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

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

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

Abstract

Background: Evidence based findings on long-term health-related consequences of a SARS-CoV-2 infection remain scarce. Data from wearable devices, well suited for continuous measurement of heart rate and physical activity, offers a unique opportunity to assess the impact of such infections on an individual's health. Here we aim to characterize comprehensively how persistent self-reported symptoms during both acute and post-acute infection correlate to changes in resting heart rate (RHR) and physical activity, as measured by consumer-grade wearable sensors.
Methods: Using a wearable-derived dataset of behavior and physiology (n = 20,815), we identified 137 individuals who are characterized by persistent fatigue and shortness of breath after a reported positive SARS-CoV-2 test. We compared this cohort with COVID-19 positive without persistent symptoms and negative controls. The comparison is based on measurements of RHR and physical activity as well as self-reported health-related Quality of Life (QoL) through WHO-5 and EQ-5D before, during, and after the infection.
Findings: We identified a unique phenotype of persistent COVID-19 symptoms and associated wearable data characteristics and compared this phenotype to COVID-19 positive and negative controls. Individuals who reported persistent symptoms (coexisting shortness of breath and fatigue) showed higher RHRs (mean difference of $2 \cdot 37/1 \cdot 49$ bpm), and lower daily step count (on average 3,030/2,909 steps less) compared to positive/negative controls, even at least three weeks prior to a SARS-CoV-2 infection. During the acute phase (0-4 weeks after a positive COVID-19 test), individuals with persistent shortness of breath and fatigue exhibited a decrease in mean RHR, $1 \cdot 86$ times that of individuals in the positive control cohort. Similarly, the persistent symptom phenotype took an average of seven days longer to return to normal compared to positive controls.

Additionally we found that self-reported persistent COVID-19 symptoms are linked to a substantial reduction in mean QoL, even before infection.
Interpretation: The analysis of individual wearable time-series suggests that the persistent symptom phenotype, characterized by shortness of breath and fatigue, may have been more exposed to pre-existing health conditions and/or exhibited lower levels of fitness prior to a SARS-CoV-2 infection. Our approach demonstrates the enormous potential in tracking the dynamics of physiological and physical activity under natural conditions in the context of infectious and chronic diseases.
Funding: This study was funded in part by 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).

Research in Context

Evidence before this study: Previous research on persistent symptoms of the post-COVID-19 condition on heart rate and physical activity (measured in step count) often lacks a valid control group and/or information on the health status of individuals prior to the SARS-CoV-2 infection. The majority of studies have been conducted in clinical settings with a potential selection bias and do not account for post-COVID-19 conditions in the general population or its imprint on everyday life. Furthermore, knowledge on how lingering symptoms affect objectively measurable vital signals (such as heart rate and step count) in different phases of acute and post-acute infection regulation, is useful for enabling timely and targeted treatment interventions in clinical monitoring.
Added value of this study: By incorporating detailed data obtained from wearables prior to, during, and after infection, including symptoms, overall wellbeing, and pre-existing health conditions, we could effectively identify and thoroughly characterize individuals with persistent COVID-19 symptoms. This unique advantage of our approach enhances the interpretation of phases in post-acute infection regulation. It further facilitates a comprehensive analysis of both, perceived and physiological health status, providing a multifaceted view of post-COVID-19 condition. Additionally, the characterization of patient demographics, comorbidities, and Quality of Life (QoL) enriches our understanding of the population at risk for developing persistent symptoms.
Implications of the entire available evidence: We found that individuals experiencing persistent shortness of breath and fatigue, previously identified as core symptoms of post-COVID-19 condition, exhibit on average elevated resting heart rate (RHR), lower daily activity levels, lower QoL, and a higher count of pre-existing conditions already prior to an infection with SARS-CoV-2 compared to two control cohorts. In addition, the average decrease in RHR (bradycardia) during the acute phase of the infection was more pronounced and prolonged in those with persistent symptoms compared to the controls. These findings have helped to identify individuals at risk of developing persistent symptoms following SARS-CoV-2 infection and potentially assist tailoring diagnosis and treatment at an individual level.

1 Introduction

SARS-CoV-2 infections exhibit a diverse range of long-term impacts on individual health. The WHO defines post-COVID-19 condition 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]. Post-acute sequelae of a SARS-CoV-2 infection (PASC) or Long COVID [2, 3] is a multi-systemic condition [4] that has led to increased demand for health care due to increased disease burden [5]. Multiple retrospective cohort studies report increased risk of PASC such as gastrointestinal diseases, cardiovascular diseases, diabetes, neurological disorders, and diseases affecting the respiratory system [6, 7, 8, 9, 10, 11, 12]. Studies on persistent symptoms post-SARS-CoV-2 infection vary widely in design, including differences in symptoms studied, follow-up duration, and study populations [13, 14, 15]. Challenges in data availability result in a lack of control adequate cohorts, insufficient information on pre-infection health status, focus on severe cases only, and typically rely exclusively on subjective self-reported data, or infrequent measurements [16, 17, 18, 19, 13, 20]. Consequently, diagnostic and treatment guidelines for long-term impacts of SARS-CoV-2 infection still remain poorly defined.

Data from wearable devices such as smartwatches or fitness trackers, offer insights into individual infection regulation and the long-term impact of SARS-CoV-2 on personal health, including heart rate, activity, step count, and sleep duration [21, 22, 23]. Combining this data with individual symptom reports not only enables us to detect critical events (e.g infection onset) [24] but also also provides a deeper understanding of how consequences of a SARS-CoV-2 infection affect individual physiology and behavior [24, 25, 26]. During the course of the COVID-19 pandemic, several initiatives were launched to collect such type of data [27, 26, 28]. One such initiative is the Corona Data Donation project [21, 29, 30] 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. Next to daily averages, wearable data is available with exceptional resolution, down to the order of seconds. This permits a deep analysis of vital 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 vital signals in all phases of a SARS-CoV-2 infection (Fig. 1a-d), 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) [2]. One can clearly distinguish between individuals exhibiting strong regulation responses following infection (Fig. 1a,b) from those with weak 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). Based on this, we discriminated our study cohort into individuals that reported long-lasting symptoms (such as shortness of breath and fatigue) after a SARS-CoV-2 infection and those who did not (see Fig. 1 and Material and Methods). We first compared weekly symptom reports of the *full cohort* consisting of a positive cohort (P) to a matched negative control cohort (NC) with respect to calendar week and weeks before and after a COVID-19 test (Fig. 1f). We then employed high temporal resolution data on RHR and activity and three age- and gender matched *match cohorts* (1:3)(Fig. 1h), i.e., individuals from P who report persistent symptoms (persistent symptoms match cohort PSM) to the remaining positive (PM) and negative match cohorts (NM). Finally, we compared survey results on well-being and quality of life across individuals within the *differentiated cohort* (Fig. 1g) consisting of those users in P who report persistent symptoms (PS) and those who don't (positive control cohort PC) and NC.

To assess these effects systematically and comprehensively, we leveraged the unique combination of wearable-derived vital and survey data at hand to discern individuals with persistent symptoms following a SARS-CoV-2 infection and analyzed their physical health across acute and post-acute COVID-19, shedding light on their characteristics, their course of infection and potential long-term consequences. We thereby focused on cohort averages of physiology (specifically RHR) and behavior (measured by step count). In order to make sense of the observed vital changes we then compared these objective measurements with self-reported QoL (WHO-5 and EQ-5D), showing how signals in passively sensed health data directly translate into subjectively experienced alterations or limitations in daily life.

2 Results

2.1 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) [2]. During the acute phase of infection, fatigue was reported as a symptom by up to $69 \cdot 76\%$ (SE $= 0 \cdot 54\%$) of individuals in P and by $19 \cdot 53\%$ (SE $= 0 \cdot 46\%$) of individuals in NC. Fatigue was reported significantly more often in P than in NC 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 $13 \cdot 55\%$ (SE $= 0 \cdot 51\%$) and $2 \cdot 46\%$ (SE $= 0 \cdot 23\%$) by individuals in P and NC, respectively. It was increasingly reported for 15 weeks in P (Fig. 2b). The combination of both symptoms was

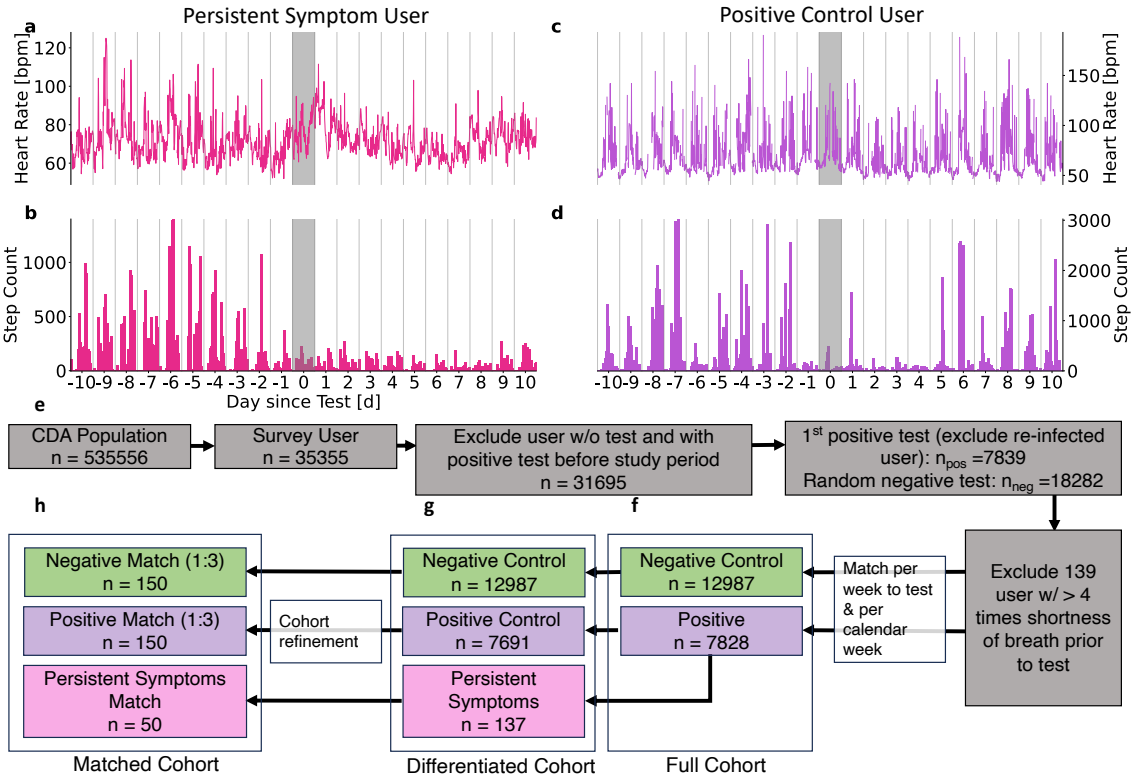

Fig 1. Time series of two representative individuals from the CDA population (a-d) and cohort diagram (e-h). (a,c) Average daily heart rate [bpm] and step counts (b,d) per 15 minutes relative to the day of the reported positive SARS-CoV-2 test (shaded grey area). Elevated heart rate and reduced step count persisted for more than 10 days after the test for an individual reporting persistent shortness of breath and fatigue (a-b). These patterns were less pronounced for an individual without persistent symptoms (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. (f) Based on all participants of the Corona Data Donation Application (CDA Population) we constructed a full cohort of individuals who reported a positive SARS-CoV-2 test (positive cohort P) and those who only reported negative SARS-CoV-2 tests (negative control cohort NC). (g) We then constructed a *differentiated cohort* by extracting all individuals from P who report shortness of breath and fatigue five times or more after a positive SARS-CoV-2 test, yielding the persistent symptoms cohort (PS) and a positive control cohort (PC) with the remaining participants. (h) We then constructed three age and gender matched *matched cohorts* (1:3), i.e. the persistent symptoms match cohort (PSM), the positive match cohort (PM) and the negative match cohort (NM). See Materials & Methods for more details.

reported by up to $12 \cdot 2\%$ (SE $= 0 \cdot 09\%$) and $1 \cdot 84\%$ (SE $= 0 \cdot 05\%$) by individuals in P and NC and at increased frequency in P 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 Post-COVID-condition (PCC) [31, 4, 32]. 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 P with NC (Fig. S1).

Based on the above findings, we defined a subset of P 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 (PS). Individuals in PS reported on average $6 \cdot 2$ (Std $= 2 \cdot 33$, SE $= 0 \cdot 2$) unique symptoms during the acute phase compared to $2 \cdot 41$ (Std $= 2 \cdot 39$, SE $= 0 \cdot 01$) symptoms reported by individuals with a positive SARS-CoV-2 test but no persistent symptoms (PC) and $0 \cdot 59$ (Std $= 1 \cdot 14$, SE $= 0 \cdot 00$) symptoms reported by individuals in NC (Fig. S2).

Fig. S3 illustrates relative frequencies of all symptoms in all phases, relative to the test date window for all cohorts, comprehensively.

	Corona Data Donation Application (CDA) Population	Positive (P) Cohort	Negative Control (NC) Cohort	Positive Control (PC) Cohort	Persistent Symptoms (PS) Cohort	Positive Match (PM) Cohort	Negative Match (NM) Cohort	Persistent Symptoms Match (PSM) Cohort
n	535,556	7,691	12,987	7,669	137	150	150	50
Vaccinated [%]	$7 \cdot 8$	$90 \cdot 0$	$88 \cdot 3$	$90 \cdot 0$	$88 \cdot 3$	$90 \cdot 0$	$91 \cdot 3$	$88 \cdot 0$
Not vaccinated [%]	$0 \cdot 1$	$0 \cdot 4$	$0 \cdot 5$	$0 \cdot 4$	$0 \cdot 7$	$0 \cdot 0$	$0 \cdot 0$	$2 \cdot 0$
Mean Age	$52 \cdot 8$	$52 \cdot 4$	$55 \cdot 0$	$52 \cdot 4$	$52 \cdot 7$	$53 \cdot 6$	$53 \cdot 6$	$53 \cdot 6$
(Std)	$(13 \cdot 5)$	$(11 \cdot 9)$	$(12 \cdot 0)$	$(11 \cdot 9)$	$(11 \cdot 7)$	$(10 \cdot 5)$	$(10 \cdot 5)$	$(10 \cdot 5)$
Reported Gender [%]	$7 \cdot 4$	$90 \cdot 7$	$88 \cdot 9$	$90 \cdot 8$	$88 \cdot 3$	$90 \cdot 0$	$90 \cdot 0$	$90 \cdot 0$
Female Ratio [%]	$36 \cdot 0$	$39 \cdot 8$	$38 \cdot 3$	$39 \cdot 6$	$53 \cdot 7$	$50 \cdot 0$	$50 \cdot 0$	$50 \cdot 0$
Mean BMI	$26 \cdot 5$	$26 \cdot 2$	$26 \cdot 5$	$26 \cdot 2$	$27 \cdot 7$	$25 \cdot 5$	$26 \cdot 9$	$27 \cdot 4$
(Std)	$(24 \cdot 4)$	$(4 \cdot 7)$	$(5 \cdot 1)$	$(4 \cdot 7)$	$(5 \cdot 7)$	$(4 \cdot 0)$	$(6 \cdot 2)$	$(4 \cdot 6)$

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.

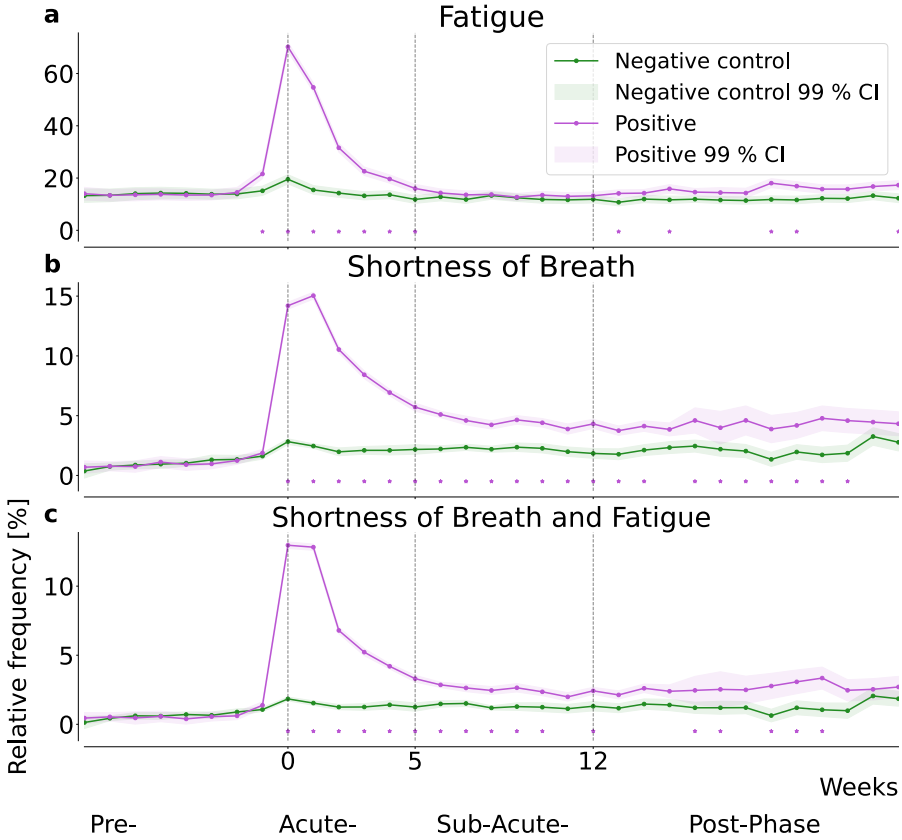

Fig 2. Relative frequency of self-reported fatigue (a), shortness of breath (b), and their combination(c), all relative to the week of the reported SARS-CoV-2 test for positive (P) individuals and a matched negative control (NC) cohort. Shading indicates the 99% confidence interval, i.e, $2 \cdot 576$ 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 $0 \cdot 01$.

2.2 Persistent symptoms correspond to more pronounced and prolonged RHR-changes

We investigated the corresponding wearable-derived vital data around the SARS-CoV-2 test (Fig.3) for the three age- and gender-matched (1:3) cohorts PSM, PM and NM (Fig.1h).

We evaluated the average RHR relative to the reporting day of a SARS-CoV-2 test within the three cohorts, see Fig.3a. For each individual, we first computed the average of all RHR intervals in the last seven days to obtain one data point per day, and adjusted for seasonal differences by subtracting the seasonal mean (daily mean RHR of the CDA population, (Fig. S4a)). To account for individual baselines and degrees of variations in RHR we ensured comparability across individuals by computing per-user Z-scores. For this, we subtracted from each RHR data point the respective user average and

divided by the standard deviation in the seven days prior to a reported SARS-CoV-2 test. That way all time series are centered around zero and are measured in units of standard deviations. Subsequently, we averaged over all individuals in the respective cohort, see Material and Methods.

We observed changes in RHR regulation for the PSM cohort compared to the PM and NM cohorts (Fig. 3 a). For the PSM cohort, we found a mean z-score of $-0 \cdot 88$ compared to $-0 \cdot 25$ for the PM cohort. This indicates that, on average, the deviation was $0 \cdot 88$ standard deviations below the pre-phase mean for the PSM cohort and $0 \cdot 25$ standard deviations below the mean for the PM cohort. The difference between maximum and minimum deviation in the period 14 to 20 days after the test was more pronounced for the PSM cohort (a factor of $1 \cdot 86$ in the corresponding z-scores). Individuals in PSM also showed evidence of transient tachycardia followed by a prolonged relative bradycardia, that did not return to baseline (z-score ≥ 0) until 18 days after infection. Individuals in PM and NM exhibited comparable levels of mean RHR except during the acute phase. The relative bradycardia of PM individuals persisted until 11 days after infection, one week less than in PSM individuals. Interestingly, we found that, on average, PSM individuals have an elevated RHR compared to the two control cohorts already *prior* to the reported SARS-CoV-2 test (inset in Fig.3a). Specifically, PSM individuals exhibited an average increase of $2 \cdot 37$ bpm and $1 \cdot 49$ (between 21 to 7 days to the test) compared to PM and NM individuals, respectively.

2.3 Physical activity profiles in acute and sub-acute infection regulation

To assess differences in physical activity patterns between individuals in the PSM, PM, and NM 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). We accounted for seasonal differences in activity (i.e., higher and lower step count during summer and winter, respectively, see (Fig. S4b)), bysubtracting the average of all participants in the CDA population from the average steps per day, see Material and Methods Section 4.3 for details. We report all results relative to the mean steps per day during pre-phase of the NM cohort. The respective median step count for PSM-individuals was consistently below the median of both control cohorts in all four phases. As the three cohorts are age and sex matched, PSM individuals should have comparable step counts to the two control cohorts, suggesting 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 PSM cohort was 3,030 steps less than that of the PM cohort and 2,909 steps less than the NM cohort. This indicates a generally lower level of activity compared to individuals from the PM and NM cohort. In addition, PSM individuals engaged at least three weeks prior to the infection in $1 \cdot 12$ and $0 \cdot 99$ days of high physical activity less than individuals from the PM and NM cohorts, respectively (Fig. S5).

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

In order to clarify the relationship between characteristics observed in wearable data and perceived (self-reported) wellbeing and QoL, we analyzed responses to an adapted version of the WHO-5 well-being index [33, 34] and a modified EQ-5D health-related QoL survey (see SI for details) across the *differentiated cohort* (Fig. 1g) , see Fig. 4. We did not distinguish whether a response occurred before or after a reported SARS-CoV-2 test because EQ-5D was not administered until after 2022. By then, most participants had already reported a test (see Fig. S8).

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 PS was skewed towards low values, (Fig. 4 a) with a mean value of $2 \cdot 62$ (Std $= 0 \cdot 55$, SE $= 0 \cdot 05$) for the PS cohort and a mean of $3 \cdot 22$ (Std $= 0 \cdot 63$, SE $= 0 \cdot 01$) and $3 \cdot 2$ (Std $= 0 \cdot 69$, SE $= 0 \cdot 01$) for the PC and NC cohort, respectively. We also observed lower wellbeing scores for all individual outcomes (Fig. 4 c-g). Users in PS reported particularly low scores for feeling energetic and rested upon waking up (Fig. 4d), possibly indicating activity and sleep issues. Statistically significant differences ($\alpha = 0 \cdot 001$) between the PS and the control cohorts were found in regards to all WHO-5 questions, as determined by the Mann-Whitney U test. Likewise, we

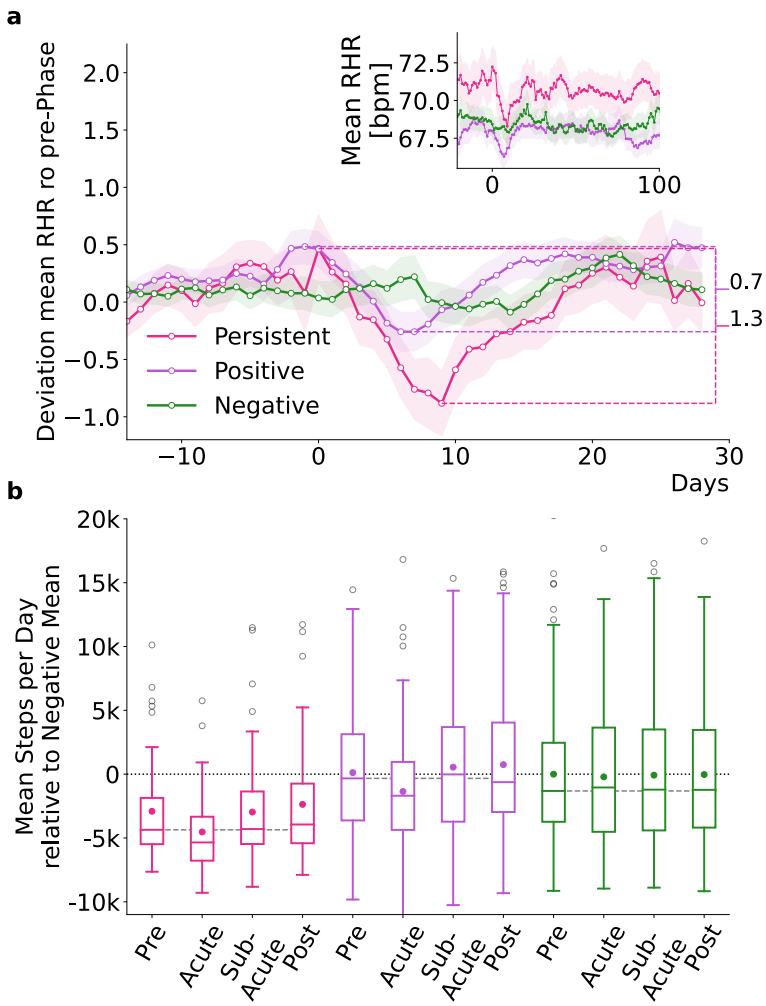

Fig 3. Wearable data analysis of the *match cohort*. (a) Z-transformed mean RHR (average of all 15-minute RHR measurements within the last seven days) relative to the seasonal mean RHR with respect to the mean and standard deviation up to seven days prior to the date of the reported test of all individuals in the PSM (pink), PM (purple) and NM (green) 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 ($1 \cdot 3$ vs $0 \cdot 7$) and more prolonged for PSM than for PM. 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, PSM-individuals showed an increased RHR compared to PM and NM. (b) Average steps per day relative to the mean of NM 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 PSM 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 steps per day values for PSM were below the mean steps per day values of the two control cohorts in all four phases. Likewise we found a reduction in the variance of the PSM compared to the two control cohorts.

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 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 PS compared to the control cohorts (Fig. 4b) with a mean of $2 \cdot 79$ (Std $= 0 \cdot 62$, SE $= 0 \cdot 06$). Again, PC and NC show similar values across modified EQ-5D questions and overall with a mean of $3 \cdot 55$ (Std $= 0 \cdot 46$, SE $= 0 \cdot 01$) and $3 \cdot 52$ (Std $= 0 \cdot 52$, SE $= 0 \cdot 01$) for PC and NC, respectively. Likewise, 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 PS cohort were significantly different from the control cohorts for all QoL questions ($\alpha = 0 \cdot 001$). Notably, also the answers of the PC compared to the NC 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 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 (Fig. S6) and after (Fig. S7) the reported SARS-CoV-2 tests and found systematically lower values of wellbeing and QoL for the PS cohort before as well as after the reported test compared to the PC and NC 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.

We additionally tested the predictive power of our results by means of a simple logistic regression model. It revealed that the variable representing pre-infection wellbeing performed best in predicting whether an individual belonged to the PM or PSM cohort. Other models that included median RHR and step count during pre- and acute phases did not improve prediction accuracy. Including wellbeing as a predictor resulted in a mean accuracy of $75 \cdot 51\%$ (std $= 6 \cdot 2$) and a moderate mean F1-score of $40 \cdot 0\%$ (std $= 5 \cdot 6 \times 10^{-15}$) over 1,000 runs.

3 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. Based on RHR, step count, survey-based QoL and predisposing health characterization we identified a phenotype of persistent COVID-19 symptoms in comparison with COVID-19 positive and negative controls. Individuals with this phenotype are characterized by persistent (sub-acute and post-phase) shortness of breath and fatigue, elevated RHR levels and reduced physical activity *prior* to the infection and on average lower wellbeing and health-related QoL scores compared to positive and negative controls.

We used weekly symptom reports to uncover notable differences in the prevalence of fatigue, shortness of breath and their co-occurrence in SARS-CoV-2 positive and negative individuals, respectively. Both symptoms have been reported as typical, persistent symptoms of COVID-19 which are common in individuals with post-COVID condition [14, 4, 35, 13]. Indeed, according to a WHO case definition, fatigue and dyspnoea (shortness of breath) are the main symptoms of post-COVID-19 condition [3].

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 in-

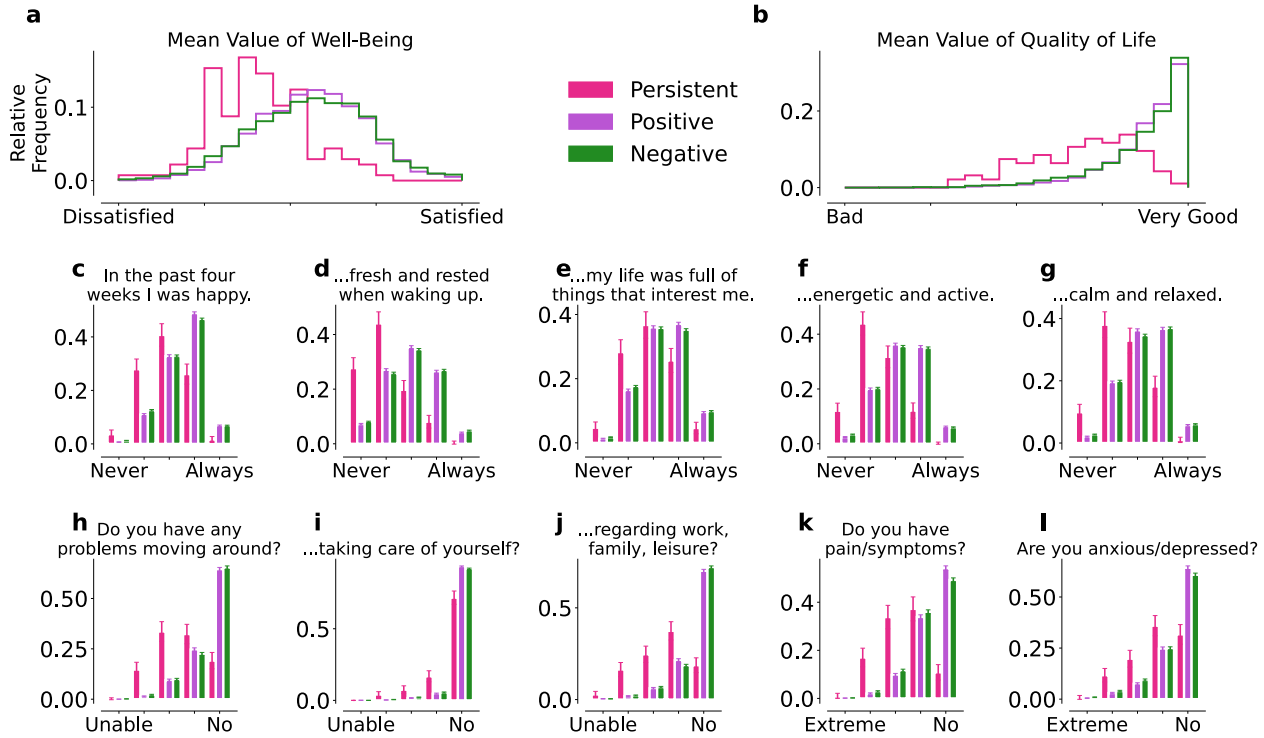

Fig 4. WHO-5 wellbeing (a,c-g) and modified EQ-5D/QoL (b,h-l) for the PS (pink), PC (purple), and NC (green) cohorts. The individual WHO-5 and modified EQ-5D scores were both averaged to obtain the overall wellbeing (a) and QoL (b) scores, respectively. Overall, PS individuals reported more issues with wellbeing (c-g) and QoL (h-l) than the control cohorts. Error bars indicate standard errors. The responses from all PS cohort were significantly different ($\alpha < 0 \cdot 001$) 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).

fection compared to positive controls. Similar differences have been observed in previous studies when comparing SARS-CoV-2 positive and negative individuals [21, 22]. Individuals that later reported persistent symptoms already exhibited an elevated average RHR (mean increase of $2 \cdot 37$ bpm/$1 \cdot 49$ 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 $1 \cdot 12/0 \cdot 99$ days less during the pre-phase than the positive/negative controls. Given that RHR and activity are generally anti-correlated [36, 37], these results indicate that PSM-users generally display lower physical fitness levels which could be caused by pre-existing health conditions while simultaneously increasing their susceptibility for developing long-term symptoms. We found that individuals with persistent symptoms reported more medical conditions with allergies being $1 \cdot 68$ times more common and mental illness being $1 \cdot 79$ to $2 \cdot 07$ times more common than in controls, Table S1. This is in line with findings on the association of these conditions and the risk of developing post-COVID condition [38, 39]. However, 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.
Finally, we found that individuals with persistent symptoms reported significantly lower levels ($\alpha = 0 \cdot 001$) of wellbeing and QoL compared to the control cohorts. This finding holds for periods before and after the reported SARS-CoV-2 test, possibly indicating physical or mental health problems (e.g. chronic conditions, allergies, etc., see Table S1.) among individuals who later report persistent symptoms. Specifically we find a mean difference of $0 \cdot 6/0 \cdot 58$ in mean well-being score and a mean difference of $0 \cdot 76/0 \cdot 73$ between PS and PC/NC individuals. This is in line with recent findings, where individuals reporting persistent symptoms also reported poorer mental health in comparison to a control group [40], as well as a study reporting disturbed sleep as an important risk factor for PCC [41].

Our data allow a detailed examination of activity and heart rate patterns, providing valuable insights not only into the many ways in which SARS-CoV-2 infection can affect behaviour 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, under natural conditions, and outside of healthcare facilities [42, 43]. Wearables thus have the potential to improve patient-centered care by empowering individuals to monitor their health and manage their symptoms more effectively [44, 45]. 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 post-COVID-19 condition research, since this condition is not exclusively linked to severe COVID-19 cases [46].

Several factors should be considered when interpreting our results. Across all participants in the Corona Data Donation project we observed an over-representation of males and an under-representation of adolescents and elderly (65+) individuals [21], the latter probably being the individuals at greatest risk of suffering from long-term consequences of a SARS-CoV-2 infection [47]. It is also reasonable to assume that the study population is more health-conscious than the general population, as the adoption of (wrist-worn) wearables is partly associated with a predominance of health-related habits [48]. We further did not explicitly account for any potential vaccination status, as the majority of participants in the CDA are at least partially vaccinated or do not report their vaccination status, see table 1. We did not distinguish infections by variants of concern (VoC), which could potentially elicit different immune responses and potentially different heart rate regulation. However, most individuals included in this study report their SARS-CoV-2 test date in 2022 (Fig. S8) when the variant B.1.1.529 (Omicron) was predominant in Germany. Keeping all these considerations in mind, it is reasonable to assume that our

results under-represent the true disease burden of PCC and rather provide a lower bound for expected long-term imprints of an infection [49, 50, 51, 52, 40].

The sample sizes of our study cohorts are limited as we demand sufficient wearable data to be present in all four phases (pre, acute, sub-acute and post) around a SARS-CoV-2 test. Additionally, we imposed strict criteria on our definition of persistent symptoms (more than five times shortness of breath and fatigue after the reported SARS-CoV-2 test) as for now those are the ones most commonly observed in individuals suffering from PCC [3, 15, 13, 35, 14]. These strict criteria lead to a sole focus on a specific subset of potential symptoms, effectively excluding other potential imprints of COVID-19 on long-term health [15, 4, 14].

We emphasize that the first two cohorts studied for symptom reporting are not matched in terms of age and gender, but only in terms of calendar week and temporal distance to a reported SARS-CoV-2 test. The resulting differences in the age ($2 \cdot 6$ years) and gender ($1 \cdot 5$ %) distribution of these cohorts are small and unlikely to influence the results (Table 1). Previous studies that stratified symptom reports by gender only found differences after the reported SARS-CoV-2 test for symptoms other than those studied here [13].

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 post-COVID-19 condition [53, 54]. Further combination of survey and wearable data would provide a deeper understanding of the relationship between subjective symptom reports and objective vital signs, 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.

4 Materials and Methods

4.1 Data Characteristics

From April 12, 2020, to December 31, 2022, a total of 535,556 individuals downloaded and used the Corona-Datenspende App (CDA) [30], actively contributing a minimum of one essential data point each. The vital 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.

Symptom reports were surveyed on a weekly basis since October 21, 2021. Until November 09, 2022, 35,355 individuals participated in the respective survey at least once. Participants could either choose to report no symptoms or one/multiple symptoms from the following list: Shortness of Breath, Fatigue, Headache, Sore Throat, Loss of Smell and Taste, Diarrhea, Runny Nose, Cough, Chills, Limb Ache, Fever. We compared reports on symptoms between individuals who reported a positive SARS-CoV-2 test (P = 7,691 individuals) and a negative control cohort (NC = 12,987 individuals) (see Section 4.2 and Fig.1e). Participants in the P cohort reported symptoms an average of $18 \cdot 4$ times (standard deviation (Std) = $11 \cdot 2$). Participants in the NC cohort submitted reports $11 \cdot 1$ times on average (Std = $11 \cdot 4$). The P cohort was further split in sub-cohorts, one with individuals that reported persistent symptoms (PS = 137 individuals) and the rest (PC = 7669 individuals), which together with the NC cohort yielding in three cohorts, see Section 4.2 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 PC cohort and 12,866 users from the NC cohort answered the survey on well-being at least once. The WHO-5 wellbeing index [33, 34] 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 exact wording of the five questions is in the SI. 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 PC cohort and 12,866 individuals from the NC cohort answered the five questions on QoL a least once (see SI for exact wording). 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* (PSM = 50), *positive match* (PM = 150) and *negative match* (NM = 150), we analyzed high frequency longitudinal data on RHR and step count (see Section 4.2 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.

4.2 Cohort creation

To study reported symptoms relative to a reported positive SARS-CoV-2 test (P) and compared them with a negative control (NC) 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 (PC) of 7,691 individuals and a negative control (NC) of 12,987 individuals (full cohort, Fig. 1f). For the analysis of WHO-5 and modified EQ-5D, we selected individuals from P 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 (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 (PC)) and the 12,987 individuals from the negative control (NC) cohort (differentiated cohort Fig. 1g). To study differences and changes in physiological parameters (RHR) and behavior (step count) between the PC, NC, and PS cohorts, we analyzed fine-grained vital 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 [2]. After excluding individuals with less than 100 vital data points across all phases, the Persistent Symptoms Match (PSM) cohort comprises 50 individuals. We then created age- and gender-matched Positive Match (PM) and Negative Match (NM) cohorts, each with 150 individuals selected from the PC and NC cohorts, respectively (match cohort, Fig. 1h). For characteristics of the cohorts see Tabs. 1 and S1.

4.3 Data processing and statistical analysis

4.3.1 Weekly reports on symptoms.

For the P and the NC 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 $0 \cdot 01$. To investigate whether pairs of symptoms are significantly concordant, we compared the relative frequency of reports where individuals report both symptoms.

4.3.2 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 PSM cohort, the PM cohort and the NM 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 (Fig. S4a).

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 [21], 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 Fig. S4b. We show the difference to the mean daily step count (adjusted for seasonal variation) of the NM cohort during pre-phase. A positive/negative value then indicates that an individual is more/less active than the NM 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.

4.3.3 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.

4.3.4 Logistic Regression.

Logistic regression analysis was performed with the target variable indicating individuals with persistent symptoms (PSM) versus a randomly selected sample of the same size from the cohort without persistent symptoms (PM). This process was iterated 1,000 times, with each iteration utilizing a different random sample. The explanatory variable was the mean WHO-5 wellbeing score up until one

week prior to the test. Individuals with missing data for any explanatory variable were excluded from the analysis.

4.3.5 Role of the funding source.

This study was funded in part with funds from the funding program of the City of Vienna MA7 with no role of the study sponsor in study design, collection, analysis, interpretation of the data, writing of the report and decision to submit.

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 vital 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).

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.

Declaration of interests

The authors declare no conflicting interests.

References

[1] World Health Organization. *Post COVID-19 condition (Long COVID)*. URL: 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] Ani Nalbandian et al. "Post-acute COVID-19 syndrome". en. In: *Nature Medicine* 27.4 (Apr. 2021), pp. 601–615. ISSN: 1078-8956, 1546-170X. DOI: 10.1038/s41591-021-01283-z. URL: https://www.nature.com/articles/s41591-021-01283-z (visited on 09/26/2023).

[3] Joan B Soriano et al. "A clinical case definition of post-COVID-19 condition by a Delphi consensus". en. In: *The Lancet Infectious Diseases* 22.4 (Apr. 2022), e102–e107. ISSN: 14733099. DOI: 10.1016/S1473-3099(21)00703-9. URL: https://linkinghub.elsevier.com/retrieve/pii/S1473309921007039 (visited on 09/26/2023).

[4] Hannah E. Davis et al. "Long COVID: major findings, mechanisms and recommendations". en. In: *Nature Reviews Microbiology* 21.3 (Mar. 2023), pp. 133–146. ISSN: 1740-1526, 1740-1534. DOI: 10.1038/s41579-022-00846-2. URL: https://www.nature.com/articles/s41579-022-00846-2 (visited on 09/11/2023).

[5] Dominik Menges et al. "Burden of post-COVID-19 syndrome and implications for healthcare service planning: A population-based cohort study". en. In: *PLOS ONE* 16.7 (July 2021). Ed. by Martin Chtolongo Simuunza, e0254523. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0254523. URL: https://dx.plos.org/10.1371/journal.pone.0254523 (visited on 10/11/2023).

[6] Evan Xu, Yan Xie, and Ziyad Al-Aly. "Long-term neurologic outcomes of COVID-19". en. In: *Nature Medicine* (Sept. 2022). ISSN: 1078-8956, 1546-170X. DOI: 10.1038/s41591-022-02001-z. URL: https://www.nature.com/articles/s41591-022-02001-z (visited on 10/04/2022).

[7] Evan Xu, Yan Xie, and Ziyad Al-Aly. "Long-term gastrointestinal outcomes of COVID-19". en. In: *Nature Communications* 14.1 (Mar. 2023), p. 983. ISSN: 2041-1723. DOI: 10.1038/s41467-023-36223-7. URL: https://www.nature.com/articles/s41467-023-36223-7 (visited on 09/13/2023).

[8] Yan Xie et al. "Long-term cardiovascular outcomes of COVID-19". en. In: *Nature Medicine* 28.3 (Mar. 2022), pp. 583–590. ISSN: 1078-8956, 1546-170X. DOI: 10.1038/s41591-022-01689-3. URL: https://www.nature.com/articles/s41591-022-01689-3 (visited on 04/11/2022).

[9] Yan Xie, Evan Xu, and Ziyad Al-Aly. "Risks of mental health outcomes in people with covid-19: cohort study". en. In: *BMJ* (Feb. 2022), e068993. ISSN: 1756-1833. DOI: 10.1136/bmj-2021-068993. URL: https://www.bmj.com/lookup/doi/10.1136/bmj-2021-068993 (visited on 04/11/2022).

[10] Ziyad Al-Aly, Benjamin Bowe, and Yan Xie. "Long COVID after breakthrough SARS-CoV-2 infection". en. In: *Nature Medicine* (May 2022). ISSN: 1078-8956, 1546-170X. DOI: 10.1038/s41591-022-01840-0. URL: https://www.nature.com/articles/s41591-022-01840-0 (visited on 06/01/2022).

[11] Ziyad Al-Aly, Yan Xie, and Benjamin Bowe. "High-dimensional characterization of post-acute sequelae of COVID-19". en. In: *Nature* 594.7862 (June 2021), pp. 259–264. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-021-03553-9. URL: http://www.nature.com/articles/s41586-021-03553-9 (visited on 04/20/2022).

[12] Benjamin Bowe et al. "Kidney Outcomes in Long COVID". en. In: *Journal of the American Society of Nephrology* 32.11 (Nov. 2021), pp. 2851–2862. ISSN: 1046-6673, 1533-3450. DOI: 10.1681/ASN.2021060734. URL: https://jasn.asnjournals.org/lookup/doi/10.1681/ASN.2021060734 (visited on 11/02/2022).

[13] Aranka V Ballering et al. "Persistence of somatic symptoms after COVID-19 in the Netherlands: an observational cohort study". en. In: *The Lancet* 400.10350 (Aug. 2022), pp. 452–461. ISSN: 01406736. DOI: 10.1016/S0140-6736(22)01214-4. URL: https://linkinghub.elsevier.com/retrieve/pii/S0140673622012144 (visited on 09/11/2023).

[14] Tahmina Nasserie, Michael Hittle, and Steven N. Goodman. "Assessment of the Frequency and Variety of Persistent Symptoms Among Patients With COVID-19: A Systematic Review". en. In: *JAMA Network Open* 4.5 (May 2021), e2111417. ISSN: 2574-3805. DOI: 10.1001/jamanetworkopen.2021.11417. URL: https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2780376 (visited on 09/26/2023).

[15] Hannah E. Davis et al. "Characterizing long COVID in an international cohort: 7 months of symptoms and their impact". en. In: *eClinicalMedicine* 38 (Aug. 2021), p. 101019. ISSN: 25895370. DOI: 10.1016/j.eclinm.2021.101019. URL: https://linkinghub.elsevier.com/retrieve/pii/S2589537021002996 (visited on 09/11/2023).

[16] Zoe Marjenberg et al. "Risk of long COVID main symptoms after SARS-CoV-2 infection: a systematic review and meta-analysis". In: *Scientific Reports* (2023). DOI: 10.1038/s41598-023-42321-9.

[17] Tracy Beth Høeg, Shamez Ladhani, and Vinay Prasad. "How methodological pitfalls have created widespread misunderstanding about long COVID". In: *BMJ Evidence-Based Medicine* (2023). DOI: 10.1136/bmjebm-2023-112338.

[18] LL O'Mahoney et al. "The prevalence and long-term health effects of Long Covid among hospitalised and non-hospitalised populations: a systematic review and meta-analysis". In: *eClinicalMedicine* (2023). DOI: 10.1016/j.eclinm.2022.101762.

[19] Ho Cheng Koc et al. "Long COVID and its Management". en. In: *International Journal of Biological Sciences* 18.12 (2022), pp. 4768–4780. ISSN: 1449-2288. DOI: 10.7150/ijbs.75056. URL: https://www.ijbs.com/v18p4768.htm (visited on 10/17/2023).

[20] Qing Han et al. "Long-Term Sequelae of COVID-19: A Systematic Review and Meta-Analysis of One-Year Follow-Up Studies on Post-COVID Symptoms". In: *Pathogens* (2022). DOI: 10.3390/pathogens11020269.

[21] Marc Wiedermann et al. "Evidence for positive long- and short-term effects of vaccinations against COVID-19 in wearable sensor metrics". en. In: *PNAS Nexus* 2.7 (July 2023). Ed. by B Levine, pgad223. ISSN: 2752-6542. DOI: 10.1093/pnasnexus/pgad223. URL: https://academic.oup.com/pnasnexus/article/doi/10.1093/pnasnexus/pgad223/7225910 (visited on 09/25/2023).

[22] Jennifer M. Radin et al. "Assessment of Prolonged Physiological and Behavioral Changes Associated With COVID-19 Infection". en. In: *JAMA Network Open* 4.7 (July 2021), e2115959. ISSN: 2574-3805. DOI: 10.1001/jamanetworkopen.2021.15959. URL: https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2781687 (visited on 09/11/2023).

[23] Aravind Natarajan, Hao-Wei Su, and Conor Heneghan. "Occurrence of Relative Bradycardia and Relative Tachycardia in Individuals Diagnosed With COVID-19". en. In: *Frontiers in Physiology* 13 (May 2022), p. 898251. ISSN: 1664-042X. DOI: 10.3389/fphys.2022.898251. URL: https://www.frontiersin.org/articles/10.3389/fphys.2022.898251/full (visited on 09/26/2023).

[24] Tejaswini Mishra et al. "Pre-symptomatic detection of COVID-19 from smartwatch data". en. In: *Nature Biomedical Engineering* 4.12 (Nov. 2020), pp. 1208–1220. ISSN: 2157-846X. DOI: 10.1038/s41551-020-00640-6. URL: https://www.nature.com/articles/s41551-020-00640-6 (visited on 09/26/2023).

[25] Matteo Gadaleta et al. "Passive detection of COVID-19 with wearable sensors and explainable machine learning algorithms". en. In: *npj Digital Medicine* 4.1 (Dec. 2021), p. 166. ISSN: 2398-6352. DOI: 10.1038/s41746-021-00533-1. URL: https://www.nature.com/articles/s41746-021-00533-1 (visited on 09/26/2023).

[26] Giorgio Quer et al. "Wearable sensor data and self-reported symptoms for COVID-19 detection". en. In: *Nature Medicine* 27.1 (Jan. 2021), pp. 73–77. ISSN: 1078-8956, 1546-170X. DOI: 10.1038/s41591-020-1123-x. URL: https://www.nature.com/articles/s41591-020-1123-x (visited on 09/26/2023).

[27] Thomas Varsavsky et al. "Detecting COVID-19 infection hotspots in England using large-scale self-reported data from a mobile application: a prospective, observational study". en. In: *The Lancet Public Health* 6.1 (Jan. 2021), e21–e29. ISSN: 24682667. DOI: 10.1016/S2468-2667(20)30269-3. URL: https://linkinghub.elsevier.com/retrieve/pii/S2468266720302693 (visited on 09/25/2023).

[28] Cristina Menni 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. In: *The Lancet Infectious Diseases* 21.7 (July 2021), pp. 939–949. ISSN: 14733099. DOI: 10.1016/S1473-3099(21)00224-3. URL: https://linkinghub.elsevier.com/retrieve/pii/S1473309921002243 (visited on 09/25/2023).

[29] Jakob J Kolb et al. "Prevalence of Positive COVID-19 Test Results Collected by Digital Self-report in the US and Germany". In: 6(1):e2253800 (2023). DOI: doi:10.1001/jamanetworkopen.2022.53800.

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

[31] Melina Michelen et al. "Characterising long COVID: a living systematic review". en. In: *BMJ Global Health* 6.9 (Sept. 2021), e005427. ISSN: 2059-7908. DOI: 10.1136/bmjgh-2021-005427. URL: https://gh.bmj.com/lookup/doi/10.1136/bmjgh-2021-005427 (visited on 10/03/2023).

[32] Liane S Canas 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. In: *The Lancet Digital Health* 5.7 (July 2023), e421–e434. ISSN: 25897500. DOI: 10.1016/S2589-7500(23)00056-0. URL: https://linkinghub.elsevier.com/retrieve/pii/S2589750023000560 (visited on 11/02/2023).

[33] Christian Winther Topp et al. "The WHO-5 Well-Being Index: A Systematic Review of the Literature". en. In: *Psychotherapy and Psychosomatics* 84.3 (2015), pp. 167–176. ISSN: 0033-3190, 1423-0348. DOI: 10.1159/000376585. URL: https://www.karger.com/Article/FullText/376585 (visited on 09/29/2023).

[34] (World Health Organisation) WHO. *WHO-5 Well-Being Index*. URL: https://www.corc.uk.net/outcome-experience-measures/the-world-health-organisation-five-well-being-index-who-5/ (visited on 09/29/2023).

[35] Nhu Ngoc Nguyen et al. "Clinical patterns of somatic symptoms in patients suffering from post-acute long COVID: a systematic review". en. In: *European Journal of Clinical Microbiology & Infectious Diseases* 41.4 (Apr. 2022), pp. 515–545. ISSN: 0934-9723, 1435-4373. DOI: 10.1007/s10096-022-04417-4. URL: https://link.springer.com/10.1007/s10096-022-04417-4 (visited on 09/26/2023).

[36] Anne Kerstin Reimers, Guido Knapp, and Carl-Detlev Reimers. "Effects of Exercise on the Resting Heart Rate: A Systematic Review and Meta-Analysis of Interventional Studies". In: *Journal of Clinical Medicine* 7.12 (2018). ISSN: 2077-0383. DOI: 10.3390/jcm7120503. URL: https://www.mdpi.com/2077-0383/7/12/503.

[37] Guoyuan Huang et al. "Resting heart rate changes after endurance training in older adults: a meta-analysis". In: *Medicine and science in sports and exercise* 37.8 (Aug. 2005), pp. 1381–1386. ISSN: 0195-9131. DOI: 10.1249/01.mss.0000174899.35392.0c. URL: https://journals.lww.com/acsm-msse/Fulltext/2005/08000/Resting_Heart_Rate_Changes_after_Endurance.18.aspx.

[38] KEJ Philip et al. "Impact of COVID-19 on people with asthma: a mixed methods analysis from a UK wide survey". In: *BMJ Open Resp Res* (2022). DOI: 10.1136/bmjresp-2021-001056.

[39] S Wang et al. "Associations of Depression, Anxiety, Worry, Perceived Stress, and Loneliness Prior to Infection With Risk of Post–COVID-19 Conditions". In: *JAMA Psychiatry* (2022). DOI: 10.1001/jamapsychiatry.2022.2640.

[40] Christina J. Atchison et al. "Long-term health impacts of COVID-19 among 242,712 adults in England". en. In: *Nature Communications* 14.1 (Oct. 2023), p. 6588. ISSN: 2041-1723. DOI: 10.1038/s41467-023-41879-2. URL: https://www.nature.com/articles/s41467-023-41879-2 (visited on 11/21/2023).

[41] C Schilling et al. "Pre-existing sleep problems as a predictor of post-acute sequelae of COVID-19". In: *Journal of Sleep Research* (2024). DOI: 10.1111/jsr.13949.

[42] Ali K. Yetisen et al. "Wearables in Medicine". en. In: *Advanced Materials* 30.33 (2018). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201706910, p. 1706910. ISSN: 1521-4095. DOI: 10.1002/adma.201706910. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201706910 (visited on 02/20/2024).

[43] Jamie L. Adams et al. "A real-world study of wearable sensors in Parkinson's disease". en. In: *npj Parkinson's Disease* 7.1 (Nov. 2021). Number: 1 Publisher: Nature Publishing Group, pp. 1–8. ISSN: 2373-8057. DOI: 10.1038/s41531-021-00248-w. URL: https://www.nature.com/articles/s41531-021-00248-w (visited on 02/20/2024).

[44] Kamil Reza Khondakar and Ajeet Kaushik. "Role of Wearable Sensing Technology to Manage Long COVID". en. In: *Biosensors* 13.1 (Dec. 2022), p. 62. ISSN: 2079-6374. DOI: 10.3390/bios13010062. URL: https://www.mdpi.com/2079-6374/13/1/62 (visited on 09/11/2023).

[45] Talha Burki. "Wearable technology and COVID-19". en. In: *The Lancet Respiratory Medicine* 10.10 (Oct. 2022), pp. 934–935. ISSN: 22132600. DOI: 10.1016/S2213-2600(22)00351-4. URL: https://linkinghub.elsevier.com/retrieve/pii/S2213260022003514 (visited on 09/11/2023).

[46] Harry Crook et al. "Long covid—mechanisms, risk factors, and management". en. In: *BMJ* (July 2021), n1648. ISSN: 1756-1833. DOI: 10.1136/bmj.n1648. URL: https://www.bmj.com/lookup/doi/10.1136/bmj.n1648 (visited on 10/17/2023).

[47] Claudia Carvalho-Schneider et al. "Follow-up of adults with noncritical COVID-19 two months after symptom onset". en. In: *Clinical Microbiology and Infection* 27.2 (Feb. 2021), pp. 258–263. ISSN: 1198743X. DOI: 10.1016/j.cmi.2020.09.052. URL: https://linkinghub.elsevier.com/retrieve/pii/S1198743X20306066 (visited on 11/20/2023).

[48] Quinge Wu, Kelli Sum, and Dan Nathan-Roberts. "How Fitness Trackers Facilitate Health Behavior Change". In: 60 (Sept. 2016), pp. 1068–1072. DOI: 10.1177/1541931213601247.

[49] SA Richard et al. "Persistent COVID-19 Symptoms at 6 Months After Onset and the Role of Vaccination Before or After SARS-CoV-2 Infection". In: *JAMA Network Open* (2023). DOI: 10.1001/jamanetworkopen.2022.51360.

[50] Ana Hernández-Aceituno, Abigail García-Hernández, and Eneko Larumbe-Zabala. "COVID-19 long-term sequelae: Omicron versus Alpha and Delta variants". In: *Infectious Diseases Now* (2023). DOI: 10.1016/j.idnow.2023.104688.

[51] Alexandre R. Marra 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. In: *Antimicrobial Stewardship & Healthcare Epidemiology* 3.1 (2023), e168. ISSN: 2732-494X. DOI: 10.1017/ash.2023.447. URL: https://www.cambridge.org/core/product/identifier/S2732494X23004473/type/journal_article (visited on 11/21/2023).

[52] Joshua Nealon and Benjamin J Cowling. "Omicron severity: milder but not mild". en. In: *The Lancet* 399.10323 (Jan. 2022), pp. 412–413. ISSN: 01406736. DOI: 10.1016/S0140-6736(22)00056-3. URL: https://linkinghub.elsevier.com/retrieve/pii/S0140673622000563 (visited on 11/20/2023).

[53] Satish A Raj et al. "Long-COVID postural tachycardia syndrome: an American Autonomic Society statement". In: *Clin Auton Res* 31 (2021), pp. 365–368. DOI: 10.1007/s10286-021-00798-2.

[54] Cameron K. Ormiston, Iwona Świątkiewicz, and Pam R. Taub. "Postural orthostatic tachycardia syndrome as a sequela of COVID-19". In: *Heart Rhythm* 19.11 (2022), pp. 1880–1889. ISSN: 1547-5271. DOI: https://doi.org/10.1016/j.hrthm.2022.07.014. URL: https://www.sciencedirect.com/science/article/pii/S1547527122021853.

	CDA Population	Positive	Negative Control	Positive Control	Persistent Symptoms	Positive Match	Negative Match	Persistent Symptoms Match
n	535,556	7,828	12,987	7,691	137	150	150	50
% answered	7 · 3	89 · 9	87 · 7	89 · 9	88 · 3	90 · 0	89 · 3	90 · 0
No disease in the past 12 months [%]	46 · 6	48 · 5	44 · 3	49 · 0	24 · 8	54 · 1	41 · 0	31 · 1
Allergies [%]	29 · 5	30 · 4	30 · 0	30 · 0	50 · 4	25 · 2	32 · 1	46 · 7
Hypertension [%]	20 · 5	18 · 2	22 · 6	18 · 1	23 · 1	16 · 3	25 · 4	20 · 0
Asthma [%]	9 · 6	9 · 3	9 · 3	9 · 0	24 · 0	11 · 1	11 · 9	22 · 2
Increased blood lipid/cholesterol levels [%]	11 · 7	10 · 4	12 · 8	10 · 3	15 · 7	8 · 1	12 · 7	17 · 8
Chronic Bronchitis [%]	2 · 8	2 · 1	3 · 1	2 · 0	6 · 6	1 · 5	1 · 5	8 · 9
Other [%]	12 · 2	10 · 5	13 · 9	10 · 3	19 · 8	10 · 4	9 · 7	22 · 2
Mental Illness Diagnose [%]	20 · 7	19 · 5	22 · 2	19 · 2	39 · 7	15 · 6	20 · 1	44 · 4

Table I. Characterization of individuals in the persistent symptoms, positive control and negative control cohort. See Materials and Methods (Section IV) for further details on the specifics of the assignments.

WHO-5 Questions on Well-Being. 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.
- ... my everyday life was full of things that interest me.

Modified EQ-5D Sruvey on Quality-of-Life. 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?
- Are you anxious or depressed?

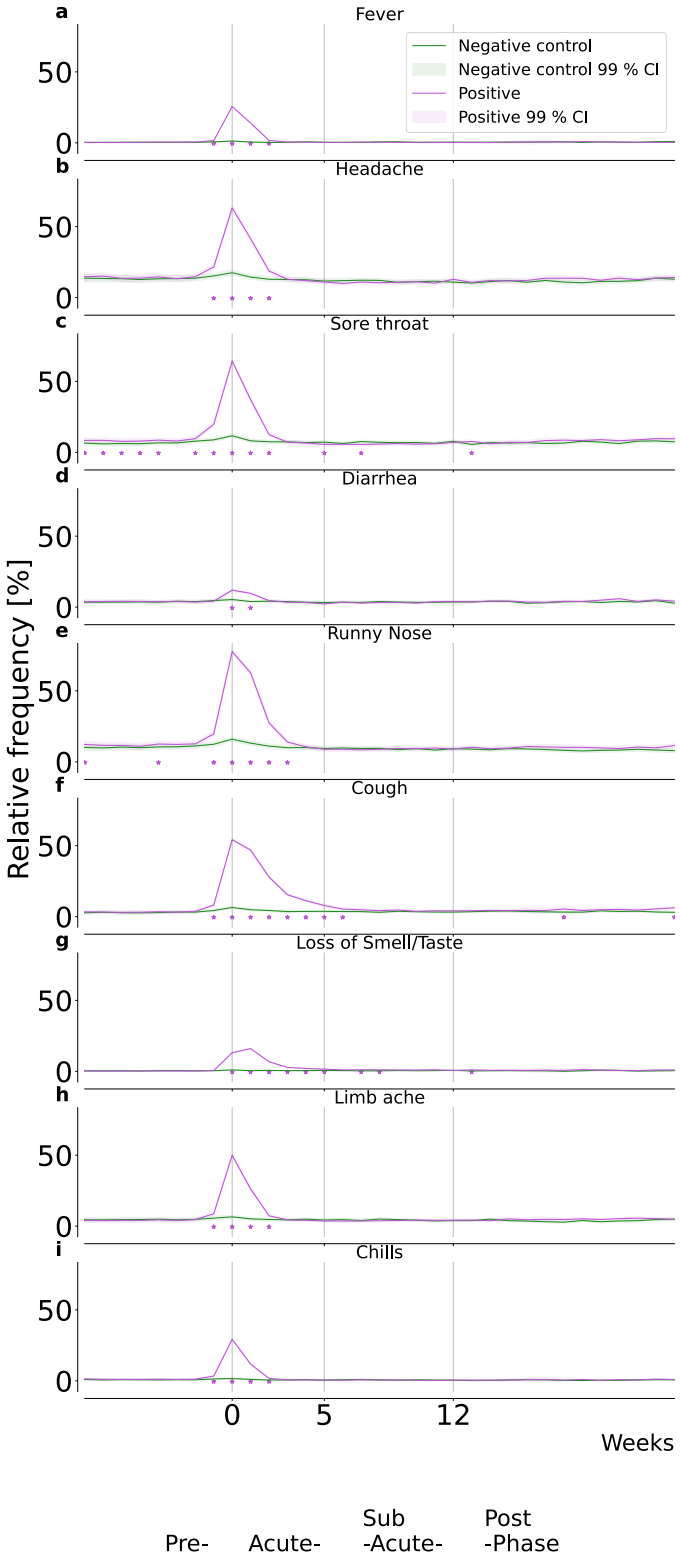

Figure 1. Relative frequency of symptoms relative to week of reported SARS-CoV-2 test for positive and negative control cohort.

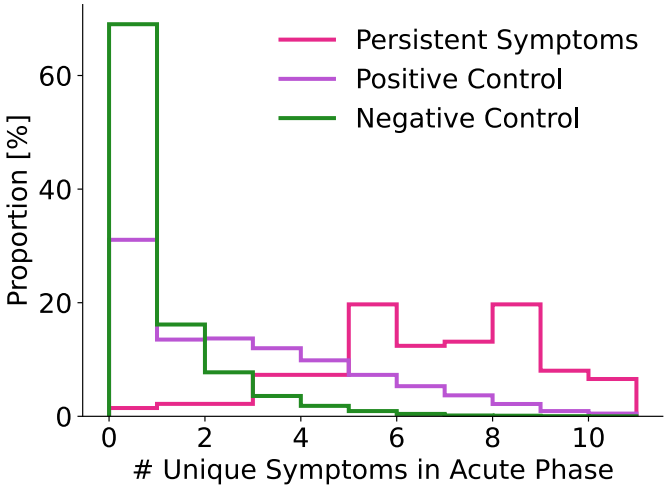

Figure 2. Distribution of number of unique reported symptoms during acute phase of infection (0-4 weeks to the reported SARS-CoV-2 test) studied cohorts.

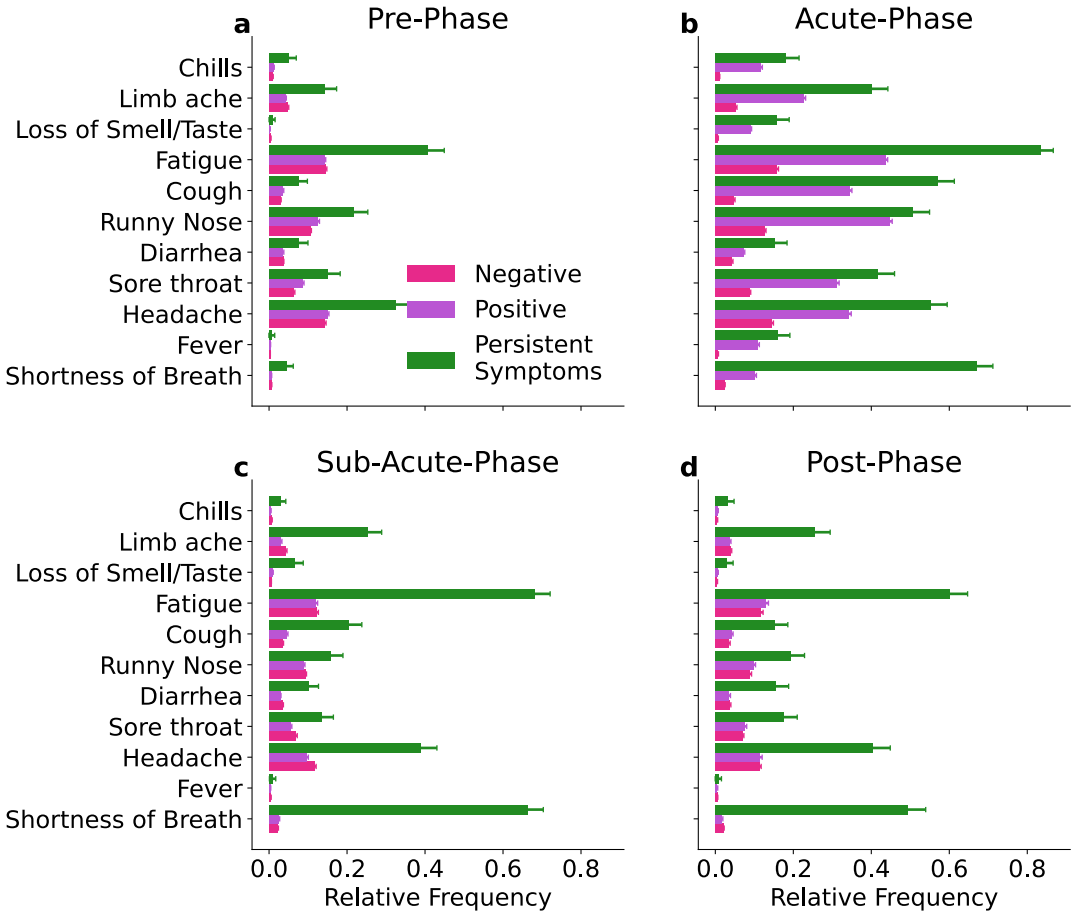

Figure 3. Relative frequency per symptom. For pre-, acute-, sub-acute and post phase we show the relative frequency per reported symptom for all three groups in the vital cohort: negative match, positive match and persistent symptoms match.

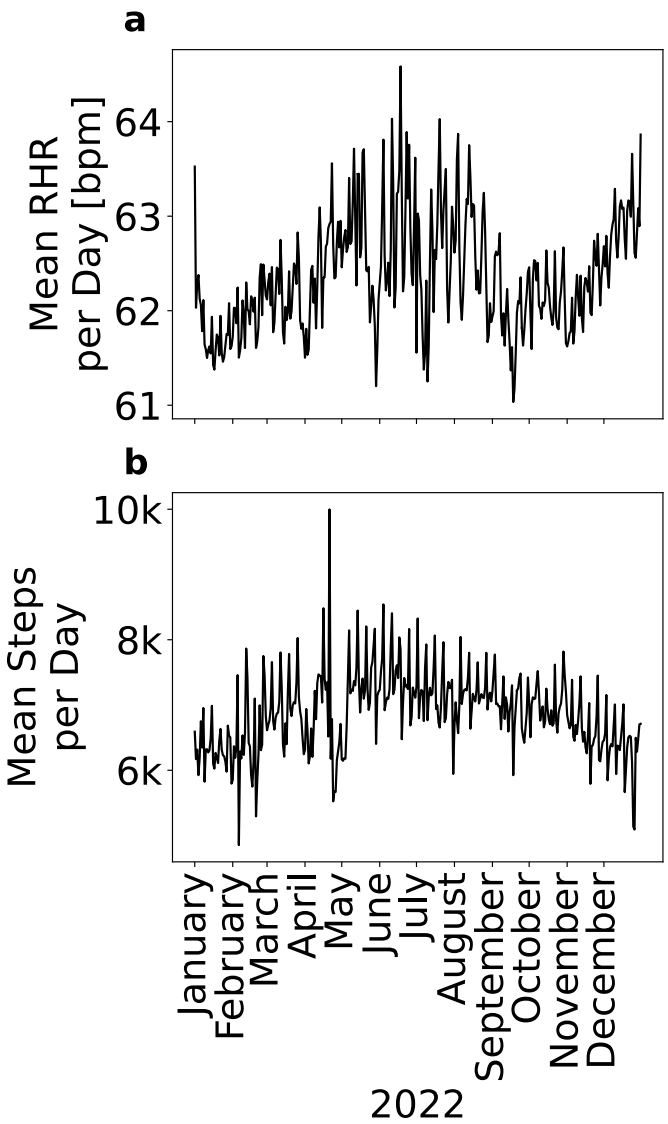

Figure 4. Seasonal trends in vital data. Exemplary timeseries of mean RHR per day (a) and mean steps per day (b) over the CDA Population for one example year (2022) and one example donation-source (Apple).

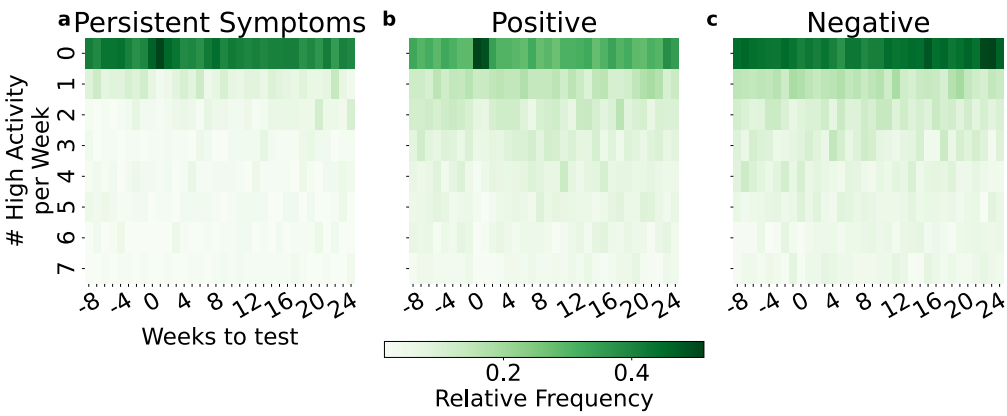

Figure 5. High activity profiles of the *fine cohort*. Relative frequency of the number of high activity days per week relative to the SARS-CoV-2 test date for PSM (a), PM (b) and NM (b). A day is defined as a high activity day when the corresponding number of steps per day exceed one standard deviation over the seasonal mean taken over the whole set of users in the Corona Data Donation Project (CDA Population).

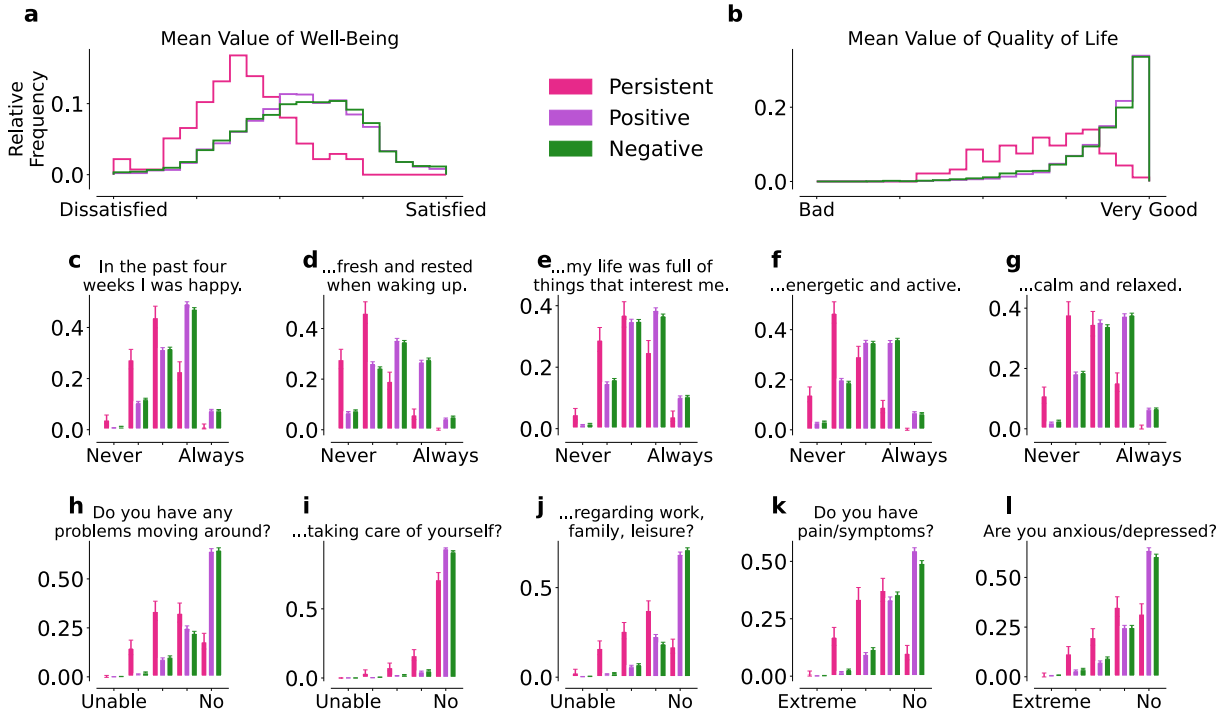

Figure 6. WHO-5 wellbeing (a,c-g) and modified EQ-5D/QoL (b,h-l) for the *PS* (pink), *PC* (purple), and *NC* (green) cohorts before the reported SARS-CoV-2 test. The individual WHO-5 and modified EQ-5D scores were both averaged to obtain the overall wellbeing (a) and QoL (b) scores, respectively. Overall, *PS* individuals reported more issues with wellbeing (c-g) and QoL (h-l) than the control cohorts. Error bars indicate standard errors.

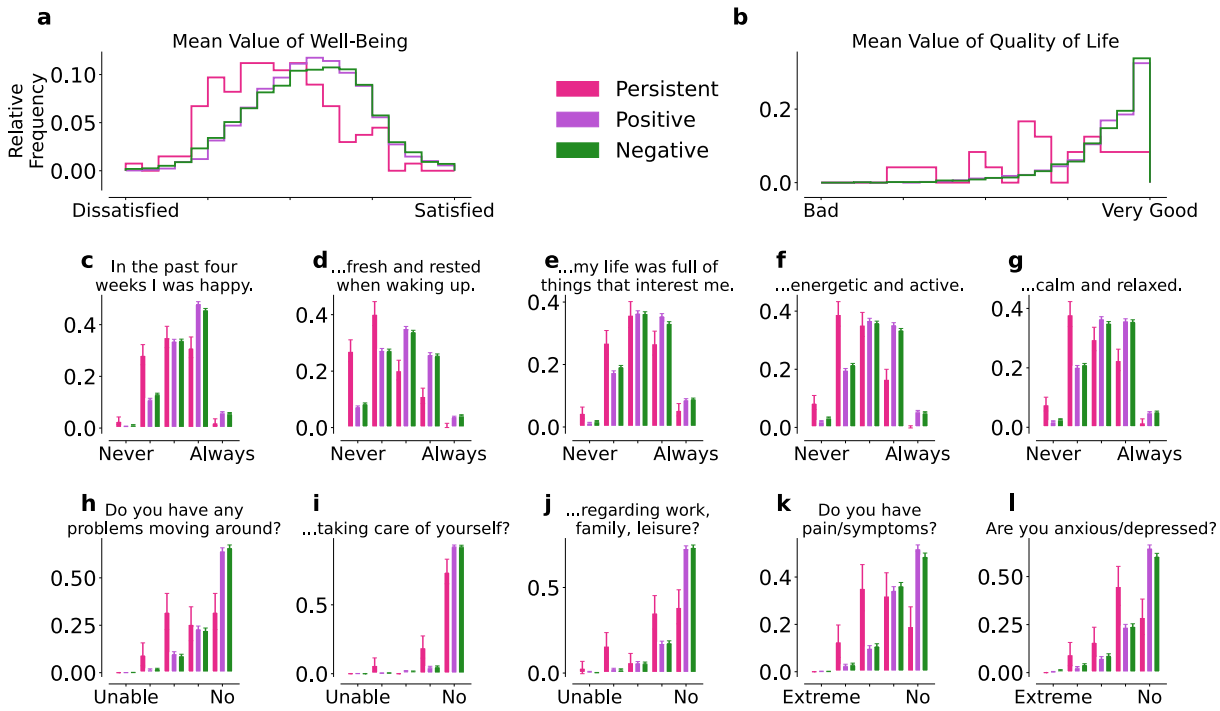

Figure 7. WHO-5 wellbeing (a,c-g) and modified EQ-5D/QoL (b,h-l) for the *PS* (pink), *PC* (purple), and *NC* (green) cohorts after the reported SARS-CoV-2 test. The individual WHO-5 and modified EQ-5D scores were both averaged to obtain the overall wellbeing (a) and QoL (b) scores, respectively. Overall, *PS* individuals reported more issues with wellbeing (c-g) and QoL (h-l) than the control cohorts. Error bars indicate standard errors.

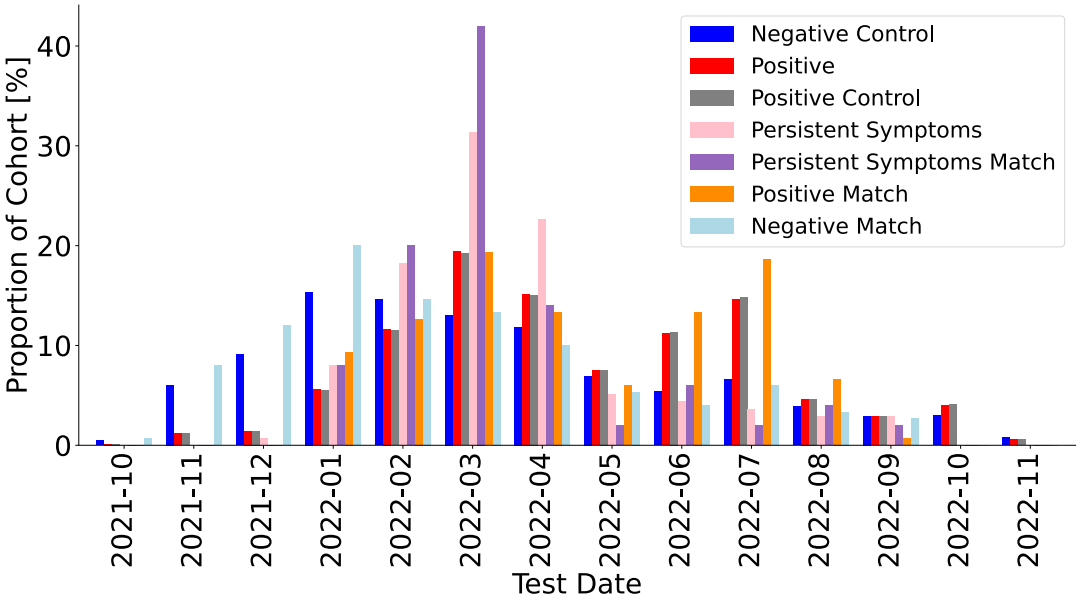

Figure 8. Distribution of reported SARS-CoV-2 testdates of studied cohorts.