characteristics of the cohorts see Tabs. 1 and S1.

Data processing and statistical analysis

Statistical analysis of group differences.
To analyze group differences in our study cohorts, we used different statistical tests tailored to the nature of the data. Continuous variables, such as age and BMI, were analyzed using the Kruskal-Wallis test. Categorical variables, including sex, vaccination status, and comorbidities, were assessed using Fisher's exact test.

We performed these comparisons in three main cohort analyses: In the *full cohort* analysis, we compared the positive and negative cohorts ($COVID-19[+]$ vs. $COVID-19[-]$). In the *differentiated cohort* analysis, we compared the positive without persistent symptoms cohort and the negative cohort ($COVID-19[+]NS$ vs. $COVID-19[-]$) as well as the positive with persistent symptoms cohort and the positive without persistent symptoms cohort ($COVID-19[+]PS$ vs. $COVID-19[+]NS$). In the *matched cohort* analysis, we compared differences between the matched positive without persistent symptoms cohort and the matched negative cohort ($M-COVID-19[+]NS$ vs. $M-COVID-19[-]$) and the matched positive with persistent symptoms cohort and the matched positive without persistent symptoms cohort ($M-COVID-19[+]PS$ vs. $M-COVID-19[+]NS$).

For all statistical tests, significance levels were set at three thresholds: $\alpha = 0.05$, $\alpha = 0.01$, and $\alpha = 0.001$.

Weekly reports on symptoms.
For the $COVID-19[+]$ and the $COVID-19[-]$ cohorts, we compared the relative frequency of symptom reports relative to the calendar week in which a SARS-CoV-2 test was reported. We detected differences between the two groups at 99% confidence using a two-sided two proportion z-test and a significance level of 0.01. To investigate whether pairs of symptoms are significantly concordant, we compared the relative frequency of reports where individuals report both symptoms.

Resting heart rate and step count.
To ensure data consistency, we only analyzed one source/device per participant (e.g., Apple Health, Fitbit, or Garmin). If a user has multiple devices, we prioritized the one with the most data points, potentially leading to missing information if different devices are used for different activities. Observations are based on measurement intervals ranging from less than one minute to several hours, resulting in overlapping intervals in a few instances and a slight overestimation of daily activity. This is rare compared to the total data volume. In our data pre-processing approach, intervals longer than one minute were split into one-minute segments and duplicate intervals are removed. These one-minute intervals were then resampled to 15-minute intervals by calculating the mean of the RHRs and the sum of the steps. In cases where a user contributes data from multiple sources, only the source with the highest number of entries is considered.

To compare RHRs between the $M-COVID-19[+]PS$ cohort, the $M-COVID-19[+]NS$ cohort and the $M-COVID-19[-]$ cohort, for each user and each day we first computed (relative to the reported SARS-CoV-2 test) an average of all 15 minute intervals within the past seven days. We only computed the average if there were more than 20 entries within that seven-day window. We then computed the average RHR over all users in a cohort, again relative to the reported SARS-CoV-2 test. To evaluate tachycardia during the acute phase of the infection, we z-transformed the average of all 15 minute intervals within the last seven days per user relative to the mean RHR prior to the test (window of 62 days to 7 days before the date of the reported test). We additionally controlled for seasonal fluctuations by z-transforming the average RHR relative to the seasonal mean of the CDA population (Supplementary Figure 4a).

To examine differences in behavioral changes around a SARS-CoV-2 infection between cohorts, we investigated the average step count for the cohort members. As the daily step count shows seasonal

variations [13], we always assessed this metric by subtracting the respective mean value per day and device within the entire set of participants of the whole Corona Data Donation project, see also Supplementary Figure 4b. We show the difference to the mean daily step count (adjusted for seasonal variation) of the $M - COVID - 19[-]$ cohort during pre-phase. A positive/negative value then indicates that an individual is more/less active than the $M - COVID - 19[-]$ cohort during pre-phase average. To assess variations or shifts in high-intensity physical activity, such as sports, we established a criterion where a day is considered to be high-intensity physical activity if the step count is more than one standard deviation above the seasonal average step count.

Monthly surveys on well-being and quality of life.

For the analysis of the surveys on well-being/QoL we calculated the mean response to each question per user over all answers (not discriminating between before and after day of reported SARS-CoV-2 test) and showed the relative frequency of each possible answer value per cohort. Error bars indicate the standard error. The answers to each one of the five WHO-5 Well-Being questions are designed so that they can be averaged to obtain a mean value of well-being. Note, that the WHO-5 well-being is usually measured on a six-point scale. Due to technical reasons however, our survey provides an ordinal five-point scale. This also applies to the five modified EQ-5D QoL questions. The modification is attributed to different translations of questions and answers, along with one missing question. For both the WHO-5 and the modified EQ-5D, we show a histogram of the distribution of the respective mean scores for all three cohorts studied.

Logistic Regression.

We performed logistic regression analysis to identify variables associated with the risk of developing persistent symptoms after SARS-CoV-2 infection. To ensure robustness, we used 10-fold stratified cross-validation to assess model performance and calculate the area under the precision-recall curve (AUPRC). Feature importance was assessed by calculating mean coefficients, standard deviations and odds ratios across cross-validation folds.

Ethical consideration

Participation in the study was voluntary and self-recruited. All individuals participating in the Corona Data Donation Project provided informed consent electronically via the app. Consent was provided separately for submitting wearable data and participating in the in-app surveys. Participation is only possible for German residents age 16 and older and data is only stored pseudonymously, using a randomly generated unique user ID. Participant age is rounded to 5 years. The study is subject to strict compliance with the data protection provisions set out in the EU General Data Protection Regulation (GDPR) and the Federal Data Protection Act (BDSG). A comprehensive privacy impact assessment was conducted through an external law-firm specialized in e-Health and research projects. The study was reviewed and approved by the Data Privacy Officer at the Robert Koch Institute (internal operation number 2021-009) in agreement with the Federal Commissioner for Data Protection and Freedom of Information (BfDI), Germany's highest independent supreme federal authority for data protection and freedom of information. Ethical approval for this study was obtained from the ethics board at the University of Erfurt (approval number 20220414).

Data availability

The data analyzed in this study concerns the health condition of individual persons. As per the General Data Protection Regulation (GDPR) such data constitutes sensitive information and is protected as a "special category of personal data". For this reason, the data must not be directly shared in a public repository, but interested parties can request access to the data following registration with the Data Privacy Officer at the Robert Koch-Institute. All initial data inquiries should be addressed to Dirk Brockmann (`dirk.brockmann@tu-dresden.de`) or the general contact address of the Corona Data Donation project (`synosys@tu-dresden.de`).

Code availability

Python 3.11.6 was used to perform the simulations and data analysis. The simulation and analysis code for this study will be available under MIT license in the repository at `https://github.com/katharinaledebur/characteristics_persistent_symptoms.git`.

Acknowlegdements

This study was financed with funds from the overall funding program of the City of Vienna MA7. Funding was also received from the Federal Ministry of Health of Germany (Grants "Corona-Datenspende", CD21, DS22 and DS23).

Author contributions

MW and KL conceptualized the study. MW supervised the project. DB devised and created the Corona Data Donation project. MW and KL devised the analytic methods. KL carried out the analysis and produced the plots and graphics. KL and MW directly accessed and verified the underlying data reported in the manuscript. KL and MW wrote the first draft of the manuscript. CP, DB, ST and PK made critical comments regarding the manuscript. CP contributed medical expertise regarding the medical interpretation of the findings and in developing medical hypotheses. KL, MW, CP, PK, DB and ST conducted reviewing and editing of the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

References

1. World Health Organization. *Post COVID-19 condition (Long COVID)* https://www.who.int/europe/news-room/fact-sheets/item/post-covid-19-condition#:~:text=It%20is%20defined%20as%20the,months%20with%20no%20other%20explanation.
2. Davis, H. E., McCorkell, L., Vogel, J. M. & Topol, E. J. Long COVID: major findings, mechanisms and recommendations. en. *Nature Reviews Microbiology* **21,** 133–146. ISSN: 1740-1526, 1740-1534. https://www.nature.com/articles/s41579-022-00846-2 (Mar. 2023).
3. Menges, D. *et al.* Burden of post-COVID-19 syndrome and implications for healthcare service planning: A population-based cohort study. en. *PLOS ONE* **16** (ed Simuunza, M. C.) e0254523. ISSN: 1932-6203. https://dx.plos.org/10.1371/journal.pone.0254523 (July 2021).
4. Radin, J. M. *et al.* Long-term changes in wearable sensor data in people with and without Long Covid. *npj digital medicine* (2024).
5. Ballering, A. V., Van Zon, S. K. R., Olde Hartman, T. C. & Rosmalen, J. G. M. Persistence of somatic symptoms after COVID-19 in the Netherlands: an observational cohort study. en. *The Lancet* **400,** 452–461. ISSN: 01406736. https://linkinghub.elsevier.com/retrieve/pii/S0140673622012144 (Aug. 2022).
6. Nasserie, T., Hittle, M. & Goodman, S. N. Assessment of the Frequency and Variety of Persistent Symptoms Among Patients With COVID-19: A Systematic Review. en. *JAMA Network Open* **4,** e2111417. ISSN: 2574-3805. https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2780376 (May 2021).
7. Davis, H. E. *et al.* Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. en. *eClinicalMedicine* **38,** 101019. ISSN: 25895370. https://linkinghub.elsevier.com/retrieve/pii/S2589537021002996 (2023) (Aug. 2021).
8. Marjenberg, Z. *et al.* Risk of long COVID main symptoms after SARS-CoV-2 infection: a systematic review and meta-analysis. *Scientific Reports* **13,** 15332 (2023).
9. Høeg, T. B., Ladhani, S. & Prasad, V. How methodological pitfalls have created widespread misunderstanding about long COVID. *BMJ Evidence-Based Medicine* **29,** 142–146 (2023).
10. O'Mahoney, L. *et al.* The prevalence and long-term health effects of Long Covid among hospitalised and non-hospitalised populations: a systematic review and meta-analysis. *eClinicalMedicine* **55,** 101762 (2023).
11. Koc, H. C., Xiao, J., Liu, W., Li, Y. & Chen, G. Long COVID and its Management. en. *International Journal of Biological Sciences* **18,** 4768–4780. ISSN: 1449-2288. https://www.ijbs.com/v18p4768.htm (2022).
12. Han, Q., Zheng, B., Daines, L. & Sheikh, A. Long-Term Sequelae of COVID-19: A Systematic Review and Meta-Analysis of One-Year Follow-Up Studies on Post-COVID Symptoms. *Pathogens* **11,** 269 (2022).
13. Wiedermann, M. *et al.* Evidence for positive long- and short-term effects of vaccinations against COVID-19 in wearable sensor metrics. en. *PNAS Nexus* **2** (ed Levine, B.) pgad223. ISSN: 2752-6542. https://academic.oup.com/pnasnexus/article/doi/10.1093/pnasnexus/pgad223/7225910 (2023) (July 2023).
14. Radin, J. M. *et al.* Assessment of Prolonged Physiological and Behavioral Changes Associated With COVID-19 Infection. en. *JAMA Network Open* **4,** e2115959. ISSN: 2574-3805. https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2781687 (July 2021).
15. Natarajan, A., Su, H.-W. & Heneghan, C. Occurrence of Relative Bradycardia and Relative Tachycardia in Individuals Diagnosed With COVID-19. en. *Frontiers in Physiology* **13,** 898251. ISSN: 1664-042X. https://www.frontiersin.org/articles/10.3389/fphys.2022.898251/full (May 2022).

16. Mishra, T. *et al.* Pre-symptomatic detection of COVID-19 from smartwatch data. en. *Nature Biomedical Engineering* **4,** 1208–1220. ISSN: 2157-846X. https://www.nature.com/articles/s41551-020-00640-6 (2023) (Nov. 2020).

17. Gadaleta, M. *et al.* Passive detection of COVID-19 with wearable sensors and explainable machine learning algorithms. en. *npj Digital Medicine* **4,** 166. ISSN: 2398-6352. https://www.nature.com/articles/s41746-021-00533-1 (Dec. 2021).

18. Quer, G. *et al.* Wearable sensor data and self-reported symptoms for COVID-19 detection. en. *Nature Medicine* **27,** 73–77. ISSN: 1078-8956, 1546-170X. https://www.nature.com/articles/s41591-020-1123-x (Jan. 2021).

19. Varsavsky, T. *et al.* Detecting COVID-19 infection hotspots in England using large-scale self-reported data from a mobile application: a prospective, observational study. en. *The Lancet Public Health* **6,** e21–e29. ISSN: 24682667. https://linkinghub.elsevier.com/retrieve/pii/S2468266720302693 (Jan. 2021).

20. Menni, C. *et al.* Vaccine side-effects and SARS-CoV-2 infection after vaccination in users of the COVID Symptom Study app in the UK: a prospective observational study. en. *The Lancet Infectious Diseases* **21,** 939–949. ISSN: 14733099. https://linkinghub.elsevier.com/retrieve/pii/S1473309921002243 (July 2021).

21. Robert-Koch Institute. *Corona Data Donation Project* https://corona-datenspende.github.io/en/. Accessed: 2024-09-16.

22. Kolb, J. J. *et al.* Prevalence of Positive COVID-19 Test Results Collected by Digital Self-report in the US and Germany. *JAMA Network Open* **6(1):e2253800** (2023).

23. Robert Koch Institute and Research on Complex Systems of Humboldt University of Berlin. *Corona Data Donation* https://corona-datenspende.github.io/en/.

24. Nalbandian, A. *et al.* Post-acute COVID-19 syndrome. en. *Nature Medicine* **27,** 601–615. ISSN: 1078-8956, 1546-170X. https://www.nature.com/articles/s41591-021-01283-z (2023) (Apr. 2021).

25. Michelen, M. *et al.* Characterising long COVID: a living systematic review. en. *BMJ Global Health* **6,** e005427. ISSN: 2059-7908. https://gh.bmj.com/lookup/doi/10.1136/bmjgh-2021-005427 (2023) (Sept. 2021).

26. Canas, L. S. *et al.* Profiling post-COVID-19 condition across different variants of SARS-CoV-2: a prospective longitudinal study in unvaccinated wild-type, unvaccinated alpha-variant, and vaccinated delta-variant populations. en. *The Lancet Digital Health* **5,** e421–e434. ISSN: 25897500. https://linkinghub.elsevier.com/retrieve/pii/S2589750023000560 (2023) (July 2023).

27. Topp, C. W., Østergaard, S. D., Søndergaard, S. & Bech, P. The WHO-5 Well-Being Index: A Systematic Review of the Literature. en. *Psychotherapy and Psychosomatics* **84,** 167–176. ISSN: 0033-3190, 1423-0348. https://www.karger.com/Article/FullText/376585 (2023) (2015).

28. WHO (World Health Organisation). *WHO-5 Well-Being Index* https://www.corc.uk.net/outcome-experience-measures/the-world-health-organisation-five-well-being-index-who-5/ (2023).

29. Nguyen, N. N. *et al.* Clinical patterns of somatic symptoms in patients suffering from post-acute long COVID: a systematic review. en. *European Journal of Clinical Microbiology & Infectious Diseases* **41,** 515–545. ISSN: 0934-9723, 1435-4373. https://link.springer.com/10.1007/s10096-022-04417-4 (2023) (Apr. 2022).

30. Soriano, J. B., Murthy, S., Marshall, J. C., Relan, P. & Diaz, J. V. A clinical case definition of post-COVID-19 condition by a Delphi consensus. en. *The Lancet Infectious Diseases* **22,** e102–e107. ISSN: 14733099. https://linkinghub.elsevier.com/retrieve/pii/S1473309921007039 (2023) (Apr. 2022).

31. Reimers, A. K., Knapp, G. & Reimers, C.-D. Effects of Exercise on the Resting Heart Rate: A Systematic Review and Meta-Analysis of Interventional Studies. *Journal of Clinical Medicine* **7.** ISSN: 2077-0383. https://www.mdpi.com/2077-0383/7/12/503 (2018).

32. Huang, G., Shi, X., Davis-Brezette, J. A. & Osness, W. H. Resting heart rate changes after endurance training in older adults: a meta-analysis. *Medicine and science in sports and exercise* **37,** 1381–1386. ISSN: 0195-9131. https://journals.lww.com/acsm-msse/Fulltext/2005/08000/Resting_Heart_Rate_Changes_after_Endurance.18.aspx (2005).

33. Bai, F. *et al.* Female gender is associated with long COVID syndrome: a prospective cohort study. *Clinical Microbiology and Infection* **28,** 611.e9–611.e16. ISSN: 1198-743X. https://www.sciencedirect.com/science/article/pii/S1198743X21006297 (2022).

34. Philip, K. *et al.* Impact of COVID-19 on people with asthma: a mixed methods analysis from a UK wide survey. *BMJ Open Resp Res* **9(1),** e001056 (2022).

35. Wang, S. *et al.* Associations of Depression, Anxiety, Worry, Perceived Stress, and Loneliness Prior to Infection With Risk of Post–COVID-19 Conditions. *JAMA Psychiatry* **79,** 1081–1091 (2022).

36. Atchison, C. J. *et al.* Long-term health impacts of COVID-19 among 242,712 adults in England. en. *Nature Communications* **14,** 6588. ISSN: 2041-1723. https://www.nature.com/articles/s41467-023-41879-2 (2023) (Oct. 2023).

37. Schilling, C. *et al.* Pre-existing sleep problems as a predictor of post-acute sequelae of COVID-19. *Journal of Sleep Research* **33,** e13949 (2024).

38. Antony, B. *et al.* Predictive models of long COVID. en. *eBioMedicine* **96,** 104777. ISSN: 23523964. https://linkinghub.elsevier.com/retrieve/pii/S2352396423003432 (2024) (Oct. 2023).

39. Zang, C. *et al.* Identification of risk factors of Long COVID and predictive modeling in the RECOVER EHR cohorts. en. *Communications Medicine* **4,** 130. ISSN: 2730-664X. https://www.nature.com/articles/s43856-024-00549-0 (2024) (July 2024).

40. Yetisen, A. K., Martinez-Hurtado, J. L., Ünal, B., Khademhosseini, A. & Butt, H. Wearables in Medicine. en. *Advanced Materials* **30,** 1706910. ISSN: 1521-4095. https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201706910 (2018).

41. Adams, J. L. *et al.* A real-world study of wearable sensors in Parkinson's disease. en. *npj Parkinson's Disease* **7.** Number: 1 Publisher: Nature Publishing Group, 1–8. ISSN: 2373-8057. https://www.nature.com/articles/s41531-021-00248-w (2024) (Nov. 2021).

42. Khondakar, K. R. & Kaushik, A. Role of Wearable Sensing Technology to Manage Long COVID. en. *Biosensors* **13,** 62. ISSN: 2079-6374. https://www.mdpi.com/2079-6374/13/1/62 (2023) (Dec. 2022).

43. Burki, T. Wearable technology and COVID-19. en. *The Lancet Respiratory Medicine* **10,** 934–935. ISSN: 22132600. https://linkinghub.elsevier.com/retrieve/pii/S2213260022003514 (2023) (Oct. 2022).

44. Crook, H., Raza, S., Nowell, J., Young, M. & Edison, P. Long covid—mechanisms, risk factors, and management. en. *BMJ* **374,** n1648. ISSN: 1756-1833. https://www.bmj.com/lookup/doi/10.1136/bmj.n1648 (2023) (July 2021).

45. Carvalho-Schneider, C. *et al.* Follow-up of adults with noncritical COVID-19 two months after symptom onset. en. *Clinical Microbiology and Infection* **27,** 258–263. ISSN: 1198743X. https://linkinghub.elsevier.com/retrieve/pii/S1198743X20306066 (2023) (Feb. 2021).

46. Wu, Q., Sum, K. & Nathan-Roberts, D. How Fitness Trackers Facilitate Health Behavior Change. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* **60,** 1068–1072 (Sept. 2016).

47. Richard, S. *et al.* Persistent COVID-19 Symptoms at 6 Months After Onset and the Role of Vaccination Before or After SARS-CoV-2 Infection. *JAMA Network Open* **6(1),** e2251360 (2023).

48. Hernández-Aceituno, A., García-Hernández, A. & Larumbe-Zabala, E. COVID-19 long-term sequelae: Omicron versus Alpha and Delta variants. *Infectious Diseases Now* **53,** 104688 (2023).

49. Marra, A. R. *et al.* The effectiveness of COVID-19 vaccine in the prevention of post-COVID conditions: a systematic literature review and meta-analysis of the latest research. en. *Antimicrobial Stewardship & Healthcare Epidemiology* **3,** e168. ISSN: 2732-494X. https://www.cambridge.org/core/product/identifier/S2732494X23004473/type/journal_article (2023) (2023).

50. Nealon, J. & Cowling, B. J. Omicron severity: milder but not mild. en. *The Lancet* **399,** 412–413. ISSN: 01406736. https://linkinghub.elsevier.com/retrieve/pii/S0140673622000563 (2023) (Jan. 2022).

51. Raj, S. A. *et al.* Long-COVID postural tachycardia syndrome: an American Autonomic Society statement. *Clin Auton Res* **31,** 365–368 (2021).

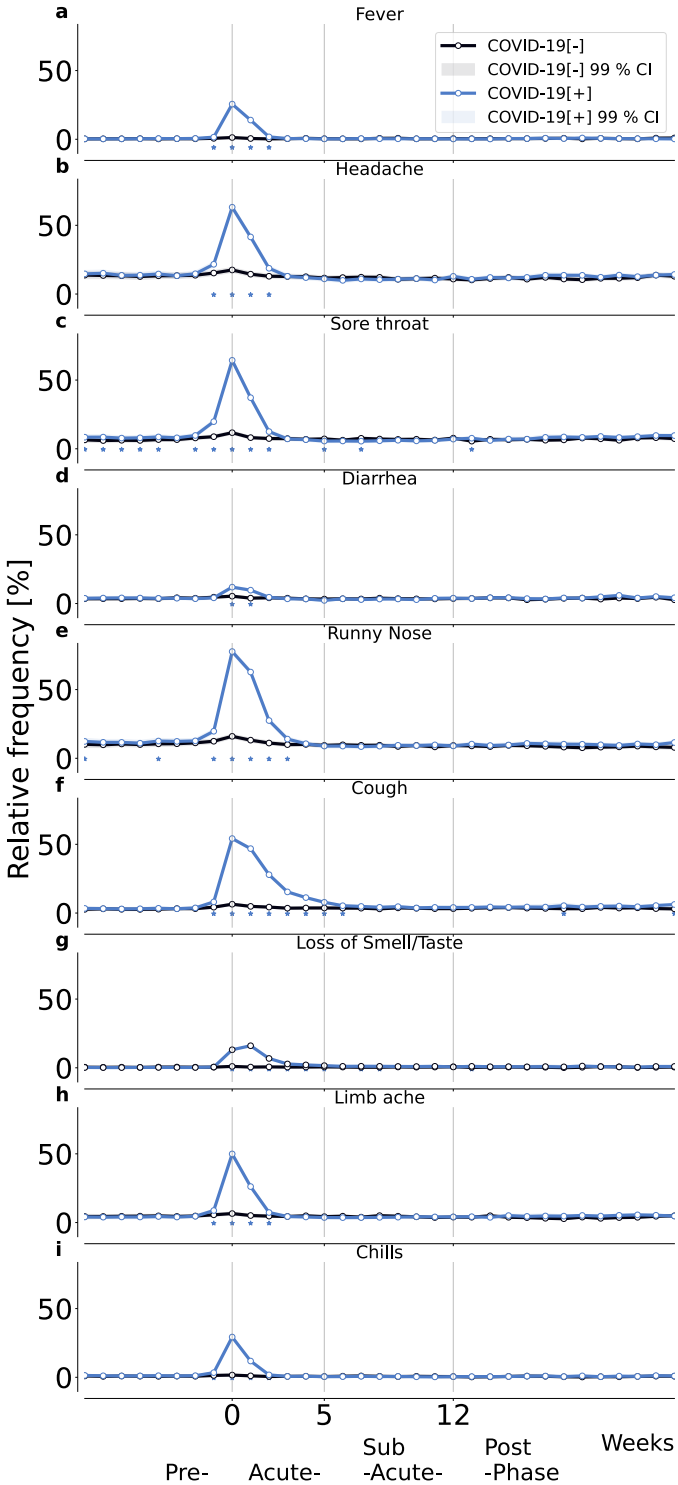
52. Ormiston, C. K., Świątkiewicz, I. & Taub, P. R. Postural orthostatic tachycardia syndrome as a sequela of COVID-19. *Heart Rhythm* **19,** 1880–1889. ISSN: 1547-5271. https://www.sciencedirect.com/science/article/pii/S1547527122021853 (2022).

SUPPLEMENTARY INFORMATION: WEARABLE DATA REVEALS DISTINCT CHARACTERISTICS OF INDIVIDUALS WITH PERSISTENT SYMPTOMS AFTER A SARS-COV-2 INFECTION

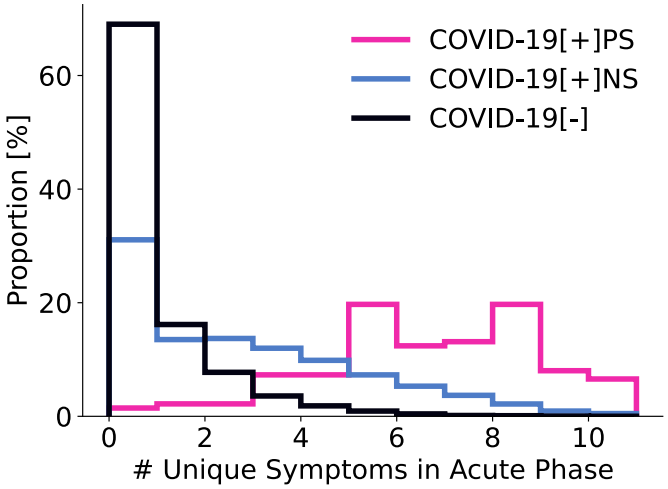
Supplementary Table I. Information on other diseases and conditions of all studied cohorts. See Materials and Methods (Section IV) for further details on the specifics of the assignments.

	COVID-19 [+][1]	**COVID-19 [-]**	**COVID-19 [+]NS**[2]	**COVID-19 [+]PS**[3]	**M-COVID-19 [+]NS**[4]	**M-COVID-19 [-]**	**M-COVID-19 [+]PS**[5]
N	7,828	12,987	7,691	137	150	150	50
% answered	89.9***	87.7	89.9***	88.3	90.0	89.3	90.0
No disease in the past 12 months [%]	48.5***	44.3	49.0***	24.8***	54.1	41.0	31.1*
Allergies [%]	30.4	30.0	30.0	50.4***	25.2	32.1	46.7**
Hypertension [%]	18.2***	22.6	18.1***	23.1	16.3	25.4	20.0
Asthma [%]	9.3	9.3	9.0	24.0***	11.1	11.9	22.2
Increased blood lipid/ cholesterol levels [%]	10.4***	12.8	10.3***	15.7	8.1	12.7	17.8
Chronic Bronchitis [%]	2.1***	3.1	2.0***	6.6**	1.5	1.5	8.9*
Other [%]	10.5***	13.9	10.3***	19.8*	10.4	9.7	22.2
Mental Illness Diagnose [%]	19.5**	22.2	19.2††	39.7***	15.6	20.1	44.4***

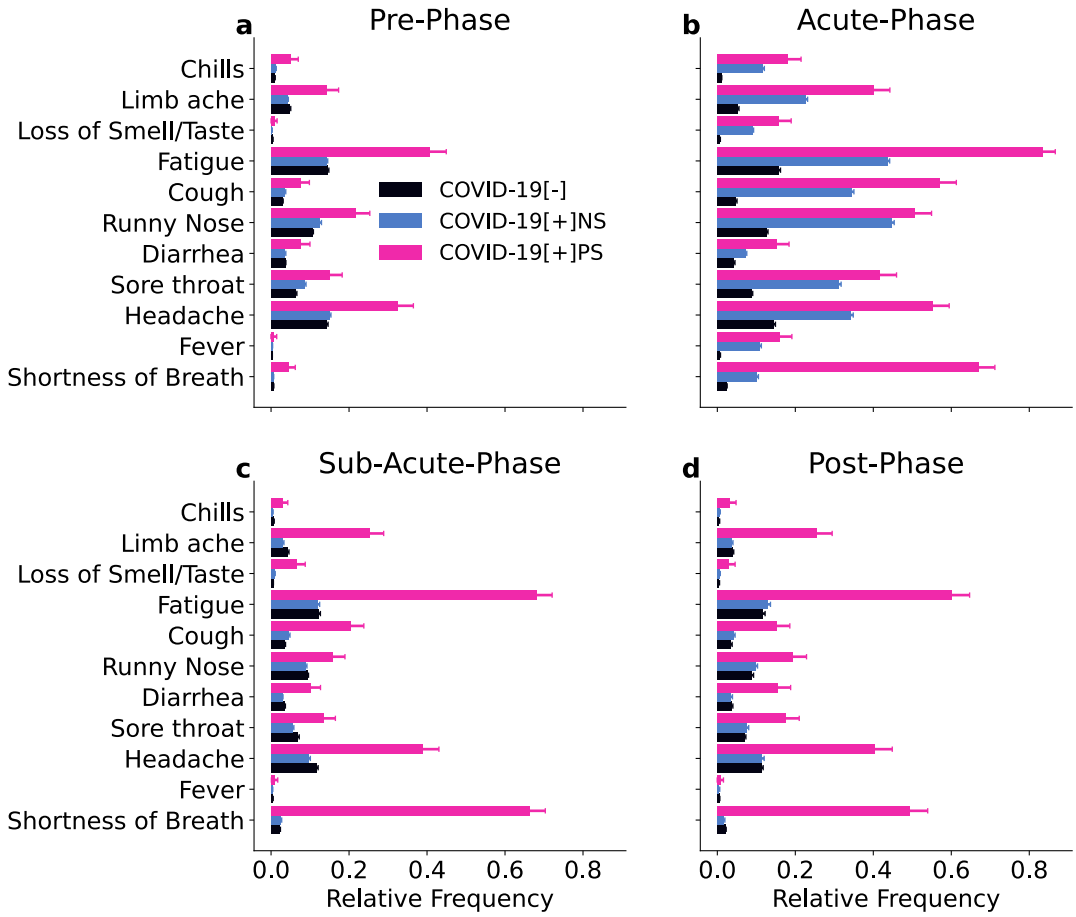
[1] $COVID-19[+]$ vs. $COVID-19[-]$; [2] $COVID-19[+]NS$ vs. $COVID-19[-]$; [3] $COVID-19[+]PS$ vs. $COVID-19[+]NS$; [4] $M-COVID-19[+]NS$ vs. $M-COVID-19[-]$;[5] $M-COVID-19[+]PS$ vs. $M-COVID-19[+]NS$. p-values: *p ¡ 0.05, **p ¡ 0.01, ***p ¡ 0.001.



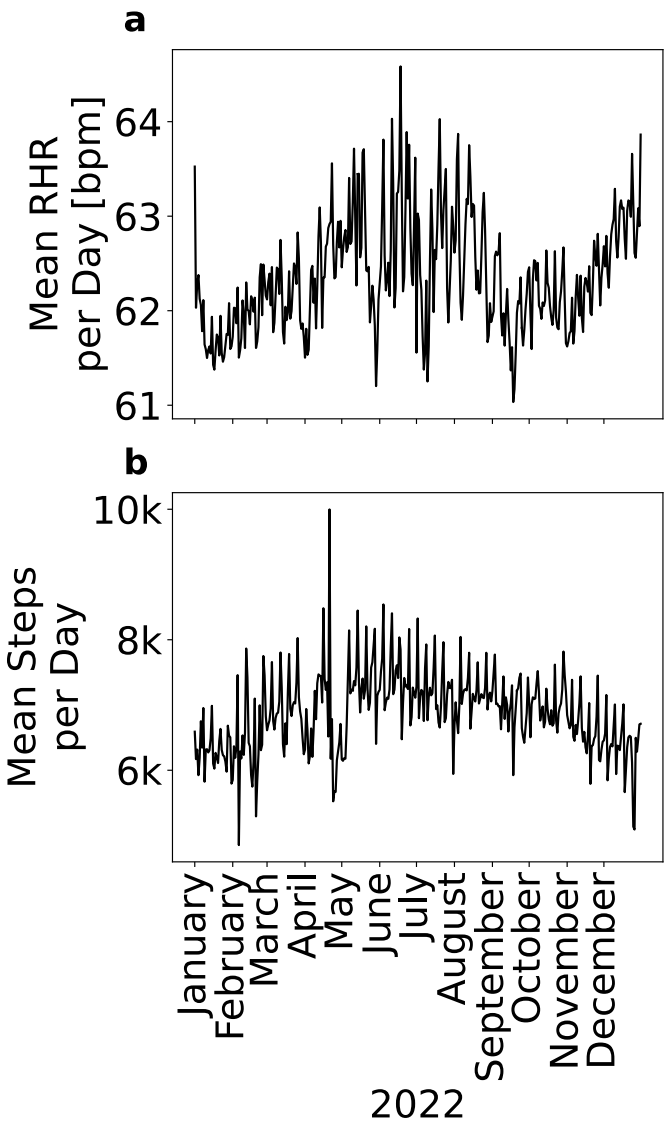
Supplementary Figure 1. Relative frequency of symptoms relative to week of reported SARS-CoV-2 test for positive cohort in blue ($COVID-19[+]$) and negative control cohort ($COVID-19[-]$) in black. The symptoms are fever (a), headache (b), sore throat (c), diarrhea (d), runny nose (e), cough (f), loss of smell/taste (g), limb ache (h), and chills (i). Shading indicates the 99% confidence interval, i.e, 2.576 times the standard error of a binomial distribution. Asterisks indicate significant differences between the cohorts using a two-sided two proportion z-test with a significance level of 0.01.



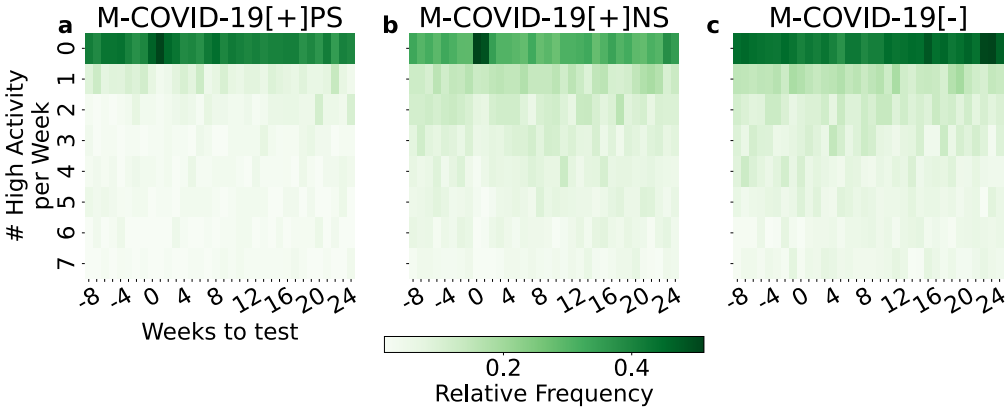
Supplementary Figure 2. Distribution of number of unique reported symptoms during acute phase of the infection (0-4 weeks to the reported SARS-CoV-2 test). We show the proportion of individuals per number of unique symptoms during the acute-phase of the infection for the persistent symptoms in pink ($COVID-19[+]PS$), positive control in blue ($COVID-19[+]NS$) and negative control cohort in black ($COVID-19[-]$).



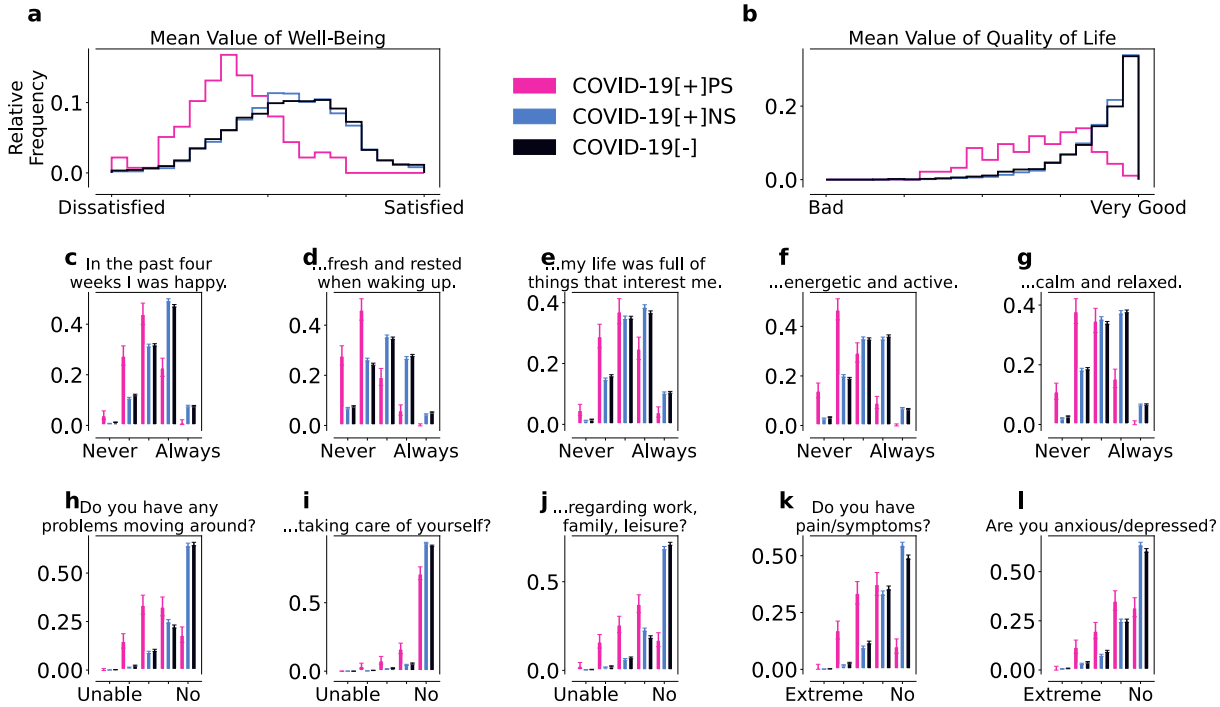
Supplementary Figure 3. Relative frequency per symptom. For pre- (a), acute- (b), sub-acute (c) and post (d) phase we show the relative frequency per reported symptom for all three groups in the matched cohort: negative match in black ($M - COVID - 19[-]$), positive match in blue ($M - COVID - 19[+]NS$) and persistent symptoms match in pink ($M - COVID - 19[+]PS$).



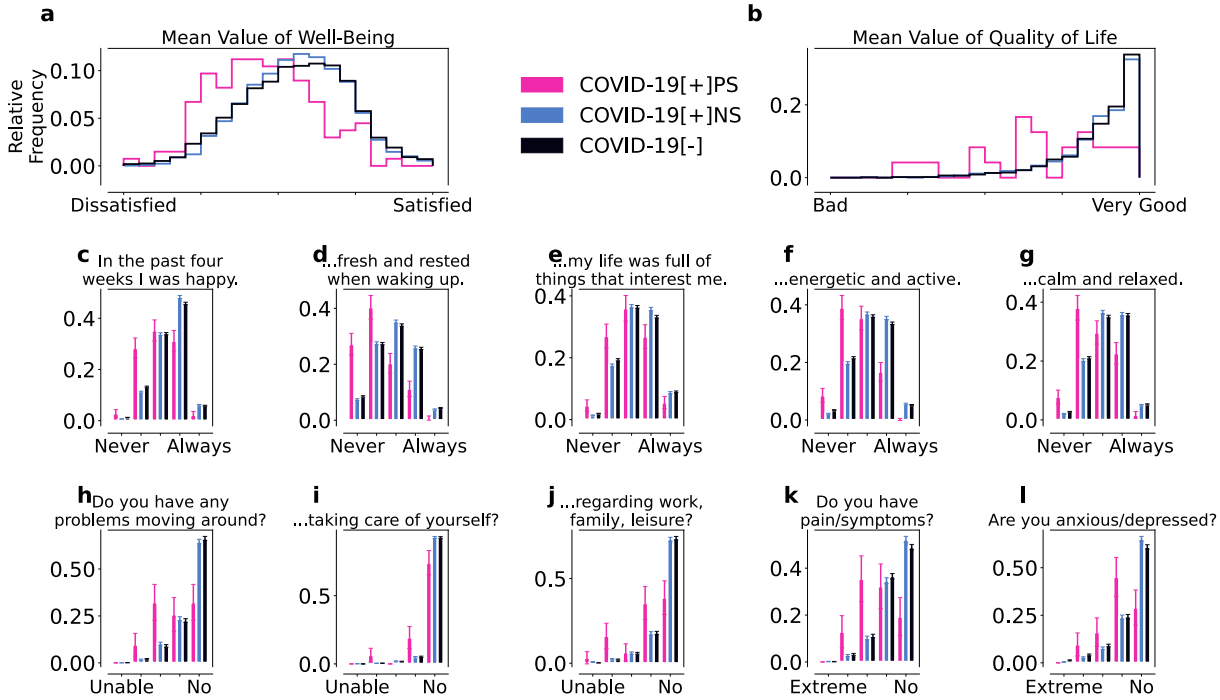
Supplementary Figure 4. Seasonal trends in wearable data. Exemplary timeseries of mean RHR per day (a) and mean steps per day (b) over the CDA Population for one example year (2022) and one example donation-source (Apple).



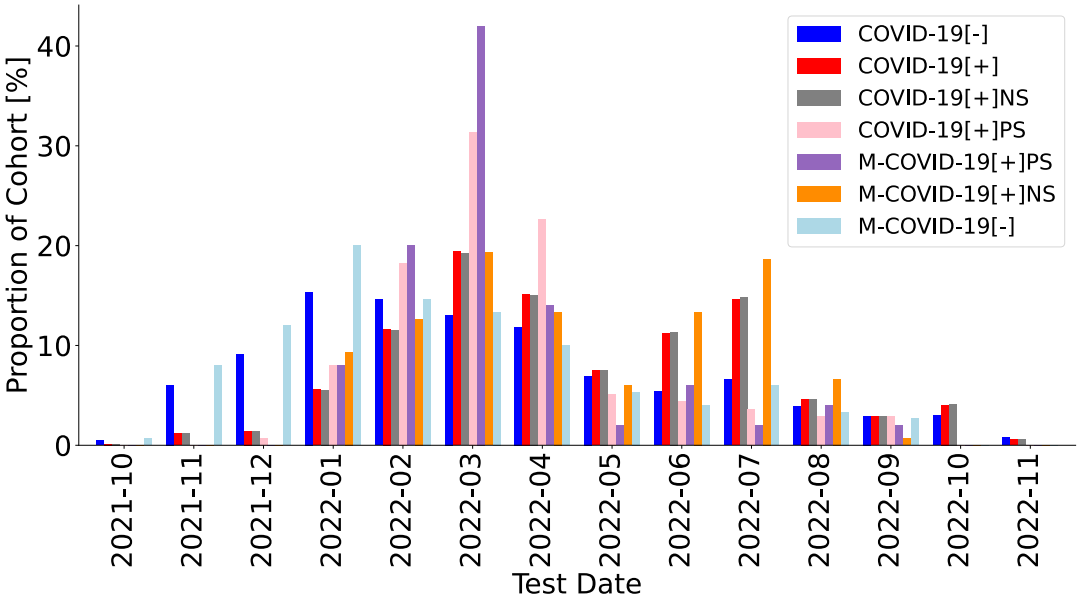
Supplementary Figure 5. High activity profiles of the *matched cohort*. Relative frequency of the number of high activity days per week relative to the SARS-CoV-2 test date for $M-COVID-19[+]PS$ (a), $M-COVID-19[+]NS$ (b) and $M-COVID-19[-]$ (c). A day is defined as a high activity day when the corresponding number of steps per day exceed one standard deviation over the seasonal mean taken over the whole set of users in the Corona Data Donation Project (CDA Population).



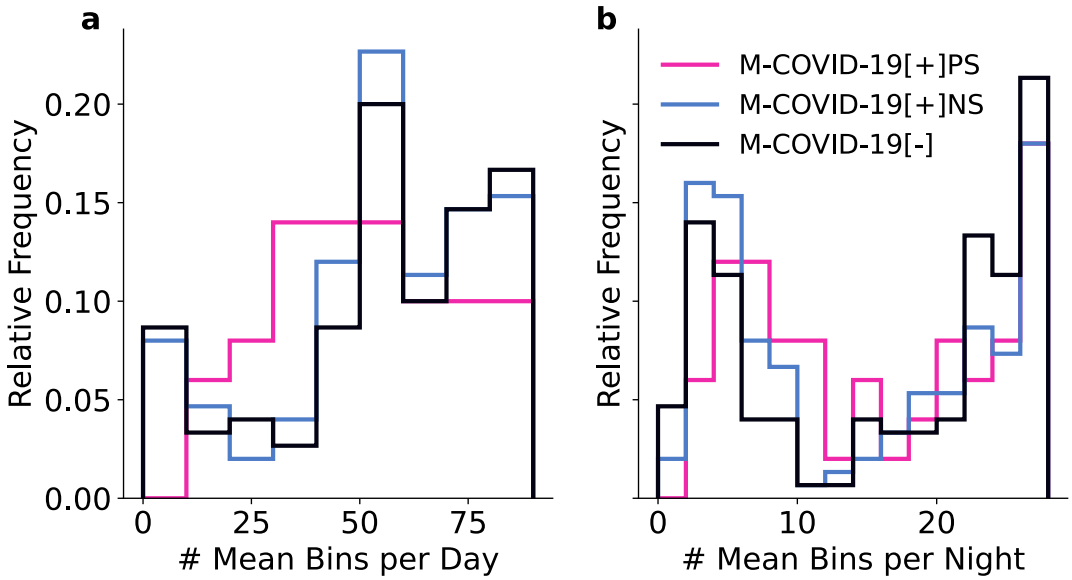
Supplementary Figure 6. Answers to WHO-5 wellbeing and EQ-5D questionnaires before the reported SARS-CoV-2 test for the persistent symptoms (pink), positive control (blue) and negative control cohort (black). WHO-5 wellbeing (a,c-g) and modified EQ-5D/QoL (b,h-l) for the $COVID-19[+]PS$ (pink), $COVID-19[+]NS$ (blue), and $COVID-19[-]$ (black) cohorts before the reported SARS-CoV-2 test. The individual WHO-5 and modified EQ-5D scores were both averaged to obtain the overall wellbeing (a) and QoL (b) scores, respectively. Overall, $COVID-19[+]PS$ individuals reported more issues with wellbeing (c-g) and QoL (h-l) than the control cohorts. Error bars indicate standard errors.



Supplementary Figure 7. Answers to WHO-5 wellbeing and EQ-5D questionnaires after the reported SARS-CoV-2 test for the persistent symptoms (pink), positive control (blue) and negative control cohort (black). WHO-5 wellbeing (a,c-g) and modified EQ-5D/QoL (b,h-l) for the $COVID-19[+]PS$ (pink), $COVID-19[+]NS$ (blue), and $COVID-19[-]$ (black) cohorts after the reported SARS-CoV-2 test. The individual WHO-5 and modified EQ-5D scores were both averaged to obtain the overall wellbeing (a) and QoL (b) scores, respectively. Overall, $COVID-19[+]PS$ individuals reported more issues with wellbeing (c-g) and QoL (h-l) than the control cohorts. Error bars indicate standard errors.



Supplementary Figure 8. Distribution of reported SARS-CoV-2 testdates of studied cohorts. We show per month ranging from October 2021 to November 2022 the proportion of individuals who indicated their SARS-CoV-2 test in the respective month for the COVID-19[-] (negative, black), COVID-19[+] (positive, red), COVID-19[+]NS (positive control, grey), COVID-19[+]PS (persistent symptoms, pink), M-COVID-19[+]PS (persistent symptoms match, blue), M-COVID-19[+]NS (positive match, orange), and M-COVID-19[-] (negative match cohort, light black).



Supplementary Figure 9. Distribution of time-worn of wearable devices. Number of 15 minute heart rate measurements as a proxy for wearing time for the three age and sex matched cohorts M-COVID-19[+]PS (persistent symptoms match, blue), M-COVID-19[+]NS (positive match, pink), and M-COVID-19[-] (negative match cohort, black) in 24 hours (a) and during night time from 11pm until 6am (b).