## Symptom experience before vs. after confirmed SARS-CoV-2 infection: a population and case control study using prospectively recorded symptom data.

## Authors

Carole H. Sudre, PhD<sup>1,2,3</sup><br>Michela Antonelli, PhD<sup>1</sup><br>Nathan J Cheetham, PhD<br>Erika Molteni, PhD<sup>1</sup><br>Liane S Canas, PhD<sup>1</sup> Michela Antonelli, PhD<br>Nathan J Cheetham, PhD<br>Erika Molteni, PhD<sup>1</sup><br>Liane S Canas, PhD<sup>1</sup><br>Vicky Bowyer, MSc<sup>4</sup>

Nathan J Cheetham, PhD\*<br>Erika Molteni, PhD<sup>1</sup><br>Liane S Canas, PhD<sup>1</sup><br>Vicky Bowyer, MSc<sup>4</sup><br>Ben Murray, MSc<sup>1</sup> Erika Molteni, PhD<sup>+</sup><br>Liane S Canas, PhD<sup>1</sup><br>Vicky Bowyer, MSc<sup>4</sup><br>Ben Murray, MSc<sup>1</sup><br>Khaled Rjoob, PhD<sup>2</sup>  $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ Liane S Canas, PhD<sup>+</sup><br>Vicky Bowyer, MSc<sup>4</sup><br>Ben Murray, MSc<sup>1</sup><br>Khaled Rjoob, PhD<sup>2</sup><br>Marc Modat, PhD<sup>1</sup>

Ben Murray, MSc<sup>+</sup><br>Khaled Rjoob, PhD<br>Marc Modat, PhD<sup>1</sup><br>Joan Capdevila Pu<br>Christina Hu <sup>5</sup>  $\frac{1}{2}$ 

Ben Murray, MSc<sup>1</sup><br>Khaled Rjoob, PhD<sup>2</sup><br>Marc Modat, PhD<sup>1</sup><br>Joan Capdevila Pujol <sup>5</sup>  $\overline{a}$ Khaled Rjoob, PhD<sup>2</sup><br>Marc Modat, PhD<sup>1</sup><br>Joan Capdevila Pujo<br>Christina Hu<sup>5</sup><br>Jonathan Wolf<sup>5</sup> Marc Modat, PhD<sup>+</sup><br>Joan Capdevila Puj<br>Christina Hu <sup>5</sup><br>Jonathan Wolf <sup>5</sup><br>Tim D Spector, Pro  $\frac{1}{2}$ 

Joan Capdevila Pujol<br>Christina Hu<sup>5</sup><br>Jonathan Wolf<sup>5</sup><br>Tim D Spector, Prof<sup>4</sup><br>Alexander Hammers, I

Christina Hu<br>Jonathan Wol<br>Tim D Spector<br>Alexander Har<br>Claire J Steves Jonathan Wolf<br>Tim D Spector, F<br>Alexander Hamr<br>Claire J Steves, F<br>Sebastien Ourse Tim D Spector, Prof<br>Alexander Hammers,<br>Claire J Steves, Prof<sup>4,</sup><br>Sebastien Ourselin, F<br>Emma L Duncan, Pro Alexander Hammers, MD, PhD<sup>4, 9</sup><br>Claire J Steves, Prof<sup>4, 7</sup><br>Sebastien Ourselin, Prof<sup>4</sup><br>Emma L Duncan, Prof<sup>4,8,</sup><br>Affiliations Claire J Steves, Prof $\tilde{a}$ <br>Sebastien Ourselin, Prof $^2$ <br>Emma L Duncan, Prof $^2$ <br>**Affiliations**<br>1. School of Biom<br>2. MRC Unit for L

- 
- Sebastien Ourselin, Prof<sup>4,8,</sup><br>Emma L Duncan, Prof<sup>4,8,</sup><br>**Affiliations**<br>1. School of Biomec<br>2. MRC Unit for Life<br>Experimental Me  $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ Emma L Duncan, Prof <sup>4,8</sup>,<br>**Affiliations**<br>1. School of Biomee<br>2. MRC Unit for Life<br>Experimental Me<br>3. Centre for Medic 1. Sch<br>2. MR<br>2. MR<br>Exp<br>3. Cen<br>4. Dep
	- 1. School of Biomedical Engineering & Imaging Sciences, King's College London, London, UK.<br>
	2. MRC Unit for Lifelong Health and Ageing at UCL, Department of Population Science and<br>
	Experimental Medicine, University College Experimental Medicine, University College London, London, UK<br>
	2. Centre for Medical Image Computing, Department of Computer Science, University Colle<br>
	1. London, London, UK<br>
	2. Department of Twin Research and Genetic Epide Experiment for Medical Image Computing, Department of Computer<br>London, London, UK<br>Department of Twin Research and Genetic Epidemiology, King's<br>ZOE Limited London, UK<br>Guy's and St Thomas' PET Centre, Guy's and St Thomas' NH 3. Centre for Medical Image Computer Computer Computer States College London, London, UI<br>3. Computer of Twin Research and Genetic Epidemiology, King's College London, London, UI<br>3. Guy's and St Thomas' PET Centre, Guy's an
	- Department of Twin<br>ZOE Limited London,<br>Guy's and St Thomas<br>Department of Agein<br>Department of Endoo
	-
	- 6. Guy's and St Thomas' PE<br>7. Department of Ageing ar<br>8. Department of Endocrino<br>words
	- 5. ZOE Limited London, UK<br>6. Guy's and St Thomas' PET Centre, Guy's and St Thomas' NHS Foundation trust, London, UK.<br>7. Department of Ageing and Health, Guy's and St Thomas' NHS Foundation trust, London, UK.<br>8. Department
- 7. Department of Ageing and Health, Guy's and St Thomas' NHS Foundation trust, London, UK.<br>8. Department of Endocrinology, Guy's and St Thomas' NHS Foundation Trust, London, UK.<br>**Key words**<br>COVID-19 symptoms; long COVID; p 7. Department of Endocrinology, Guy's and St Thomas' NHS Foundation Trust, London, UK.<br>8. Department of Endocrinology, Guy's and St Thomas' NHS Foundation Trust, London, UK.<br>1D-19 symptoms; long COVID; post-COVID.

## Key words

8. Beptartment of Endocrinology, Guy's and St Thomas (MD-19 symptoms; long COVID; post-COVID.<br>The Symptoms; long COVID; post-COVID.<br>The count

## Word count

Abstract: 256<br>Number of figures: 8<br>Number of tables: 1 Abstract: 256<br>Number of figures: 8

Number of tables: 1

#### Abstract

#### Background:

## Methods

Some marviadals experience prolonged illness duration.<br> **Methods**<br>
Survival analysis was performed in adults (n=23,452) with community-managed SARC-CoV-2<br>
infection prospectively self-logging data through the ZOE COVID Sym infection symptoms affected post-COVID illness duration.<br>**Methods**<br>Survival analysis was performed in adults (n=23,452) with<br>infection prospectively self-logging data through the ZOE (<br>weekly, from 8 weeks before to 12 wee Survival analysis was performed in adults (n=23,452) with community-managed SARC-CoV-2<br>infection prospectively self-logging data through the ZOE COVID Symptom Study app, at least<br>weekly, from 8 weeks before to 12 weeks aft infection prospection, and the Logging and the Z<sub>O</sub>U (D-19 onset, conditioned on presence vs.<br>absence of baseline symptoms (4-8 weeks before COVID-19). A case-control study was perform<br>1350 individuals with long illness (≥ weekly, from 8 weekly here are the 12 weeks before COVID-19). A case-control study was perfor<br>1350 individuals with long illness (≥8 weeks, 906 [67.1%] with illness ≥12 weeks), matched 1:2<br>age, sex, body mass index, testin 1350 individuals with long illness ( $\geq$ 8 weeks, 906 [67.1%] with illness  $\geq$ 12 weeks), matched 1:1 (for<br>age, sex, body mass index, testing week, prior infection, vaccination, smoking, index of multiple<br>deprivation) wi 1350 individuals with short life in the specific order in the specific state of multiple<br>deprivation) with 1350 individuals with short illness (<4 weeks). Baseline symptoms were compare<br>between the two groups; and against

age, sex, body mass individuals with short illness (<4 weeks). Baseline symptoms were comp<br>between the two groups; and against post-COVID symptoms.<br>Findings: Individuals reporting baseline symptoms had longer post-COVID sy deprediently that 1350 individuals with short COVID symptoms.<br>
Findings: Individuals reporting baseline symptoms had longer post-COVID symptom duration (from<br>
10 to 15 days) with baseline fatigue nearly doubling duration. Findings: Individuals reporting baseline symptoms had longer<br>10 to 15 days) with baseline fatigue nearly doubling duration.<br>individuals with long illness were asymptomatic beforehand.<br>symptoms, vs. 255 (18.9%) of 1350 ind Finangs: Individuals reporting baseline symptoms had longer post-COVID symptom duration (from<br>10 to 15 days) with baseline fatigue nearly doubling duration. Two-thirds (910 of 1350 [67.4%]) of<br>individuals with long illnes 10 to 15 days) with baseline fatigue nearly doubling duration. Two-thirds (910 of 1350 [67.4%]) of<br>individuals with long illness were asymptomatic beforehand. However, 440 (32.6%) had baseline<br>symptoms, vs. 255 (18.9%) of symptoms, vs. 255 (18.9%) of 1350 individuals with short illness (p<0.0001). Baseline symptoms<br>increased the odds ratio for long illness (2.14 [CI: 1.78; 2.57]). Prior comorbidities were more<br>common in individuals with lo increased the odds ratio for long illness (2.14 [CI: 1.78; 2.57]). Prior comorbidities were more<br>common in individuals with long vs. short illness. In individuals with long illness, baseline<br>symptomatic (vs. asymptomatic) common in individuals with long vs. short illness. In individuals with long illness, baseline<br>symptomatic (vs. asymptomatic) individuals were more likely to be female, younger, and have prior<br>comorbidities; and baseline an symptomatic (vs. asymptomatic) individuals were more likely to be female, younger, and<br>comorbidities; and baseline and post-acute symptoms and symptom burden correlated st<br>Interpretation: Individuals experiencing symptoms symptomatic (vs. asymptomatic) individuals in the interestinct) is a seriestinct, younger, and have prior comorbidities; and baseline and post-acute symptoms and symptom burden correlated strongly.<br>Interpretation: Individu Interpretation: Individuals experiencing symptoms before COVID-19 have longer illness duration<br>increased odds of long illness. However, many individuals with long illness are well before SARS-C<br>2 infection. Interpretation: Individuals experiencing symptoms before COVID-19 have longer imress duration and<br>increased odds of long illness. However, many individuals with long illness are well before SARS-CoV-<br>2 infection. 2 infection.

## Introduction

SARS-CoV-2 has infected more than half a billion individuals to date. Individuals of older age, male<br>sex, and with prior comorbidities have poorer outcomes after acute infection, including higher rates<br>of hospitalization a sex, and mortality [1]. many individuals hospitalized for COVID-19 experience<br>protracted convalescence [2,3] particularly individuals requiring ventilatory support; a longitudinal<br>UK study found the majority (71%) were not protracted convalescence [2,3] particularly individuals requiring ventilatory support; a longi<br>UK study found the majority (71%) were not fully recovered six months post-discharge [4].<br>Imany community-managed individuals a protracted convariations (2,3) particularly individuals requiring ventilatory supportions of the majority (71%) were not fully recovered six months post-discharge [4]. Howeve many community-managed individuals also report many community-managed individuals also report protracted post-acute illness. An early<br>community-based study found 13.3% experienced illness beyond 4 weeks, and 2.8% beyond 12<br>weeks, with longer duration associated with fe many community-based study found 13.3% experienced illness beyond 4 weeks, and 2.8% beyo<br>weeks, with longer duration associated with female sex, older age, more severe acute illn<br>prior comorbidities [3]. In the UK, ongoing community-based study found 2010 in performanced study from 13.3% provided illness, are prior comorbidities [3]. In the UK, ongoing symptomatic COVID-19 (OCS) and the Post COVID-19 syndrome (PCS) are defined as otherwise-u prior comorbidities [3]. In the UK, ongoing symptomatic COVID-19 (OCS) and the Post COVID-19<br>syndrome (PCS) are defined as otherwise-unexplained symptoms and signs for 4-12 weeks (OSC), after an acute illness attributable prior comorbidities [3]. In the UK, ongoing symptomate COVID-19 (COU) more than 12 weeks (OSC) more than 12 weeks (PCS), after an acute illness attributable to SARS-CoV-2 infection [5] with so variation internationally in syndrome (PCS), are defined as otherwise analytical symptoms and signs for  $T = 12$  week (PCS), after an acute illness attributable to SARS-CoV-2 infection [5] with some variation internationally in terminology and symptom

variation internationally in terminology and symptom duration (e.g., 8 vs. 12-week threshold) [6]).<br>PCS prevalence estimates vary substantially. A recent meta-analysis [7] highlighted the<br>heterogeneity of published studie PCS prevalence estimates vary substantially. A recent meta-analysis [7] highlighted the<br>heterogeneity of published studies ( $l^2$  = 100%) with widely differing prevalence estimates (9%-81%<br>varying globally, regionally, and PERTURENCE EDMINISTRIP, THE EMPTON META-META-ANDRETIC MAP, THE INDEEDED META-<br>Photography and by the meta-analysis (PCS prevalence estimates der<br>predominantly hospitalized cohorts. In March 2023 the UK Office of National heterogeneity of published studies ( $I = 100\%$ ) with widely differing prevalence estimates (9%-81%),<br>varying globally, regionally, and by hospitalization status. PCS prevalence estimates derive from<br>predominantly hospitali predominantly hospitalized cohorts. In March 2023 the UK Office of National Statistics estimated 1.9<br>million citizens (2.9% of the population) had self-reported long COVID (defined as symptoms for<br>more than four weeks afte million citizens (2.9% of the population) had self-reported long COVID (defined as symptoms for<br>more than four weeks after either test-positive or suspected SARS-CoV-2 infection)[8]. An earlier UK<br>study reported prevalence million chain four weeks after either test-positive or suspected SARS-CoV-2 infection][8]. An earlier<br>study reported prevalence of PCS of 1.2%-4.8% in test-positive individuals, considering only<br>symptoms that limited day-t study reported prevalence of PCS of 1.2%-4.8% in test-positive individuals, considering only<br>symptoms that limited day-to-day functioning [9]. Neither study included a control group. In<br>contrast, a large UK primary-care st symptoms that limited day-to-day functioning [9]. Neither study included a control group. In<br>contrast, a large UK primary-care study comparing community-managed adults with contrast, a large UK primary-care study comparing sum transit, a large UK primary-care study comparing community-managed adults with confirent SARS-CoV-2 infection to a matched contemporaneous cohort reported symptom prevalen<br>(here, at least one symptom) at 12 weeks of 5. contrast, a large UK primary-care study comparing community-managed adults with confirmed<br>SARS-CoV-2 infection to a matched contemporaneous cohort reported symptom prevalence<br>(here, at least one symptom) at 12 weeks of 5.4 SARS-CoV-2 infection to a matched contemporaneous cohort reported symptom prevalence (here, at least one symptom) at 12 weeks of 5.4% in infected vs. 4.3% in uninfected individuals  $[10]$ .

Infections, including bacteria (e.g., Borrelia) viruses (e.g., Epstein-Barr Virus) and parasites (e.g.,<br>Giardia), which share many characteristics including fatigue, exertional intolerance, and<br>neurocognitive symptoms ('br Giardia), which share many characteristics including fatigue, exertional intolerance, and<br>neurocognitive symptoms ('brain fog') (recently comprehensively reviewed [11]). Post-acute<br>infection syndromes are more common in fe Etherica, which share many characteristics including targety such share meaning that<br>infection syndromes are more common in females and younger individuals, though a rel<br>with initial illness severity is less clear [11]. Fe infection syndromes are more common in females and younger individuals, though a relation<br>with initial illness severity is less clear [11]. Few studies assess pre-morbid risk factors for post-<br>syndromes prospectively [12] infection symplements are more common termined and younger individuals, in eigenstanding<br>with initial illness severity is less clear [11]. Few studies assess pre-morbid risk factors for post-acu<br>syndromes prospectively [12 syndromes prospectively [12] despite possible influence of premorbid conditions on post-acute<br>illness symptomatology [13]. Postulated non-exclusive mechanisms include remnant infection,<br>4 illness symptomatology [13]. Postulated non-exclusive mechanisms include remnant infection,<br>illness symptomatology [13]. Postulated non-exclusive mechanisms include remnant infection, illness symptomatology [13]. Postulated non-exclusive mechanisms include remnant infection,

autoimmunity induction, and/or maladaptive tissue repair; however, for most affected individuals,<br>definitive pathophysiology is unclear despite extensive investigations. Whether similar processes<br>underpin PCS, and whether

underpin PCS, and whether common to all PCS individuals, is unclear [14].<br>The ZOE COVID Symptom Study began in March 2020 with participating adults logging their healt<br>data contemporaneously across the pandemic. Thus, symp The ZOE COVID Symptom Study began in March 2020 with participating ad<br>data contemporaneously across the pandemic. Thus, symptoms could be a<br>longitudinally in individuals subsequently contracting SARS-CoV-2, and irre<br>profil data contemporaneously across the pandemic. Thus, symptoms could be assessed prospectively at<br>longitudinally in individuals subsequently contracting SARS-CoV-2, and irrespective of ultimate illne<br>profile.<br>We hypothesised t data contemporance are pandemic. Thus, symptoms could be above a properties and the illness<br>profile.<br>We hypothesised that symptoms and comorbidities before SARS-CoV-2 infection might contribute to<br>post-acute symptomatology

longitudinal, in individuals subsequently subsequently experience of unintersement profile.<br>We hypothesised that symptoms and comorbidities before SARS-CoV-2 infection might contribute to<br>post-acute symptomatology, includi profile.<br>We hypothesised that symptoms and comorbidities before SARS-CoV-2 infection might contribute to<br>post-acute symptomatology, including illness duration. We assessed:<br>A) symptoms reported before COVID-19, in individu

- Metal symptomatology, including illness duration. We assessed:<br>
A) symptoms reported before COVID-19, in individuals subsequently experiencing long vs. short<br>
illness;<br>
B) symptom correlation before and after SARS-CoV-2 in examine symptomatology, including initial that the subseque<br>
illness;<br>
B) symptom correlation before and after SARS-CoV-2 infection; a<br>
C) whether prior symptoms and comorbidities affect post-COVID A) symptoms reported before COVID-19, in individuals subsequently inperioding and individuals subsequently experience COVID-19, in the COVID-19, in the COVID-19, whether prior symptoms and comorbidities affect post-COVID i
	-
	- iversified<br>interface<br>s<br>s B) whether prior symptoms and comorbidities affect post-COVID illn<br>
	shods<br>
	. C) whether prior symptoms affect post-COVID in the prior symptoms affect post-COVID in the profile. The prior symptoms are computed to the profile of  $\mathcal{C}$  in the prior symptom of  $\mathcal{C}$  is a computed to the profile.

## Methods

## **Cohort**

 $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ The Ltd, King's College London, Massachusetts General Hospital, Lund University and Uppsala<br>University (ethics approval: KCL ethics committee REMAS no. 18210, review reference LRS-19/20–<br>18210, with all individuals providi University (ethics approval: KCL ethics committee REMAS no. 18210, review reference LRS-19/<br>18210, with all individuals providing informed consent for use of their data in COVID-19 resear<br>registration). After initial loggi University (ethics approximates committee Remark of Lethical, COVID-19 research a<br>
registration). After initial logging of baseline demographics including comorbidities, (Suppl. Table :<br>
participants were prompted daily to 1821-19 mm annumed providing incomponent of the control and monoidities, (Suppl. Table S1<br>participants were prompted daily to report any symptoms (direct questions and free text (Suppl.<br>Table S1)), SARS-CoV-2 testing and r registration). After initial logging of baseline demographics including comorbidities, (Suppl. Table 31)<br>participants were prompted daily to report any symptoms (direct questions and free text (Suppl.<br>Table S1)), SARS-CoVparticipants were prompted daily to report any symptoms (direct questions and rice text (Supplit<br>Table S1)), SARS-CoV-2 testing and results, and vaccination(s) using a phone-based app. Data<br>collection expanded on 4 Novembe Table S1)), SARS-COV-2 testing and results, and vaccination(s) using a phone-based app. Data<br>collection expanded on 4 November 2020 to include more direct symptom questions. The col<br>was surveyed regarding pre-pandemic ment was surveyed regarding pre-pandemic mental health diagnoses in February-April 2021 [15] (Supp<br>Table S2). The current dataset was cut on 30 May 2022 with symptom assessment altering the<br>following day.<br>As previously [3], COV was surveyed regarding pre-pandemic mental health diagnoses in rebruary-April 2021 [15] (Suppl.<br>Table S2). The current dataset was cut on 30 May 2022 with symptom assessment altering the<br>following day.<br>As previously [3], C

Table S2). The current dataset was cut on 30 May 2022 with symptom assessment altering the<br>following day.<br>Table S1) commencing between 14 days before and 7 days after a self-reported positive polym<br>chain reaction (PCR) or Forening any.<br>As previously [<br>**Table S1**) coming<br>chain reaction<br>subsequent illr Table S1) commencing between 14 days before and 7 days after a self-reported positive polymerase<br>chain reaction (PCR) or lateral flow antigen test (LFAT). For individuals with multiple positive tests,<br>subsequent illness p Table S1) commencing between 14 days before and 7 days after a sen-reported positive polymerase<br>chain reaction (PCR) or lateral flow antigen test (LFAT). For individuals with multiple positive tests,<br>subsequent illness pr subsequent illness profiles were defined for tests spaced >90 days apart. If  $\leq$ 90 days apart, illness subsequent illness profiles were defined for the stage spaced  $\sim$  10 days apart. If  $\sim$ 90 days apart. If  $\sim$ 

profile was defined according to the first positive test. Illness duration was calculated from first<br>symptomatic day until return to asymptomatic (i.e., logging as healthy) [3,16]. Consideration was<br>given to possible right siven to possible right censoring in duration calculation (ongoing illness at final data censoring;<br>logging discontinuation while still symptomatic). To calculate illness duration attributable to acute<br>SARS-CoV-2 infection logging discontinuation while still symptomatic). To calculate illness duration attributable to act<br>SARS-CoV-2 infection, individuals were required to log as healthy for at least one week immedia<br>before COVID-19 commenceme SARS-COVID-19 commencement [3,16]. If symptoms were again logged within one week of a<br>healthy report, illness was considered ongoing, thus allowing for illness fluctuation [3,16].<br>The baseline period was defined as 4 to 8

healthy report, illness was considered ongoing, thus allowing for illness fluctuation [3,16].<br>The baseline period was defined as 4 to 8 weeks before, and the post-COVID period as 8-12 weeks bealthy report, illness was considered ongoing, thus allowing for illness fluctuation [3,16].<br>The baseline period was defined as 4 to 8 weeks before, and the post-COVID period as 8-12 we<br>after, COVID-19 onset. To ensure co he baseline period was defined as 4 to 8 weeks before, and the post-COVID period as 8-1.<br>after, COVID-19 onset. To ensure consistent assessment in the entire cohort for the necess<br>weeks, data were constrained to individual after, COVID-19 onset. To ensure consistent assessment in the entire cohort for the necessary 20<br>weeks, data were constrained to individuals in whom COVID-19 commenced between 30 Decembe<br>2020 (8 weeks after 4 November 2020 after, COVID-19 commenced between 30 December 2020 (8 weeks, data were constrained to individuals in whom COVID-19 commenced between 30 December 2020 (8 weeks after 4 November 2020, date of symptom question expansion) and weeks, data were constrained to individuals in whom COVID-19 commenced between 30<br>2020 (8 weeks after 4 November 2020, date of symptom question expansion) and 2 Marc<br>weeks before 1 June 2022, date of symptom assessment alt h 2022 (12<br><sub>Is</sub><br>eriod of<br>rders that weeks before 1 June 2022, date of symptom assessment alteration). Symptom burden was<br>calculated as number of individual symptoms reported at least once during the defined period of<br>assessment. Self-reported mental health d weeks before 1 June 2022, and the property alternative interation, property alternative rate<br>calculated as number of individual symptoms reported at least once during the defined per<br>assessment. Self-reported mental health assessment. Self-reported mental health diagnoses were considered overall, and for disorders that<br>can include psychosis (Suppl. Table S2).<br>Short illness was defined as <4 weeks and long illness as ≥8 weeks.<br>Inclusion crite can include psychosis (Suppl. Table S2).<br>Short illness was defined as <4 weeks and long illness as ≥8 weeks.<br>Inclusion criteria were:

- can include psychosis (Suppl. Table S2).<br>Short illness was defined as <4 weeks and long illness as ≥8 weeks.<br>Inclusion criteria were:<br>a) self-reporting UK adults presenting with PCR- or LFAT-confirmed COVID-19, between 30 S<br>Short inclusion criteria were:<br>Short in the self-reporting UK adults presenting with PCR- or LFAT-confi<br>December 2020 and 2 March 2022.<br>Short long in the self once weekly, from ≥8 weeks before until ≥1
- a) self-reporting U<br>December 2020<br>b) logging at least<br>commencemen Becember 2020 and 2 March 2022.<br>
a) logging at least once weekly, from ≥8 weeks before until ≥12 weeks after COVID-19<br>
commencement.<br>
c) logging as healthy in the week before COVID-19 commencement.<br>
d) co-morbidity and de Degging at least once weekly, from ≥<br>Commencement.<br>logging as healthy in the week befor<br>co-morbidity and demographic data<br>the mental health survey
	-
	- b) logging at least once were week before COVID-19 commencement.<br>
	b) logging as healthy in the week before COVID-19 commencement.<br>
	d) co-morbidity and demographic data logged at registration, with subsequent particip<br>
	the Economization<br>
	logging as healthy<br>
	co-morbidity and<br>
	the mental health<br>
	on to non-adhere

d) co-morbidity and demographic data logged at registration, with su<br>the mental health survey.<br>ddition to non-adherence to inclusion criteria, individuals vaccinated v<br>c-COVID periods or one week before these periods were d) co-mortal health survey.<br>
ddition to non-adherence to inclusion criteria, individuals vaccinated within either baseline or<br>
:-COVID periods or one week before these periods were excluded, given symptom overlap<br>
veen vac on to non-adherence to in<br>VID periods or one week b<br>n vaccination side-effects a<br>aining individuals, a Cox m In addition to non-addition to include the periods were excluded, given symptom overlap<br>between vaccination side-effects and COVID-19 [17].<br>For remaining individuals, a Cox model was performed, adjusting for demographic cr

postmeen vaccination side-effects and COVID-19 [17].<br>For remaining individuals, a Cox model was performed, adjusting for demographic criteria inclu<br>age, sex, body mass index (BMI), week of test, number of test-positive SAR For remaining individuals, a Cox model was performe<br>age, sex, body mass index (BMI), week of test, numbe<br>smoking status, index of multiple deprivation (IMD), a<br>effect of any symptom present at baseline on median For remaining individuals, a Commodel was performed, any adjusting of themographic criteria including<br>age, sex, body mass index (BMI), week of test, number of test-positive SARS-CoV-2 infections,<br>smoking status, index of m anoking status, index of multiple deprivation (IMD), and number of vaccinations, evaluating the<br>effect of any symptom present at baseline on median duration overall, and for each symptom<br>individually. effect of any symptom present at baseline on median duration overall, and for each symptom individually. effect of any symptom present at baseline on median duration overall, and for each symptom<br>individually. individually.

Selected first. Individuals with short illness (<4 weeks) were then selected, matched 1:1 per<br>previously listed demographic criteria using the Hungarian algorithm [18], minimizing the Euclidean<br>distance cost and ensuring e previously listed demographic criteria using the Hungarian algorithm [18], minimizing the Eu<br>distance cost and ensuring equal weighting across all characteristics (normalising baseline v:<br>before matching); without replacem previously interacting depine anticia using the Hungarian algorithm [19], minimizing the Euclidean<br>distance cost and ensuring equal weighting across all characteristics (normalising baseline variables<br>before matching); wit

before matching); without replacement for controls.<br>Data were compared across groups using the McNemar test for counts and Wilcoxon signed rank<br>tests for continuous variables.<br>Using conditional logistic regression models,

before matching); Data were compared across groups using the McNem<br>tests for continuous variables.<br>Using conditional logistic regression models, we assem<br>to baseline symptom presentation (considered overa Late were compared across groups using the McNemar test of the McNemar theorem, given rank<br>tests for continuous variables.<br>Using conditional logistic regression models, we assessed the odds of long illness duration accordi Using conditional logistic regre<br>to baseline symptom presentat<br>individual symptom) using thre<br>1) Model 1 – no adjustme Using conditional regression models, we assessed the odds of an american models of adjustment and set and for each individual symptom) using three levels of adjustment for covariates:<br>
1) Model 1 – no adjustment<br>
2) Model

- 
- to baseline symptom) using three levels of adjustment for covariates:<br>
1) Model 1 no adjustment<br>
2) Model 2: As for Model 1, with additional adjustment for presence of any comorbidity logged<br>
at registration (allergic rh individual symptom) and given the levels of any administrator to contain the levels of adjustment for pressors<br>2) Model 2: As for Model 1, with additional adjustment for pressors at registration (allergic rhinitis [hay fev 2) Model 2: As for Model 1,<br>at registration (allergic rhi<br>lung disease, asthma)<br>3) Model 3: As for Model 2, 2) at registration (allergic rhinitis [hay fever], cancer, diabetes, kidney disease, heart disease,<br>lung disease, asthma)<br>3) Model 3: As for Model 2, with additional adjustment for prior mental health diagnosis.<br>2) Model 3
	-

at registration (and given rhinitis [hay] feveral), cancer, manifestively measure, measure, heart and sum all<br>Model 3: As for Model 2, with additional adjustment for prior mental health diagnosis.<br>tigate relationships betw Model 3: As for Model<br>tigate relationships be<br>tigate relationships be<br>ncing each individual sy 3) Model 2: And a displayment displayment for procedure the diagnosis.<br>Notes figate relationships between baseline and post-COVID symptoms, we assessed the odd<br>Priencing each individual symptom in the post-COVID period acc Experiencing each individual symptom in the post-COVID period according to its presence at<br>baseline, separately in individuals with long and with short illness, using logistic regression models<br>with the three levels of adj baseline, separately in individuals with long and with short illness, using logistic regression m<br>with the three levels of adjustment detailed above, and adjusting for previously listed demog<br>variables. We further investig baseline, separate levels of adjustment detailed above, and adjusting for previously listed demographic<br>variables. We further investigated possible sex-based difference in reporting at baseline and in the<br>post-COVID period wariables. We further investigated possible sex-based difference in reporting at baseline and in the<br>post-COVID period according to duration group given previous evidence of sex differences in post-<br>acute infection syndrom variables. We further investigated positive since the interfaction of sex-differences in post-<br>acute infection syndromes [11], and OSC/PCS [8]. For symptoms with a severity scale (fatigue and<br>dyspnoea), we compared severit symptom separately for long and short illness groups using summertime (May-September) as dyspnoea), we compared severity during baseline and post-COVID periods in individuals reporting<br>these symptoms at baseline; we also considered dyspnoea specifically in individuals with prior<br>asthma/lung disease. Seasonal e dysposition, we compared severity during discussion provides in position individuals with prior asthma/lung disease. Seasonal effects on symptom reporting were assessed for each individual<br>symptom separately for long and s asthma/lung disease. Seasonal effects on symptom reporting were assessed for each individua<br>symptom separately for long and short illness groups, using summertime (May-September) as<br>reference. As some co-morbidities exhibi symptom separately for long and short illness groups, using summertime (May-September) as<br>reference. As some co-morbidities exhibit symptom seasonality (e.g., allergic rhinitis) adjustme<br>previously listed demographic varia symptom separately for long and short inherent groups, and same lines (e.g., allergic rhinitis) adjustm<br>previously listed demographic variables and Model 3 variables was applied.<br>For individuals with long illness, we inves

previously listed demographic variables and Model 3 variables was applied.<br>For individuals with long illness, we investigated demographic differences according to baseline<br>symptomatology (≥1 symptom at baseline) using Chipreviously listed demographic differences are<br>previously listed demographic differences and<br>symptomatology (21 symptom at baseline) using Chi-squared test for catego<br>U-Test for continuous variables. We investigated the rel symptomatology ( $\geq 1$  symptom at baseline) using Chi-squared test for categorical and Mann Whi<br>U-Test for continuous variables. We investigated the relationship between baseline and post-CC<br>symptom burden using linear r symptomatology (≥2 symptom at baseline) using Chi-square decrease large into analyticity, increased U-Test for continuous variables. We investigated the relationship between baseline and post-COVID symptom burden using lin U-TEST FORMALES THANKER TEST FOR CONTENDANT PRODUCED THE RELATIONSHIP BETWEEN BASELINE AND provide a<br>symptom burden using linear regression, adjusting for demographic variables. symptom burden using linear regression, adjusting for demographic variables.

For baseline comorbidities that differed between individuals with long vs. short illness (Table 1), we assessed whether presence of any of these comorbidities affected baseline symptom presentation and overall post-COVID s and overall post-COVID symptom burden, using Chi-squared and Mann-Whitney U tests respectivel<br>We investigated odds for individual baseline and post-COVID symptoms in individuals with long<br>illness according to these comorbi and overall post-COVID symptoms in individuals with long<br>We investigated odds for individual baseline and post-COVID symptoms in individuals with long<br>illness according to these comorbidities, using a logistic regression m Uness according to these comorbidities, using a logistic regression model adjusted for demogra<br>Variables.<br>We repeated the conditional logistic regression analysis and other analyses for individuals with l<br>Uness, with minim

in the according to these comorbidities, using a logistic regression  $\frac{1}{2}$ . Then in the largest also according to the variables.<br>We repeated the conditional logistic regression analysis and other analyses for individua We repeat<br>illness, wit<br>of less assi<br>False disco We repeated the conditional regression analysis and control analysis conditions, illness, with minimal logging frequency of at least fortnightly (vs. weekly), thus allowing for inclusion of less assiduous loggers and longe

illess assiduous loggers and longer periodicity of symptom fluctuation.<br>False discovery rate adjustment using the Benjamini-Hochberg procedure was applied across all<br>tested symptoms, in each analysis. False discovery rate adjustment using the Benjamini-Hochberg procedur<br>tested symptoms, in each analysis.<br>. False discovery rate adjustment using the Benjaministed symptoms, in each analysis. tested symptoms, in each analysis.

## Results:

Figure 1 shows participant selection. Table 1 presents the cohort before and after matching.<br>Compared to the overall cohort of test-positive individuals, the selected cohort of regular loggers<br>was older, with slightly more Examples of the slightly more comorbidities. Censoring on mental health survey participation did not affect the cohort greatly except for sex (females more likely to participate); other characteristics<br>were stable (data no affect the cohort greatly except for sex (females more likely to participate); other characteristics<br>were stable (data not shown).<br>The majority (906 of 1350, 67.1%) individuals with long illness had illness duration beyond affect the control of the cohort greatly except for sex (females more likely to participate); once characteristics<br>The majority (906 of 1350, 67.1%) individuals with long illness had illness duration beyond 12 we<br>Impact of

## Impact of baseline symptoms on median illness duration

The majority (906 of 1350, 67.)<br>Impact of baseline symptoms<br>Reporting of any symptom at k The major of baseline symptoms on median illness duration<br>The matricon median illness duration<br>(1008: 13;16] days). Right censoring (unfinished illness) was identified for some long illness individuals<br>due to data censorsh Reporting of any symptom at baseline increased median increasing (i.e.m. 2) (i.e.m. 3) Legister 2) and [IQR: 13;16] days). Right censoring (unfinished illness) was identified for some long illness individual<br>due to data ce

(unit 2) any spanned internal (unit internal internal internal internal internal internal internal internal due<br>to data censorship (n=338) and logging interruption while unhealthy (n=166).<br>Considering individual symptoms, due to data censorship (n=338) and logging interruption while unhealthy (n=166).<br>Considering individual symptoms, baseline reporting of fatigue, headache, sneezing, sore throat, and<br>rhinorrhea increased median illness dura Primorrhea increased median illness duration by 9, 7, 5, 5, and 4 days respectively.<br>
Matched case-control cohort analysis<br>
Comorbidities

## Matched case-control cohort analysis

## **Comorbidities**

rhin china increased median increased median in  $\mathcal{S}_1$ , 7, 7, 9, 9, and 4 days respectively.<br> **Comorbidities**<br>
Considering the matched cohort (n=2700 individuals), 463 individuals reported lung Considering the matched cohort (n=2700 individuals), 463 individuals reported lung disease and/or<br>asthma: 310 (67.0%) reported both, 115 (24.8%) only asthma, and 38 (8.2%) only lung disease. Thus<br>'lung disease' and 'asthma 'lung disease' and 'asthma' categories were jointly considered. In contrast, 1326 individuals reported<br>allergic rhinitis and/or asthma: 317 (23.9%) reported both, 910 (67.9%) only allergic rhinitis, and 108<br>(8.1%) only as (8.1%) only asthma; thus, allergic rhinitis was considered both, 910 (67.9%) only allergic rhinitis, and 108<br>(8.1%) only asthma; thus, allergic rhinitis was considered separately.<br>Individuals with long (vs. short) illness

(8.1%) only asthma; thus, allergic rhinitis was considered separately.<br>
Individuals with long (vs. short) illness were more likely to report co-morbidities of allergic rhinitis<br>
(p<0.001), asthma/lung disease (p<0.001), h (8.1%) only as the long (vs. short) illness were more likely to report co-rodic resolutions (p<0.001), asthma/lung disease (p<0.001), heart disease (p=0.044], doprior mental health diagnosis (p=0.003).<br>Baseline symptoms i Individuals with long (vs. short) illness were more likely to report co-morbidities of allergic rhinities (p=0.037), and/or a<br>prior mental health diagnosis (p=0.003).<br>Baseline symptoms in individuals with short vs. long il

## Baseline symptoms in individuals with short vs. long illness.

(p=0.001), assuming disease (p=0.003),<br>prior mental health diagnosis (p=0.003).<br>Baseline symptoms in individuals with short vs. long illness.<br>Individuals with long (vs. short) illness were more likely to report baseline sy prior mental mental health diagnost in the spaceline symptoms in individuals with spaceling individuals were asymptomatic before COVID illness were asymptomatic before COVID Individuals with long (version) interest with long of the protein since symptoms (16.8%), p-value <0.0001) (Table 1). However, over two-thirds (67.4%) of individuals with long illness were asymptomatic before COVID-19.<br>Log

255 (16.8%), p-value <0.0001) (Table 1). However, over two-thirds (67.4%) of hidividuals with long<br>illness were asymptomatic before COVID-19.<br>Logging frequency during the baseline period did not differ between individuals Logging frequency during the baseline period<br>illness, and only marginally) in the post-COVII<br>short illness). Illness, and only marginally) in the post-COVID period (20 vs. 21 logs for individuals with long vs.<br>short illness).  $\frac{1}{2}$  in the post-COVID period (20 vs. 21 logs for individuals with long vs. 21 logs for individuals with long vs. 21 lon short illness).

Considering individual symptoms at baseline (Figure 2), the five commonest symptoms were the<br>same regardless of ultimate illness duration but were more prevalent in individuals with subsequent<br>long (vs. short) illness: hea same regardless of ultimates of us. short) illness: headache (18.9% vs. 9.4%), fatigue (13.0% vs. 5.6%), sore throat (15.0% vs.<br>8.8%), rhinorrhea (14.3% vs. 8.5%), and sneezing (11.9% vs.6.8%) [descriptive data only].<br>Base

dong (variably material and the term of the computing and called the computer material (19.0%), shorted (18.0%), and sneezing (11.9% vs. 6.8%) [descriptive data only].<br>Baseline symptoms associated with increased odds of lo Baseline symptoms associated with increased odds of long illness.<br>Adjusting for demographic variables, the odds ratio for long illness in symptomatic (vs.<br>asymptomatic) baseline status was 2.14 [CI:1.78; 2.57].

Baseline symptoms associated with increased odds of long illness.<br>Adjusting for demographic variables, the odds ratio for long illness in<br>asymptomatic) baseline status was 2.14 [Cl:1.78; 2.57].<br>Considered per-symptom, with Adjusting formulation in the status was 2.14 [CI:1.78; 2.57].<br>Adjustment considered per-symptom, with identical covariates: reporting of almost any individual state increased the odds of long illness (**Figure 3**). However, Considered per-symptom, with identical covariates: rep<br>baseline increased the odds of long illness (Figure 3). Ho<br>between long illness and baseline cutaneous symptoms<br>adjustment for prior comorbidities (Model 2) and prior baseline increased the odds of long illness (**Figure 3**). However, no evidence of association was seer<br>between long illness and baseline cutaneous symptoms (red welts, 'blisters', alopecia), and, after<br>adjustment for prior baseline increased the odds of long illness (Figure 3). However, no evidence of association was seen<br>between long illness and baseline cutaneous symptoms (red welts, 'blisters', alopecia), and, after<br>adjustment for prior c between long inners and baseliness are graphed (red welts), and the person, and graphed adjustment for prior comorbidities (Model 2) and prior mental health diagnoses (Model 3), anore:<br>('low appetite').<br>**Symptom concordanc** 

## Symptom concordance over time

adjustment for prior (Model 2) and the prior of the symptom concordance over time<br>adjustment for prior compares symptoms during baseline or post-COVID periods, in individuals with long or<br>short illness (categories: stayed ('low appetite').<br> **Symptom concordance over time**<br> **Figure 4** compares symptoms during **baseline or post-COVID periods,** in individuals with long or<br>
short illness (categories: stayed absent/appeared/disappeared/stayed pr Figure 4 compares symptoms during baseline or post-COVID periods, in individuals with long or<br>short illness (categories: stayed absent/appeared/disappeared/stayed present). Importantly, our<br>inclusion criteria meant individ inclusion criteria meant individuals with long illness had at least one symptom during the post-COVID period, whereas individuals with short illness had returned to asymptomatic for at least or<br>week within four weeks of de COVID-19 (though they might subsequently report symptoms<br>again).<br>In individuals with long illness, individual symptoms were more likely in the post-COVID period if<br>present at baseline, with exceptions of some cutaneous man

inclusion criteria meant individuals with long illness had at least one symptom during the postagain).<br>again).<br>In individuals with long illness, individual symptoms were more likely in the post-COVID period if<br>present at baseline, with exceptions of some cutaneous manifestations (Figure 5). Adjusting for<br>baseline co again).<br>In indiv<br>present<br>baselin In the present at baseline, with exceptions of some cutaneous manifestations (**Figure 5**). Adjusting for<br>baseline comorbidities (Model 2) and prior mental health diagnosis (Model 3) made minimal<br>difference.<br>For comparison, present at baseline, with exceptions of some cutaneous manifestations (Figure 5). Adjusting for<br>baseline comorbidities (Model 2) and prior mental health diagnosis (Model 3) made minimal<br>difference.<br>For comparison, and as e

difference.<br>For comparison, and as expected given our inclusion criteria, fewer symptoms were reported<br>post-COVID period in individuals with short illness, considered overall (Suppl. Fig S1) and by k<br>individual symptom (Su For compar<br>post-COVID<br>individual s<br>COVID peri For comparison, and individuals with short illness, considered overall (Suppl. Fig S1) and by baseline<br>
individual symptom (Suppl. Fig S2). Symptoms more likely to be present during baseline and post-<br>
COVID periods in thi post-COVID period in individuals with short illness, considered overall (Suppl. Fig S1) and by baseline<br>individual symptom (Suppl. Fig S2). Symptoms more likely to be present during baseline and post-<br>COVID periods in this Individual symptom (Suppl. Fig S2). Symptoms more incry to be present during baseline and post-<br>COVID periods in this group included several upper respiratory symptoms (tinnitus, sore throat,<br>rhinorrhoea, sneezing, hoarse Thinorrhoea, sneezing, hoarse voice) and some systemic symptoms (headache, lymphadenopath<br>dizziness, myalgia, rigors, and fatigue). Adjusting for baseline comorbidities (Model 2) reduced<br>significance for the rarest symptom dizziness, myalgia, rigors, and fatigue). Adjusting for baseline comorbidities (Model 2) reduced<br>significance for the rarest symptoms; adjusting further for prior mental health diagnosis (Model 3)<br>did not alter these resul dignificance for the rarest symptoms; adjusting further for prior mental health diagnosis (Model 2) reduced id<br>did not alter these results substantially.  $s$  did not alter these results substantially. did not alter these results substantially.

## Symptom reporting by sex.

Symptom prevalence during both baseline and post-COVID periods varied numerically by sex,<br>whether illness was of long or short duration. Most symptoms were more commonly reported by<br>females than males (descriptive data: Fi symptom prevalences appeared least for post-COVID symptoms in individuals with long illness Fable S3 (baseline symptoms) and Suppl. Table S4 (post-COVID symptoms). Sex differences in<br>symptom prevalences appeared least for post-COVID symptoms in individuals with long illness<br>(Figure 6, lower right panel), and for Table 33 (baseline symptoms) and Suppl. Table 34 (post-COVID symptoms). Sex differences in<br>symptom prevalences appeared least for post-COVID symptoms in individuals with long illness<br>(Figure 6, lower right panel), and for symptom prevalences appeared visitive previously symptoms in individuals with short illness (Figure<br>upper left panel). These findings should be interpreted descriptively as we did not formally test<br>an interaction between s (Figure 6, lower right panel), and for baseline symptoms in individuals with short liness (Figure 6, upper left panel). These findings should be interpreted descriptively as we did not formally test for<br>an interaction betw

## Symptom severity over time

upper left panel). These finalligs should be interpreted descriptively as we did not formally test for<br>an interaction between symptoms and sex.<br>Symptom severity over time<br>Considering individuals with long illness and fatig an interaction between symptoms and sex.<br> **Symptom severity over time**<br>
Considering individuals with long illness and fatigue at baseline (176 individuals): 101 (57%) reported<br>
unchanged, 45(26%) improved, and 30 (17%) wor Considering individuals with<br>
long illness with dyspnoea at baseline (35 individuals): 22 (65%) reported unchanged, 11(31%)<br>
improved, and 2 (6%) worsened severity (**Suppl. Table S1; Suppl. Fig. S3**).<br>
For comparison, indi long illness with dyspnoea at baseline (35 individuals): 22 (65%) reported unchanged, 11(31%)<br>improved, and 2 (6%) worsened severity (**Suppl. Table S1; Suppl. Fig. S3**).<br>For comparison, individuals with short illness durat

long illness with dysplaced at baseline (35 individuals): 22 (33), improved, and 2 (6%) worsened severity (**Suppl. Table S1; Suppl. Fig. S3)**.<br>For comparison, individuals with short illness duration showed improvement or r

For comparison, individuals with short illness duration showed improvement symptom if reported at baseline, noting again the bias imposed by our inc<br>Considering the 22 individuals with asthma/lung disease reporting dyspno<br> For comparison, individuals with anti-random short interesting to unrinduction criteria.<br>Considering the 22 individuals with asthma/lung disease reporting dyspnoea at baseline<br>(independent of disease duration): 1 (4.5%) re symptom is a bias in the considering the 22 individuals with asthma/lung disease reporting dyspnoea at baseline<br>(independent of disease duration): 1 (4.5%) reported worsening, 13 (59%) unchanged; are<br>decreasing or resolved Considering the 22 individuals with asthma, rang at taxos reporting a, species at a at successing<br>(independent of disease duration): 1 (4.5%) reported worsening, 13 (59%) unchanged; an<br>decreasing or resolved dyspnoea.<br>Base

## Baseline symptoms and seasonality

(independent of disease duration): 1 (institution): 1 (institution): 1 (1997) unchanged; and a correlation)<br>decreasing or resolved dyspnoea.<br>Individuals with long illness were more likely to be asymptomatic at baseline if Baseline symptoms and seasonali<br>Individuals with long illness were reprinting May-September vs. other the Reciprocally, the odds ratio for exp Individuals with long since the more likely to be asymptomate in a baseline into period steams of during May-September vs. other times of year (280 of 910 [30.8%] vs. 97 of 440 [22.0%], p=0.002).<br>Reciprocally, the odds rat Reciprocally, the odds ratio for experiencing some specific symptoms at baseline was higher in<br>November-March vs. May-September (for low mood, headache, rhinorrhoea, lymphadenopathy,<br>chest pain, sneezing, sore throat, and Reciprocally, the odds ratio for experiencing some specific symptoms of anticiating ratio for November-March vs. May-September (for low mood, headache, rhinorrhoea, lymphadenopath<br>chest pain, sneezing, sore throat, and oph November-March view of the method of the method with sheet pain, sneezing, sore throat, and ophthalmodynia), after correcting for all Model 3 covariates<br>(data not shown).<br>Individuals with short illness also appeared to hav

chest pain, sneezing, see sinest, and ophthalmodynia), after correcting for all Model 3 comments<br>(data not shown).<br>Individuals with short illness also appeared to have more rhinorrhea, sneezing, and headache at<br>baseline be (and not shown).<br>Individuals with sh<br>baseline symptom

## Baseline symptoms in individuals with long illness

Individuals with short and March, though not significantly after FDR adjustment.<br>Baseline symptoms in individuals with long illness<br>In individuals with long illness, 440 (32.6%) reported at least one symptom at baseline wh Baseline symptoms in individuals with long illness<br>In individuals with long illness, 440 (32.6%) reported at least one symptom at baseline w<br>(67.4%) were asymptomatic. In individuals with long index, 440 (22.40%) were asymptomatic.<br>
(67.4%) were asymptomatic. (67.4%) were asymptomatic.

p=0.0004); younger (54 years [46; 62] vs. 59 years [52; 65], p<0.0001); have allergic rhinitis (248 of 440 [56.4%] vs. 404 of 910 [44.3%], p<0.0001) and/or a prior mental health diagnosis (141 of 440 [32.0%] vs. 230 of 91 p=0.0001); younger (1991); younger (1991); younger (1981); younger (1981); younger (1988); younger (1988); you<br>
[32.0%] vs. 230 of 910 [25.7%], p=0.011). Thirty-six of 507 (7.1%) symptomatic and 61 of 1088 (5.69)<br>
asymptom 440 [32.0%] vs. 230 of 910 [25.7%], p=0.011). Thirty-six of 507 (7.1%) symptomatic and 61 of 1088 (5.6<br>asymptomatic individuals were health care workers (p=0.29)).<br>Post-COVID symptom burden was higher in individuals with l

asymptomatic individuals were health care workers (p=0.29)).<br>Post-COVID symptom burden was higher in individuals with long illness who were symptomatic (vs.<br>asymptomatic) at baseline (median [IQR] symptom burden (6 [3; 9] asymptomatic individuals were health care workers (p=0.29)).<br>Post-COVID symptom burden was higher in individuals with lo<br>asymptomatic) at baseline (median [IQR] symptom burden (6<br>symptom burden was associated with post-COV asymptomatic) at baseline (median [IQR] symptom burden (6 [3; 9] vs. 4 [2; 7], p<0.0001). Baseline<br>symptom burden was associated with post-COVID symptom burden (p<0.0001; β=+5.6% CI [4.4; 6.8]<br>per additional baseline symp baseline burden was associated with post-COVID symptom burden (p<0.0001;  $\beta$ =+5.6% CI [4.4; 6.8<br>per additional baseline symptom) after adjustment for all initial matching criteria.<br>Baseline symptomatic status (here, any

symptom burden was associated with post-COVID symptom burden (p<0.0001; β=+5.6% CI [4.4; 6.8]<br>per additional baseline symptom) after adjustment for all initial matching criteria.<br>Baseline symptomatic status (here, any sym per additional baseline symptom (here, any symptom) did not affect the odds of exper<br>COVID symptoms of cutaneous manifestations, palpitations, dyspnoea, cough, tinr<br>fever, rigors, anorexia, or dizziness. All other post-COV From Covid Symptoms were more common in<br>
individuals with long illness and any symptom at baseline, with the exception of anosmia/dysosmia<br>
which was less likely post-COVID (OR=0.75 [0.58; 0.96], p=0.022) (Figure 7)<br>
Influ individuals with long illness and any symptom at baseline, with the exception of anosmia/dy<br>which was less likely post-COVID (OR=0.75 [0.58; 0.96], p=0.022) (Figure 7)<br>Influence of prior comorbidities on baseline and post-

## Influence of prior comorbidities on baseline and post-COVID symptoms

individuals with long interesting and any symptom and anosismicy such the exception of anosing which was less likely post-COVID (OR=0.75 [0.58; 0.96], p=0.022) (Figure 7)<br>Influence of prior comorbidities on baseline and po Influence of prior comorbidities on baseline and post-COVID symptoms<br>Having at least one prior comorbidity (here, including prior mental health d<br>common in individuals with long vs. short illness (926 of 1350 (68.6%) indi Frammon in individuals with long vs. short illness (926 of 1350 (68.6%) individuals with long illne<br>668 of 1350 (49.5%) individuals with short illness; p<0.001). Individuals with long illness duration<br>at least one prior c common in individuals with short illness; p<0.001). Individuals with long illness duration and<br>at least one prior comorbidity (vs. those without) were more likely to have baseline symptoms (here,<br>any symptom) (326 out of 9 at least one prior comorbidity (vs. those without) were more likely to have baseline symptoms (here,<br>any symptom) (326 out of 926 [35.2%] vs. 114 out of 424 [26.9%], p=0.003); and experienced greater<br>post-COVID symptom bur any symptom) (326 out of 926 [35.2%] vs. 114 out of 424 [26.9%], p=0.003); and experienced greater<br>post-COVID symptom burden (5 [2; 9] vs. 3 [2; 6], p<0.0001) (**Figure 8**).<br>Odds ratios for individual symptoms experienced

any symptom) (326 surfaces (30.2%) (31.2%) or 32.4% (30.0001) (Figure 8).<br>post-COVID symptom burden (5 [2; 9] vs. 3 [2; 6], p<0.0001) (Figure 8).<br>Odds ratios for individual symptoms experienced during baseline and post-COV post-COVID symptom burden (5 [2; 9] vs. 5 [2; 0], p<0.0001) (Figure 8).<br>Odds ratios for individual symptoms experienced during baseline and pindividuals with at least one prior comorbidity, per comorbidity, are sho<br>Sensiti

## Sensitivity analysis

Individuals with at least one prior comorbidity, per comorbidity, are shown in **Suppl. Fig S4**.<br>**Sensitivity analysis**<br>Easing logging regularity to fortnightly-minimum reporting and (consequently) defining end o<br>as two wee Sensitivity analysis<br>Sensitivity analysis<br>Easing logging regularity to fortnightly-minimum reporting and (consequently) defining end<br>as two weeks of healthy reports, then reselecting individuals for the matched study, gave Easing regularity to fortning in y minimum reporting and (consequently) at minimum cases and the matched study, gave<br>
remarkably stable results. Samples sizes increased to 2496 individuals per group. Prevalence of<br>
baselin as two weeks of healthy reports, then reselecting individuals for the matched study, gave<br>remarkably stable results. Samples sizes increased to 2496 individuals per group. Prevalence of<br>baseline symptoms in individuals wit remarkably stable results. Samples stable results and the sixtem per group. The again nearly twist<br>that of individuals with short illness (794 [31.8%] vs 441 [17.7%]). The five commonest symptom<br>baseline were unchanged, an baseline were unchanged, and again the same in both groups, in slightly different order and<br>baseline were unchanged, and again the same in both groups, in slightly different order and<br>proportions (data not shown). As previ that committed that the intervention (1994) process (1911). The five the commence symptoms in<br>baseline were unchanged, and again the same in both groups, in slightly different order and<br>proportions (data not shown). As pre baseline were unchanged, and again the same in both groups, in signit, interest order and<br>proportions (data not shown). As previously, all comorbidities (except kidney disease and ca<br>and a prior mental health diagnosis wer proportions (data not shown). As previously, all comorbiditations (encopt kidney discussed and cancer),<br>and a prior mental health diagnosis were more common in individuals with long illness. Suppl. Fig<br>the same of the stat and a prior mental health diagnosis were more common in individuals with long illness. Suppl. Fig S5

presents ray in the constraints rate of symptom consistency between baseline and follow-up in individuals with short and long duration respectively.<br>
Iong duration respectively.<br>
Discussion STR duration respectively.<br>Signal duration respectively.<br>There we have shown a relationship between symptoms before COVID-19 and subsequent illness

## Discussion

long duration respectively.<br>Discussion<br>Here we have shown a relationship between symptoms before COVID-19 and subsequent illness<br>duration. Overall, individuals with long illness were nearly twice as likely to report sympto Here we have shown a relationship between ying term is triveled to the convergent interest.<br>Here weeks before SARS-CoV-2 infection than individuals with short illness (32.5% vs. 18.0%). Howeve<br>two-thirds of individuals wit duration. The sets before SARS-CoV-2 infection than individuals with short illness (32.5% vs. 18.0%). However<br>two-thirds of individuals with long illness were asymptomatic before COVID-19.<br>The odds of long illness increase

Wo-thirds of individuals with long illness were asymptomatic before COVID-19.<br>The odds of long illness increased for any baseline symptom, and for most individual symptoms.<br>Baseline and post-COVID individual symptoms corre two-thrids of individuals with long illness were asymptomatic before COVID-15.<br>The odds of long illness increased for any baseline symptom, and for most indivi<br>Baseline and post-COVID individual symptoms correlated closely The case is a controlled for any parameter property and for incrementation cymperical Baseline and post-COVID individual symptoms correlated closely in individuals with long illness, I<br>evident in individuals with short ill

Baseliant in individuals with short illness, acknowledging the bias created by our selection criteria.<br>Commonest baseline symptoms, regardless of illness duration, were rhinorrhea, sneezing, sore<br>throat, fatigue and headac Commonest baseline symptoms, regardless of illness duration, were rhinorrhea, sneezing, sore<br>throat, fatigue and headache, each with higher prevalence in individuals with subsequent long (vs.<br>short) illness. The cause of t Commontant Artigue and headache, each with higher prevalence in individuals with subsequent long<br>
Short) illness. The cause of these baseline non-specific symptoms is unclear, noting here low UK<br>
Circulation of respiratory throat, fatigue and headaches, parameters in individuals with the cause of these baseline non-specific symptoms is unclear, noting here low UK<br>circulation of respiratory viruses beyond SARS-CoV-2 during the time period of short)<br>Circulation of respiratory viruses beyond SARS-CoV-2 during the time period of this study [19] an<br>that health care workers (with possibly greater workplace exposure to respiratory viruses) did no<br>differ in baseline charactery consisting the same of respiratory circulation of that health care workers (with possibly greater workplace exposure to respiratory viruses) did not<br>differ in baseline symptom status. Baseline symptoms might ref differ in baseline symptom status. Baseline symptoms might reflect non-infectious comorbidities<br>and individuals with several prior comorbidities (most commonly: allergic rhinitis, asthma/lung<br>disease, and a prior mental he and individuals with several prior comorbidities (most commonly: allergic rhinitis, asthma/lung<br>disease, and a prior mental health diagnosis) were more likely to report symptoms during baseline<br>and post-COVID periods (Figu disease, and a prior mental health diagnosis) were more likely to report symptoms during basel<br>and post-COVID periods (Figure 8) though no clear differential symptom pattern within these g<br>was evident (Suppl. Figure S4).<br>D and post-COVID periods (**Figure 8**) though no clear differential symptom pattern within these group<br>was evident (**Suppl. Figure S4**).<br>Despite higher UK pollen counts in May-September, individuals with long illness were les

and post-COVID periods (Figure 8) though no clear differential symptom pattern within these groups<br>was evident (Suppl. Figure 84).<br>Despite higher UK pollen counts in May-September, individuals with long illness were less l was evident (**Suppl. Figure S4**).<br>Despite higher UK pollen counts in May-September, individuals with long illness were less likely to<br>have baseline symptoms during this time. Several symptoms, including low mood, were more Despite higher en pentalisation thanks, experiment, inatitation thanks in the optical and the baseline symptoms during this time. Several symptoms, including low mood, were more common in individuals with long illness and common in individuals with long illness and November-March baselines, noting seasonal affec<br>disorder was not solicited in the mental health questionnaire [15].<br>Our data suggests that some post-COVID symptoms, particularly

disorder was not solicited in the mental health questionnaire [15].<br>Our data suggests that some post-COVID symptoms, particularly in individuals with prior<br>comorbidities, may reflect other, serious, non-COVID illness(es). discree was not solid in the mental health question of  $\frac{1}{2}$ .<br>Our data suggests that some post-COVID symptoms, particularly in<br>comorbidities, may reflect other, serious, non-COVID illness(es). If s<br>OSC/PCS might cause Comorbidities, may reflect other, serious, non-COVID illness(es). If so, symptom misattrib<br>OSC/PCS might cause suboptimal management of these other illness(es), with persistence<br>worsening of the other condition consequentl COSC/PCS might cause suboptimal management of these other illness(es), with persistence and/or<br>worsening of the other condition consequently. Alternatively, individuals with these comorbiditie<br>might be at greater risk of S Worsening of the other condition consequently. Alternatively, individuals with these comorbidities<br>might be at greater risk of SARS-CoV-2 infection, or of more severe COVID-19 [3]); their underlying<br>comorbidities might be might be at greater risk of SARS-CoV-2 infection, or of more severe COVID-19 [3]); their underlying<br>comorbidities might be exacerbated by SARS-CoV-2 infection; and/or they may be more vulnerable<br>comorbidities might be exac comorbidities might be exacerbated by SARS-CoV-2 infection; and/or they may be more vulnerable<br>comorbidities might be exacerbated by SARS-CoV-2 infection; and/or they may be more vulnerable comorbidities might be exactly by SARS-CoV-2 infection; and<br>and  $\alpha$  infection; and  $\alpha$  infection; and  $\alpha$  individuals with asthma might represent usual asthma, post-viral asthma exacerbation, and/or<br>superimposed pathologies specific to SARS-CoV-2 infection such as post-pneumonitis fibrosis or<br>pulmonary microembolism (pertinent superimposed pathologies specific to SARS-CoV-2 infection such as post-pneumonitis fibrosis o<br>pulmonary microembolism (pertinently, our data did not support worsening dyspnoea in indivi<br>with asthma/lung disease; and recent superimposed pathologies specific to state of a simulation of properties proposed in individual<br>with asthma/lung disease; and recent systematic reviews and meta-analyses show asthma was<br>associated with lower risk of SARS-C with asthma/lung disease; and recent systematic reviews and meta-analyses show asthma was<br>associated with lower risk of SARS-CoV-2 infection or of severe COVID-19 [20]. Lastly, altered<br>pandemic health-care access might dis associated with lower risk of SARS-CoV-2 infection or of severe COVID-19 [20]. Lastly, altered<br>pandemic health-care access might disproportionately affect individuals with prior comorbiditie<br>however, there was no evidence pandemic health-care access might disproportionately affect individuals with prior comorbidit<br>however, there was no evidence of such differential access to UK primary care during the pane<br>Our data concord with two large re

however, there was no evidence of such differential access to UK primary care during the pandem<br>Our data concord with two large retrospective studies using primary care data [10,21]. As<br>mentioned, a large UK study of commu our data concord with two large retrospective studies using primary care data [10,21]. As<br>mentioned, a large UK study of community-managed adults with (n=486,149) and without<br>(n=1,944,580) SARS-CoV-2 infection showed moder Our data concern with the large retrospective stands and primary care data [10,101] and without<br>(n=1,944,580) SARS-CoV-2 infection showed moderately higher symptom prevalences in in<br>with (vs. without) SARS-CoV-2 twelve wee (n=1,944,580) SARS-CoV-2 infection showed moderately higher symptom prevalences in in<br>with (vs. without) SARS-CoV-2 twelve weeks after index event, though this difference narro<br>time [10]. Similar to our analysis, longer s with (vs. without) SARS-CoV-2 twelve weeks after index event, though this difference narrowed over<br>time [10]. Similar to our analysis, longer symptom duration associated with female sex, younger age<br>and several prior comor time [10]. Similar to our analysis, longer symptom duration associated with female sex, younger age,<br>and several prior comorbidities including respiratory illnesses and mental health diagnoses; in<br>contrast to our analysis, and several prior comorbidities including respiratory illnesses and mental health diagnoses; in<br>contrast to our analysis, this study did not compare pre- vs. post-infection symptoms per-individual<br>[10]. A German study of 5 contrast to our analysis, this study did not compare pre- vs. post-infection symptoms per-indiv<br>
[10]. A German study of 51,630 general practice patients with COVID-19 reported PCS prevale<br>
8.3% (without comparison populat contrast to our analysis, the study of 51,630 general practice patients with COVID-19 reported PCS prevalence o<br>8.3% (without comparison population), associated with female sex, comorbidities of asthma and<br>several mental h Engineer Comparison population), associated with female sex, comorbidities of asthma and<br>several mental health disorders, and, in contrast to our data, older age; prior symptoms (as opposed<br>to prior co-morbidities) were no several mental health disorders, and, in contrast to our data, older age; prior symptoms (as oppo<br>to prior co-morbidities) were not assessed [21]. Our data also concord with a study of three<br>longitudinal cohorts (54960 par sexercy in the section prior co-morbidities) were not assessed [21]. Our data also concord with a study of three<br>longitudinal cohorts (54960 participants, 3193 testing positively for SARS-CoV-2, of whom 1403<br>developed OSC/ to prior co-morbidity, the correlation (21). Our data also concord with a study of more longitudinal cohorts (54960 participants, 3193 testing positively for SARS-CoV-2, of whom 14<br>developed OSC/PCS), which showed pre-infe developed OSC/PCS), which showed pre-infection psychological distress associated with increas<br>risk of post-COVID symptoms [22]. Pertinently, the UK lockdown abrupt increased mental distre<br>particularly in females, younger a risk of post-COVID symptoms [22]. Pertinently, the UK lockdown abrupt increased mental distress,<br>particularly in females, younger adults, individuals with young children, and individuals with pre-<br>existing mental health co particularly in females, younger adults, individuals with young children, and individuals with pre-<br>existing mental health conditions [23]; and lockdown had a disproportionate effect on symptom<br>experience in individuals wi

prospectively collected symptom data across the pandemic. Symptoms were considered pre- and experience in individuals with pre-existing mental health vulnerabilities [24].<br>Our data also concord with a sub-study of the observational Dutch Lifelines cohort [25], which<br>prospectively collected symptom data across the experience in individuals with a sub-study of the observational Dutch Lifelines corrospectively collected symptom data across the pandemic. Symptoms were post-infection (out to 90-150 days after illness onset, or matched t prospectively collected symptom data across the pandemic. Symptoms were considered pre- a<br>post-infection (out to 90-150 days after illness onset, or matched time point) in 4231 individual<br>COVID-19 (both community- and hosp post-infection (out to 90-150 days after illness onset, or matched time point) in 4231 individuals w<br>COVID-19 (both community- and hospital-managed individuals) and 8462 controls. Symptom<br>severity worsened more in the COVI post-infection (out to 90-150) and hospital-managed individuals) and 8462 controls. Symptom<br>severity worsened more in the COVID-19 group (vs. uninfected controls), both during acutely and<br>during days 90-150, for several sy Severity worsened more in the COVID-19 group (vs. uninfected controls), both during acutely<br>during days 90-150, for several symptoms including breathlessness, chest pain, myalgias, ano<br>and fatigue: overall, 21.4% of cases severity interting more in the COVID-19 group (ver annualized controls) wentering acutely and<br>during days 90-150, for several symptoms including breathlessness, chest pain, myalgias, anosmia<br>and fatigue: overall, 21.4% of and fatigue: overall, 21.4% of cases (vs. 8.7% of controls) had substantial symptom increases 90-150<br>days after infection compared with pre-infection, an increased symptom severity burden of 12.7%<br>above background. Again, days after infection compared with pre-infection, an increased symptom severity burden of 12.7%<br>above background. Again, a gender effect was evident: females reported greater symptom severity<br>14 above background. Again, a gender effect was evident: females reported greater symptom severity above background. Again, a gender effect was evident: females reported greater symptom severity burden of 12.7% above background. Again, a gender effect was evident: females reported greater symptom severity

and individuals with confirmed COVID-19, although our baseline pre-infection data also provide<br>insights into background community symptom prevalences during the pandemic.<br>Lastly, nocebo effects need consideration. Media co

insights into background community symptom prevalences during the pandemic.<br>Lastly, nocebo effects need consideration. Media commentary regarding OSC/PSC has been<br>widespread, often featuring 'floating numerators'. Individu Lastly, nocebo effects need consideration. Media commentary regarding OSC/PS<br>widespread, often featuring 'floating numerators'. Individuals with anxiety and ps<br>are particularly vulnerable to nocebo effects for pain [26]; w undespread, often featuring 'floating numerators'. Individuals with anxiety and psychological<br>are particularly vulnerable to nocebo effects for pain [26]; whether applicable to the pandem<br>experience is unknown. Relevantly, where pread, security transmiting the analysis and the transmitted with annually and psychological antitions<br>are particularly vulnerable to nocebo effects for pain [26]; whether applicable to the pandemic<br>experience is unk experience is unknown. Relevantly, a high nocebo effect was observed post-SARS-CoV-2 vaccin<br>[27].<br>Se**x and age effects**<br>In both short and long illness groups, both sexes were similarly assiduous and persistent in repc

# Sex and age effects

experience is unknown. Relevantly, a high noce both the statistic post-statistic is a calculation<br>[27].<br>In both short and long illness groups, both sexes were similarly assiduous and persistent in reporting<br>(data not shown [27].<br>**Sex a**<br>In bo<br>(data In a since the shown), contributed to by study design. Nonetheless, whether ultimately experiencing<br>Iong or short illness, individuals with baseline symptoms were more likely to be female, and younger,<br>Ihan their asymptoma (d) ong or short illness, individuals with baseline symptoms were more likely to be female, and your<br>than their asymptomatic counterparts. Our results concord with previous studies showing higher<br>symptom reporting by femal In their asymptomatic counterparts. Our results concord with previous studies showing higher<br>symptom reporting by females vs. males for many conditions, including symptom prevalence in daily<br>life [28] and post-acute infect symptom reporting by females vs. males for many conditions, including symptom prevalence in diffe [28] and post-acute infection syndromes [11]. Sex differences in illness presentation are increasingly recognised, with pote symptom reporting by females variables virining conditions, including symptom prevalence in daily<br>life [28] and post-acute infection syndromes [11]. Sex differences in illness presentation are<br>increasingly recognised, with Increasingly recognised, with potential for differential gender-discriminatory healthcare. Our<br>also concord with previous studies showing decreased symptom reporting (particularly stress<br>related symptoms) with age [29].<br>St increased symptom reporting (particularly stress-<br>increased symptoms) with age [29].<br>Strengths and Limitations<br>Our study used prospective, dense, and granular symptom reporting by each person in a large coh

## Strengths and Limitations

over an extended time period, across confirmed SARS-CoV-2 infection, irrespective of ultimate Framer Fymptoms, Jamerage [29].<br>Strengths and Limitations<br>Our study used prospective, dense<br>over an extended time period, acr<br>illness duration, with each person over an extended time period, across confirmed SARS-CoV-2 infection, irrespective of ultimate<br>illness duration, with each person serving as their own control. Our design and matching approach<br>limited reporter bias, with no illness duration, with each person serving as their own control. Our design and matching appro<br>limited reporter bias, with no (or minimal) difference in logging frequency, duration, or<br>assiduousness by baseline status or b ilmited reporter bias, with no (or minimal) difference in logging frequency, duration, or<br>assiduousness by baseline status or by illness duration. Although predominant circulating SARS-Co\<br>2 variants altered during this st limited reporter bias, with no (or minimal) difference in logging frequency, duration, or<br>assiduousness by baseline status or by illness duration. Although predominant circulating SARS-CoV-<br>2 variants altered during this s anyone with self-defined long COVID of more than four weeks  $[8]$  -Lastly, our data are from 2 variants altered during this every, with during the streng interest variant [16], we<br>controlled for this by matching by testing week. We have avoided the phrase 'Long COVID': \<br>not have health records access and individu not have health records access and individuals were not asked about this diagnosis formally. Thus,<br>our data are agnostic to self-identification, in contrast to ONS data that included in their definitions<br>anyone with self-d not have health records actes and individuals in contrast to ONS data that included in their definitions<br>anyone with self-defined long COVID of more than four weeks [8]. Lastly, our data are from<br>individuals with community anyone with self-defined long COVID of more than four weeks [8]. Lastly, our data are from<br>individuals with community-managed COVID-19, whereas most papers interrogating long symptom<br>duration post-COVID-19 are dominated by andividuals with community-managed COVID-19, whereas most papers interrogating long syr<br>duration post-COVID-19 are dominated by data from hospitalised individuals in whom mechi<br>underlying symptom duration are likely to be individuals with community-managed COVID-19, managed COVID-19, whereas most papers interrogating long symptom<br>duration post-COVID-19 are dominated by data from hospitalised individuals in whom mechanisms<br>underlying symptom duration post-COVID-19 are dominated by data from hospitalised individuals in which included<br>underlying symptom duration are likely to be different. underlying symptom duration are likely to be different.

Which may affect symptom experience (for example, varying personal/regional lockdown<br>requirements resulting in varying exposure to other (i.e., non SARS-Cov-2) viruses [19]; varying<br>extent and duration of social isolation requirements resulting in varying exposure to other (i.e., non SARS-Cov-2) viruses [19]; values than divided to social isolation and loneliness; differential home-schooling responsionally experienced by women and associate requirements resulting in varying experience cancel (i.e., normalize to a pyrameter (i.e.), varying<br>extent and duration of social isolation and loneliness; differential home-schooling responsibiliti<br>which burden was dispro extending which burden was disproportionally experienced by women and associated with markedly increas<br>psychological distress [15].<br>The requirement for one week's asymptomatic logging record prior to COVID-19 symptom onset

which burden was disproportionally experienced by a single and associated with markedly increased<br>psychological distress [15].<br>The requirement for one week's asymptomatic logging record prior to COVID-19 symptom onset (to<br> provide the requirement for one w<br>Primable illness duration to b<br>less affected by pre-existing<br>symptoms and comorbidition The requirement for one weekending regaing record prior to COVID-19 and the relationship of prior symptoms and comorbidities with subsequent illness duration. Additionally, all individuals had community-managed COVID-19; o less affected by pre-existing co-morbidities, resulting in underestimation of the relationship of prio<br>symptoms and comorbidities with subsequent illness duration. Additionally, all individuals had<br>community-managed COVIDless affects and comorbidities with subsequent illness duration. Additionally, all individuals had<br>community-managed COVID-19; our data cannot be extrapolated to hospitalised individuals.<br>Although the direct symptom questi

symmunity-managed COVID-19; our data cannot be extrapolated to hospitalised individuals.<br>Although the direct symptom questions expansion from November 2020 was informed by feed<br>from individuals experiencing OSC/PCS, pertin community-managed COVID-19; our data cannot be entropolated to hospitalised individuals<br>Although the direct symptom questions expansion from November 2020 was informed by fee<br>from individuals experiencing OSC/PCS, pertinen From individuals experiencing OSC/PCS, pertinent symptoms may have been missed - although our direct questions covered the symptom groups in other PCS studies [10,21,25]. Further, all but two symptom questions were binary from individuals experiencing OSC/PCS, permission, proposition, yindividuals and there, all but two<br>symptom questions were binary (yes/no) with no available quantitative assessment or health record<br>linkage.<br>Our analysis us direct questions were binary (yes/no) with no available quantitative assessment or health reco<br>
linkage.<br>
Our analysis used data from the first reported SARS-CoV-2 positive result and subsequent illness<br>
duration. Evidence

symptom and integrals in the first of the binary of the<br>Surranalysis used data from the first reported SARS-CoV-2 positive result and subsequent illness<br>duration. Evi ی<br>Our anal<br>duration<br>group); ł<br>Overall, Our also know the first report of the first reported and the first reported in the first report of the group); however, as it transpired, only one illness episode was included per participant.<br>Overall, ZOE/CSS app users ar

group); however, as it transpired, only one illness episode was included per participant.<br>Overall, ZOE/CSS app users are not representative of the UK population (younger, more female,<br>higher educational status, lower ethni group); however, as it transpired, only one interest per a matter per participant.<br>Overall, ZOE/CSS app users are not representative of the UK population (younger, more<br>higher educational status, lower ethnic diversity, ov Migher educational status, lower ethnic diversity, over-representative of healthcare workers).<br>
Moreover, individuals had to log assiduously, at length, which might bias our data towards<br>
individuals with specific app usag higher educational status, lower ethnic diversity, over-representative of healthcare workers).<br>Moreover, individuals had to log assiduously, at length, which might bias our data towards<br>individuals with specific app usage individuals with specific app usage behaviours and potentially exclude individuals with app<br>fatigue. To assess this we compared our study participants to all ZOE app symptomatic test<br>users, across the same time period: our fatigue. To assess this we compared our study participants to all ZOE app symptomatic test-positive<br>users, across the same time period: our cohort were more persistent than positive ZOE app users<br>overall (time between symp factors across the same time period: our cohort were more persistent than positive ZOE app users<br>overall (time between symptom commencement and last report: median 206 days [IQR: 140;292] fo<br>our cohort vs. 99 days [IQR: 15 overall (time between symptom commencement and last report: median 206 days [IQR: 140;292]<br>our cohort vs. 99 days [IQR: 15;159] for test-positive symptomatic ZOE app users overall); and app<br>fatigue was not evident but the our cohort vs. 99 days [IQR: 15;159] for test-positive symptomatic ZOE app users overall); and app<br>fatigue was not evident but the opposite, with strong correlation (Spearman rho 0.465, p<0.00001)<br>in all symptomatic test-p fatigue was not evident but the opposite, with strong correlation (Spearman rho 0.465, p<0.00001<br>in all symptomatic test-positive app users between symptom duration and continuing app usage<br>(both calculated from date of sy Fatigue in all symptomatic test-positive app users between symptom duration and continuing app usage<br>(both calculated from date of symptom onset). Although we cannot comment more granularly (e.g., possible differential dr In an experiment test-positive app users a sinter-positive and an example and comment more granularly (e<br>possible differential drop-out of individuals with particular symptoms) these data do not support<br>fatigue; and pertin possible differential drop-out of individuals with particular symptoms) these data do not support app<br>fatigue; and pertinently, two-thirds of our long illness group had symptoms for ≥12 weeks. fatigue; and pertinently, two-thirds of our long illness group had symptoms for ≥12 weeks.<br>16 fatigue; and pertinently, two-thirds of our long illness group had symptoms for  $\alpha$ 12 weeks.

Our inclusion criteria required participation in the mental health questionnaire in early 2021 [15]. Of 1,257,278 million app users at the time, 715,324 (56.8%) participated; whether participation was affected by presence/ 1,257,278 affected by presence/absence of a prior mental health diagnoses cannot be determined. Howeve<br>1,257,232,257,278 censoring for mental health survey participation only significantly affected the male/female ratio<br>1, affected by presence, absorber a protonomial and angular summanded be male/female ratio<br>(females more likely to participate; data not shown).<br>We considered the impact of stringent logging frequency possibly precluding indi

censoring for more likely to participate; data not shown).<br>Character mental health survey participate is the mand/or more severe illness. However, our sensitivity analysis with loosened stringency did not<br>change our result (females of stringent logging freque<br>
(female) we considered the impact of stringent logging freque<br>
and/or more severe illness. However, our sensitivity<br>
change our results significantly (**Supp. Fig. S5-S7**).<br>
Lastly, we Merceman considered the impact of stringent logging frequency precluintly precluing in an analoger<br>and/or more severe illness. However, our sensitivity analysis with loosened stringency did not<br>change our results significa

change our results significantly (Supp. Fig. S5-S7).<br>Lastly, we cannot exclude an effect of app participation *per se* on symptom reporting (regardle:<br>disease duration), noting that use of symptom tracking apps can inflate change our results significantly (Supp. Fig. 35-S7).<br>Lastly, we cannot exclude an effect of app particip<br>disease duration), noting that use of symptom trad<br>individuals [30].<br>Conclusion Lastly, we cannot exclude an effect of app participation *per se* on symptom reporting (regardless of<br>disease duration), noting that use of symptom tracking apps can inflate symptom reporting in some<br>individuals [30].<br>Symp

## Conclusion

individuals [30].<br>Conclusion<br>Symptoms prevalence before SARS-CoV-2 infection differed in individuals who subsequently<br>manifest long vs. short illness, with correlation of symptom burden and specific symptom experience **Conclusion<br>Symptoms preva<br>manifest long vs<br>between baselin** Symptoms previously and start of the american interest in individuals and specific symptom explotive beforeen baseline and post-COVID periods. At least some of this risk is influenced by prior comorbidities. However, the m between baseline and post-COVID periods. At least some of this risk is influenced by prior<br>comorbidities. However, the majority of individuals with post-COVID symptoms were asymptomatic<br>before SARS-CoV-2 infection. Our dat between baseline and post-COVID symptoms were asym<br>before SARS-CoV-2 infection. Our data highlight the importance of parsing symptoms relate<br>COVID-19 vs. other diseases – but also the difficulties of this. Genetic studies before SARS-CoV-2 infection. Our data highlight the importance of parsing symptoms related to<br>COVID-19 vs. other diseases – but also the difficulties of this. Genetic studies in OSC/PCS, including<br>Mendelian randomization a before SARS-COVID-19 vs. other diseases – but also the difficulties of this. Genetic studies in OSC/PCS, includ<br>Mendelian randomization approaches, may prove unbiased means of disentangling association f<br>causation. Practic COVID-19 Mendelian randomization approaches, may prove unbiased means of disentangling association fron<br>Causation. Practically speaking, the clinical implications are unchanged: an open and holistic<br>approach is needed to m Mendelian randomization approaches, may prove unbiased means of disentangling assessmental randomization.<br>
Practically speaking, the clinical implications are unchanged: an open and holistic<br>
approach is needed to manage p approach is needed to manage post-COVID symptoms, whatever their aetiology.<br>
An open and holistical inputs are unchanged: and holistical implications are unchanged: and holistical implications are unchanged:  $\frac{1}{2}$  in t approach is needed to manage post-COVID symptoms, whatever the intervals  $\mathcal{O}_I$ .

## Acknowledgements

## Data sharing statement

We actizen science enabling this research.<br> **Data sharing statement**<br>
Data collected in the COVID Symptom Study smartphone app can be shared with other health<br>
researchers through the UK National Health Service-funded Heal The citizen science enabling the cleared.<br>
Data collected in the COVID Symptom Stresearchers through the UK National Head<br>
Anonymised Information Linkage consort<br>
UK). Anonymised data are available to be<br>
mublic interact h The Covid-Bird in the COVID Symptom Study Service - The COVID States and Secure Research UK and Secure Anonymised Information Linkage consortium, housed in the UK Secure Research Platform (Swatch UK). Anonymised data are a Anonymised Information Linkage consortium, housed in the UK Secure Research Platform (Swans<br>UK). Anonymised data are available to be shared with researchers according to their protocols in<br>public interest <u>https://web.www.</u> Anonymised Information Linking Constraintly, it also the UK Secure Research Platform (Straintly, 2000)<br>UK). Anonymised Infar are available to be shared with researchers according to their protocols in the<br>public interest <u></u> public interest <u>https://web.www.healthdatagateway.org/dataset/594cfe55-96e3-45ff-874c-</u><br>2c0006eeb881.<br>Conflict of Interest<br>Tim Spector and Jonathan Wolf are co-founders and founder shareholders of ZOE Ltd. Christina Hu,

## Conflict of Interest

and Joan Capdevila Pujol are employees of ZOE Ltd. Claire J Steves and Sebastien Ourselin ha **COMPTERTIFY**<br> **COMPTERT COMPTERT**<br>
Tim Spector and<br>
and Joan Capder<br>
consulted for ZC  $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$ and Joan Capdevila Pujol are employees of ZOE Ltd. Claire J Steves and Sebastien Ourselin have<br>consulted for ZOE Ltd. All other authors (CHS, EM, LSC, NJC, AH, ELD, HVB, KR) declare no conflict of<br>interest.<br>This work is su

## Funding declaration.

consulted for ZOE Ltd. All other authors (CHS, EM, LSC, NJC, AH, ELD, HVB, KR) declare no conflicinterest.<br>
Funding declaration.<br>
This work is supported by the Wellcome Engineering and Physical Sciences Research Council C<br> consulted for ZOE Ltd. All other and Consumer (CHS) 2007, ALL of ALL of ALL ONES An<br>
Funding declaration.<br>
This work is supported by the Wellcome Engineering and Physical Sciences Research Council Centr<br>
for Medical Engine merest.<br>Funding<br>This worl<br>for Med<br>Innovatic<br>the UK D for Medical Engineering at King's College London (WT 203148/Z/16/Z), the UK Research and<br>Innovation London Medical Imaging & Artificial Intelligence Centre for Value-Based Healthcare, and<br>the UK Department of Health and So For Michael Engineering at Times Corresponsive Limitary (12021-12), 2012), the Collection Condon Medical Imaging & Artificial Intelligence Centre for Value-Based Healthcare, and<br>the UK Department of Health and Social Care, Innovation London Medical Institute for Health Research<br>
Innovational Institute for Health Research<br>
(NIHR)- funded BioResource, Clinical Research Facility and comprehensive Biomedical Research<br>
Centre (BRC) award to Guy's (NIHR)- funded BioResource, Clinical Research Facility and comprehensive Biomedical Research<br>Centre (BRC) award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College<br>London and King's College Hospit (NIHR)- Funded Biomarchical Research Council Centre (BRC) award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College<br>London and King's College Hospital NHS Foundation Trust. This research was also London and King's College Hospital NHS Foundation Trust. This research was also funded in part by<br>the Wellcome Trust grant [215010/Z/18/Z]. Investigators also received support from the Chronic<br>Disease Research Foundation ( the Wellcome Trust grant [215010/Z/18/Z]. Investigators also received support from the Chronic<br>Disease Research Foundation (CDRF), HMT/UKRI/MRC COVID-19 Longitudinal Health and Wellbeing -<br>National Core Study (LHW-NCS) (MC The Wellcome Trust grant (CDRF), HMT/UKRI/MRC COVID-19 Longitudinal Health and Wellbeing -<br>National Core Study (LHW-NCS) (MC\_PC\_20030, MC\_PC\_20059 and NIHR COV-LT-0009), Medical<br>Research Council (MRC), British Heart Founda National Core Study (LHW-NCS) (MC\_PC\_20030, MC\_PC\_20059 and NIHR COV-LT-0009), Medical<br>Research Council (MRC), British Heart Foundation (BHF), Alzheimer's Society, European Union. ZOE<br>Limited provided in-kind support for a National Council (MRC), British Heart Foundation (BHF), Alzheimer's Society, European Union. ZOE<br>Limited provided in-kind support for all aspects of building, running, and supporting the app and<br>service to users worldwide. Research Council (MRC), British Heart Council (MRC), Align, Timing, and supporting the app and<br>service to users worldwide. SO was supported by the French government, through the 3IA Côte<br>d'Azur Investments in the Future pr Exercice to users worldwide. SO was supported by the French government, through the 3IA Côte<br>d'Azur Investments in the Future project managed by the National Research Agency (ANR) with the<br>reference number ANR-19-P3IA-0002 d'Azur Investments in the Future project managed by the National Research Agency (ANR) with the<br>reference number ANR-19-P3IA-0002. EM received funds from the Medical Research Council UK<br>(Skills Development Scheme) and from reference number ANR-19-P3IA-0002. EM received funds from the Medical Research Council UK<br>(Skills Development Scheme) and from the National Institute for Health Research UK (grant<br>n.134293).<br>For the purpose of Open Access, reference number Andrew Presh Presh Presh Presh Path and the Medical Research UK (grant<br>n.134293).<br>For the purpose of Open Access, the author has applied a CC BY public copyright licence to any<br>Author Accepted Manuscript (

(Skills Development Scheme) and from the National Institute for Nethin Nethinal Institute for<br>
n.134293).<br>
For the purpose of Open Access, the author has applied a CC BY public copyright licence to any<br>
Author Accepted Man

n.134293).<br>For the purpose of Open Access, the author has applied a CC BY public copyright licence to any<br>Author Accepted Manuscript (AAM) version arising from this submission.<br>**Contribution statement**<br>ELD and CHS conceptu For the purpose of purpose of purpose of operator of purpose of purpose of purpose of any purpose of  $A$ .<br> **Contribution statement**<br>
ELD and CHS conceptualised the study design. MA, BM, JCP, CHS contributed to data curatio Contribution statement<br>ELD and CHS conceptualised the study design. MA, BM, JCP, CHS contribut<br>associated software tools. CJS, TS contributed to acquisition design. CHS p<br>verified the underlying data. CHS interpreted the r ( Fi t ( i l ELD and CHS conceptualised the study design. MA, BM, JCP, CHS contributed to data curation and ELD and CHS contributed to acquisition design. CHS performed the analysis are verified the underlying data. CHS interpreted the results with ELD who supervised the study. ELD a CHS drafted the manuscript. TDS, SO and CJS a verified the underlying data. CHS interpreted the results with ELD who supervised the study. ELD and<br>CHS drafted the manuscript. TDS, SO and CJS acquired funding. All authors contributed to<br>interpretation of data and criti verified the underlying and the undependent of the underlying and change. All authors contributed to interpretation of data and critical revision of the manuscript.<br>Interpretation of data and critical revision of the manus CHS drafted the manuscript. The manuscript. The manuscript. The manuscript. interpretation of data and critical revision of the manuscript.

## Transparency statement

The lead and a manuscript is an interest of the study have been omitted; and that any<br>discrepancies from the study as planned have been explained.<br>Webpages studiscrepancies from the study as planned have been explained.<br>Studies of the study aspects of the study have been omitted; and the study of th discrepancies from the study as planned have been explained.

# Webpages

## References

- |
|
|
- with COVID-19-related death using OpenSAFELY. Nature 2020;584:430–6.<br>https://doi.org/10.1038/s41586-020-2521-4.<br>[2] Gardashkhani S, Ajri-Khameslou M, Heidarzadeh M, Rajaei Sedigh SM. Post-Intensive Care<br>Syndrome in Covid-1 with COVID-19-related death using OpenSAFELY. Nature 2020;584:430–6.<br>https://doi.org/10.1038/s41586-020-2521-4.<br>Gardashkhani S, Ajri-Khameslou M, Heidarzadeh M, Rajaei Sedigh SM. Post-Intensive Care<br>Syndrome in Covid-19 Pa https://doi.org/10.1038/s41586-020-2521-4.<br>Gardashkhani S, Ajri-Khameslou M, Heidarzad<br>Syndrome in Covid-19 Patients Discharged Fre<br>2021;23:530–8. https://doi.org/10.1097/NJH.<br>Sudre CH, Murray B, Varsavsky T, Graham MS<br>pre
- Syndrome in Covid-19 Patients Discharged From the Intensive Care Unit. J Hosp Palliat Nur<br>2021;23:530–8. https://doi.org/10.1097/NJH.0000000000000789.<br>[3] Sudre CH, Murray B, Varsavsky T, Graham MS, Penfold RS, Bowyer RC, Syndrome in Correction Correction Covid-19 Patients Discharged From the International Covid-2021;23:530–8. https://doi.org/10.1097/NJH.00000000000000789.<br>Sudre CH, Murray B, Varsavsky T, Graham MS, Penfold RS, Bowyer RC, e Sudre CH, Murray B, Varsavsky T, Graham MS, Penfold RS, Bowyer<br>predictors of long COVID. Nat Med 2021;27:626–31. https://doi.org<br>01292-y.<br>Evans RA, McAuley H, Harrison EM, Shikotra A, Singapuri A, Sereno<br>cognitive, and men
- predictors of long COVID. Nat Med 2021;27:626–31. https://doi.org/10.1038/s41591-021-<br>01292-y.<br>[4] Evans RA, McAuley H, Harrison EM, Shikotra A, Singapuri A, Sereno M, et al. Physical,<br>cognitive, and mental health impacts 01292-y.<br>
Evans RA, McAuley H, Harrison EM, Shikotra A, Singapuri A, Sereno M, et al. Physical,<br>
cognitive, and mental health impacts of COVID-19 after hospitalisation (PHOSP-COVID): a UK<br>
multicentre, prospective cohort s Evans RA, McAuley H, Harrison EM, Shikotra A, Singapuri A, Sereno M, et al. Physical, cognitive, and mental health impacts of COVID-19 after hospitalisation (PHOSP-COVID<br>multicentre, prospective cohort study. Lancet Respir Med 2021;9:1275–87.<br>https://doi.org/10.1016/S2213-2600(21)00383-0.<br>[5] A clinical cas
- multicentre, prospective cohort study. Lancet Respir Med 2021;9:1275–87.<br>https://doi.org/10.1016/S2213-2600(21)00383-0.<br>A clinical case definition of post COVID-19 condition by a Delphi consensus, 6 October 2021<br>n.d. https https://doi.org/10.1016/S2213-2600(21)00383-0.<br>A clinical case definition of post COVID-19 condition by a Delphi consensus,<br>n.d. https://www.who.int/publications/i/item/WHO-2019-nCoV-Post\_COVI<br>Clinical\_case\_definition-2021 A clinical case definition of post COVID-19 condition.d. https://www.who.int/publications/i/item/WH<br>Clinical\_case\_definition-2021.1 (accessed July 10,<br>Boscolo-Rizzo P, Polesel J, Spinato G, Fabbris C, Ca<br>Predominance of an
- m.d. https://www.who.int/publications/i/item/WHO-2019-nCoV-Post\_COVID-19\_condition-<br>Clinical\_case\_definition-2021.1 (accessed July 10, 2022).<br>[6] Boscolo-Rizzo P, Polesel J, Spinato G, Fabbris C, Calvanese L, Menegaldo A, Boscolo-Rizzo P, Polesel J, Spinato G, Fabbris C, Calvanese<br>Predominance of an altered sense of smell or taste amon<br>with mildly symptomatic COVID-19\* n.d. https://doi.org/<br>Chen C, Haupert SR, Zimmermann L, Shi X, Fritsche
- Post-Coronavirus Disease 2019 (COVID-19) Condition or Long COVID: A Meta-Analysis and Fredominance of an altered sense of smell or taste among long-lasting symptom<br>with mildly symptomatic COVID-19\* n.d. https://doi.org/10.4193/Rhin20.263.<br>[7] Chen C, Haupert SR, Zimmermann L, Shi X, Fritsche LG, Mukherjee B Predominal Symptomatic COVID-19\* n.d. https://doi.org/10.4193/Rhin20.263.<br>Chen C, Haupert SR, Zimmermann L, Shi X, Fritsche LG, Mukherjee B. Global Prevalence of<br>Post-Coronavirus Disease 2019 (COVID-19) Condition or Long C Chen C, Haupert SR, Zimmermann L, Shi X, Fritsche LG, Mukherjee B. Global Prost-Coronavirus Disease 2019 (COVID-19) Condition or Long COVID: A Meta-A<br>Systematic Review. J Infect Dis 2022. https://doi.org/10.1093/infdis/jia
- For Coronavirus Disease 2019 (COVID-19) Condition or Long COVID: A Meta-Analysis and<br>Systematic Review. J Infect Dis 2022. https://doi.org/10.1093/infdis/jiac136.<br>[8] Prevalence of ongoing symptoms following coronavirus (C Systematic Review. J Infect Dis 2022. https://doi.org/10.1093/infdis/jiac136.<br>Prevalence of ongoing symptoms following coronavirus (COVID-19) infection in the UK -<br>Office for National Statistics n.d.<br>https://www.ons.gov.uk Prevalence of ongoing symptoms following coronavirus (COVID-19) infection in the UK -<br>Office for National Statistics n.d.<br>https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsan<br>ddiseases/bulle Coffice for National Statistics n.d.<br>
https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditional<br>
diseases/bulletins/prevalence of ongoing symptoms following coronavirus covid 19 infection euk/30 m https://www.ons.gov.uk/peopler<br>https://www.ons.gov.uk/peopler<br>ddiseases/bulletins/prevalenceof<br>euk/30march2023 (accessed Mar<br>Thompson EJ, Williams DM, Wall<br>burden and risk factors in 10 UK<br>Commun 2022;13:3528. https://
- ddiseases/bulletins/prevalenceofongoingsymptomsfollowing coronavirus covid19infectioninth<br>euk/30march2023 (accessed May 10, 2023).<br>Thompson EJ, Williams DM, Walker AJ, Mitchell RE, Niedzwiedz CL, Yang TC, et al. Long COVID euk/30march2023 (accessed May 10, 2023).<br>Thompson EJ, Williams DM, Walker AJ, Mitchell RE, Niedzwiedz CL, Yang TC, et al. Long COVID<br>burden and risk factors in 10 UK longitudinal studies and electronic health records. Nat<br> euk/30march2023 (accession CDM, Walker AJ, Mitch<br>2023 (accession 2022;13:3528. https://doi.org/10.10<br>2023:13:3528. https://doi.org/10.10<br>2023:13:3528. https://doi.org/10.10 burden and risk factors in 10 UK longitudinal studies and electronic health records. Nat<br>Commun 2022;13:3528. https://doi.org/10.1038/s41467-022-30836-0.<br>19 0 Commun 2022;13:3528. https://doi.org/10.1038/s41467-022-30836-0.<br>
Electronic health records. National studies and electronic health records. National studies and electronic he<br>
Studies and the correction of the correctio Commun 2022;13:3528. https://doi.org/10.1038/s41467-022-30836-0.

- Symptoms and risk factors for long COVID in non-hospitalized adults. Nat Med 2022;28<br>14. https://doi.org/10.1038/s41591-022-01909-w.<br>[11] Choutka J, Jansari V, Hornig M, Iwasaki A. Unexplained post-acute infection syndrome
- 14. https://doi.org/10.1038/s41591-022-01909-w.<br>Choutka J, Jansari V, Hornig M, Iwasaki A. Unexplained post-acute infection syndromes. Nat<br>Med 2022;28:911–23. https://doi.org/10.1038/s41591-022-01810-6.<br>Jason LA, Katz B, G
- 14. https://doi.org/10.1038/s41<br>14. https://doi.org/10.1038/s41<br>14. https://doi.org/10.1038/s41<br>14. https://doi.org/10.1038/s41<br>14. https://doi.org/10.1038/s41<br>14. https://doi.org/10.1038/s41<br>14. https://doi.org/10.1038/s4
- Med 2022;28:911–23. https://doi.org/10.1038/s41591-022-01810-6.<br>
[12] Jason LA, Katz B, Gleason K, McManimen S, Sunnquist M, Thorpe T. A Prospective Study of<br>
Infectious Mononucleosis in College Students. Int J Psychiatry Med 2022;28:911–23. https://doi.org/10.1038/s41591-022-01810-6.<br>Jason LA, Katz B, Gleason K, McManimen S, Sunnquist M, Thorpe T. A Prospective Study of<br>Infectious Mononucleosis in College Students. Int J Psychiatry (Overl Infectious Mononucleosis in College Students. Int J Psychiatry (Overl Park) 2017;2:1–14.<br>
[13] Klem F, Wadhwa A, Prokop LJ, Sundt WJ, Farrugia G, Camilleri M, et al. Prevalence, Risk<br>
Factors, and Outcomes of Irritable Bow Infermal Monometer Mononucleon Mononucleon Mononucleon Mononucleon Risk<br>Infectors, and Outcomes of Irritable Bowel Syndrome After Infectious Enteritis: A Systema<br>Review and Meta-analysis. Gastroenterology 2017;152:1042-105 Factors, and Outcomes of Irritable Bowel Syndrome After Infectious Enteritis: A System:<br>
Review and Meta-analysis. Gastroenterology 2017;152:1042-1054.e1.<br>
https://doi.org/10.1053/j.gastro.2016.12.039.<br>
[14] Yong SJ. Long Review and Meta-analysis. Gastroenterology 2017;152:1042-1054.e1.<br>https://doi.org/10.1053/j.gastro.2016.12.039.<br>Yong SJ. Long COVID or post-COVID-19 syndrome: putative pathophysiology, risk factors, an<br>treatments. Infect D
- 
- Review Meta-analysis. Julie 2016.12.039.<br>Review SJ. Long COVID or post-COVID-19 syndrome: putative pathophysitreatments. Infect Dis 2021;53:737–54. https://doi.org/10.1080/23744<br>Klaser K, Thompson EJ, Nguyen LH, Sudre CH, Mapper, Marchel, 2012-19, 30.000 MM SV and SV appression symptoms after COVID-19 infection<br>In Neurol Neuros after COVID-19 infection<br>In Neurol Neurosurg treatments. Infect Dis 2021;53:737–54. https://doi.org/10.1080/23744235.2021.1924397.<br>
[15] Klaser K, Thompson EJ, Nguyen LH, Sudre CH, Antonelli M, Murray B, et al. Anxiety and<br>
depression symptoms after COVID-19 infectio The Masser K, Thompson EJ, Nguyen LH, Sudre CH, Antonelli M, Murray B, et al. Anxiety and<br>depression symptoms after COVID-19 infection: results from the COVID Symptom Study ap<br>J Neurol Neurosurg Psychiatry 2021;92:1254–8.
- depression symptoms after COVID-19 infection: results from the COVID Symptom Study<br>
J Neurol Neurosurg Psychiatry 2021;92:1254–8. https://doi.org/10.1136/jnnp-2021-327<br>
Kläser K, Molteni E, Graham M, Canas LS, Österdahl MF depression symptoms after COVID-19 in Neurol Neurosurg Psychiatry 2021;92:1254-8. https://doi.org/10.1136/jnnp-2021-327565.<br>Kläser K, Molteni E, Graham M, Canas LS, Österdahl MF, Antonelli M, et al. COVID-19 due to<br>the B.1 Maser K, Molteni E, Graham M, Canas LS, Österdahl MF, Antonelli M, et al. COVID-19 due to<br>the B.1.617.2 (Delta) variant compared to B.1.1.7 (Alpha) variant of SARS-CoV-2: a<br>prospective observational cohort study. Sci Rep 2 the B.1.617.2 (Delta) variant compared to B.1.1.7 (Alpha) variant of SARS-CoV-2: a<br>prospective observational cohort study. Sci Rep 2022;12:10904.<br>https://doi.org/10.1038/s41598-022-14016-0.<br>[17] Canas LS, Österdahl MF, Den
- prospective observational cohort study. Sci Rep 2022;12:10904.<br>https://doi.org/10.1038/s41598-022-14016-0.<br>Canas LS, Österdahl MF, Deng J, Hu C, Selvachandran S, Polidori L, et al. Disentangl<br>vaccination symptoms from earl prospective org/10.1038/s41598-022-14016-0.<br>
https://doi.org/10.1038/s41598-022-14016-0.<br>
Canas LS, Österdahl MF, Deng J, Hu C, Selvachandran S, Polidori I<br>
vaccination symptoms from early COVID-19. EClinicalMedicine 2<br>
ht https://doi.org/10.1038/s41598-022-14016-0.<br>Canas LS, Österdahl MF, Deng J, Hu C, Selvachandran S, Polidori L, et al. Disentangling post-<br>vaccination symptoms from early COVID-19. EClinicalMedicine 2021;42.<br>https://doi.org
- 
- [19] National flu and COVID-19 surveillance reports GOV.UK n.d.<br>https://www.gov.uk/government/statistics/national-flu-and-covid-19-surveillance-reports vaccination symptoms for the 2021.101212.<br>https://doi.org/10.1016/j.eclinm.2021.101212.<br>Kuhn HW. The Hungarian method for the assignment problem. Naval Re.<br>Quarterly 1955;2:83–97. https://doi.org/10.1002/nav.3800020109.<br>Na Quarterly 1955;2:83–97. https://doi.org/10.1002/nav.3800020109.<br>National flu and COVID-19 surveillance reports - GOV.UK n.d.<br>https://www.gov.uk/government/statistics/national-flu-and-covid-1<br>(accessed August 18, 2022). (18) Quarterly 1955;2:83–97. https://doi.org/10.1002/nav.3800020109.<br>
[19] National flu and COVID-19 surveillance reports - GOV.UK n.d.<br>
https://www.gov.uk/government/statistics/national-flu-and-covid-19-surveillance-repor National flu and COVID-19 surveillance reports - GOV.UK n.d.<br>https://www.gov.uk/government/statistics/national-flu-and-covid-1<br>(accessed August 18, 2022).<br>Sunjaya AP, Allida SM, Di Tanna GL, Jenkins C. Asthma and risk of i
- https://www.gov.uk/government/statistics/national-flu-and-o<br>
(accessed August 18, 2022).<br>
[20] Sunjaya AP, Allida SM, Di Tanna GL, Jenkins C. Asthma and risl<br>
ICU admission and mortality from COVID-19: Systematic revie<br>
As (accessed August 18, 2022).<br>Sunjaya AP, Allida SM, Di Tanna GL, Jenkins C. Asthma and risk of infection, hospitalization,<br>ICU admission and mortality from COVID-19: Systematic review and meta-analysis. Journal<br>Asthma 2022; (accessed AP, Allida SM, Di Tar<br>ICU admission and mortality<br>Asthma 2022;59:866–79. htt<br>Kostev K, Smith L, Koyanagi<br>Coronavirus Disease 2019 (C
- [21] CU admission and mortality from COVID-19: Systematic review and meta-analysis. Journal<br>Asthma 2022;59:866–79. https://doi.org/10.1080/02770903.2021.1888116.<br>[21] Kostev K, Smith L, Koyanagi A, Jacob L. Prevalence of a ICU Admin 2022;59:866–79. https://doi.org/10.1080/02770903.2021.1888116.<br>ICONID-19: Systematic Review and Meta-Coronavirus Disease 2019 (COVID-19) Condition in the 12 Months After the Diagnosis of<br>COVID-19 in Adults Follow Asthma 2022;59:866–79. https://doi.org/10.1080/02770903.2021.1888116.<br>Kostev K, Smith L, Koyanagi A, Jacob L. Prevalence of and Factors Associated With Post-<br>Coronavirus Disease 2019 (COVID-19) Condition in the 12 Months A
- Post-COVID-19 Conditions. JAMA Psychiatry 2022. COVID-19 in Adults Followed in General Practices in Germany. Open Forum Infect Dis 20:<br>https://doi.org/10.1093/ofid/ofac333.<br>Wang S, Quan L, Chavarro JE, Slopen N, Kubzansky LD, Koenen KC, et al. Associations of<br>Depression https://doi.org/10.1093/ofid/ofac333.<br>Wang S, Quan L, Chavarro JE, Slopen N, Kubzansky LD, Koenen KC, et al. Associations of<br>Depression, Anxiety, Worry, Perceived Stress, and Loneliness Prior to Infection With Risk of<br>Post Wang S, Quan L, Chavarro JE, Slopen N,<br>Depression, Anxiety, Worry, Perceived<br>Post-COVID-19 Conditions. JAMA Psych<br>https://doi.org/10.1001/jamapsychiatr Depression, Anxiety, Worry, Perceived Stress, and Loneliness Prior to Infection With Risl<br>Post-COVID-19 Conditions. JAMA Psychiatry 2022.<br>https://doi.org/10.1001/jamapsychiatry.2022.2640. Depression, Analytical Prior, Anxiet British, and Loneliness, and Loneliness Post-COVID-19 Conditions. JAMA Psychiatry 2022.<br>https://doi.org/10.1001/jamapsychiatry.2022.2640.<br>https://doi.org/10.1001/jamapsychiatry.2022.264 https://doi.org/10.1001/jamapsychiatry.2022.2640<br>https://doi.org/10.1001/jamapsychiatry.2022.2640<br>https:// https://doi.org/10.1001/jamapsychiatry.2022.2640.

- 
- the COVID-19 pandemic: a longitudinal probability sample survey of the UK population.<br>
Lancet Psychiatry 2020;7:883–92. https://doi.org/10.1016/S2215-0366(20)30308-4.<br>
[24] Ben-Ezra M, Hamama-Raz Y, Goodwin R, Leshem E, Le Lancet Psychiatry 2020;7:883–92. https://doi.org/10.1016/S2215-0366(20)30308-4.<br>Ben-Ezra M, Hamama-Raz Y, Goodwin R, Leshem E, Levin Y. Association between menta<br>health trajectories and somatic symptoms following a second Ben-Ezra M, Hamama-Raz Y, Goodwin R, Leshem E, Levin Y. Association between me<br>health trajectories and somatic symptoms following a second lockdown in Israel: A<br>longitudinal study. BMJ Open 2021;11:1–9. https://doi.org/10.
- health trajectories and somatic symptoms following a second lockdown in Israel: A<br>longitudinal study. BMJ Open 2021;11:1-9. https://doi.org/10.1136/bmjopen-2021-0504<br>[25] Ballering A V., van Zon SKR, olde Hartman TC, Rosma health trajectories and somations. Here, the sometime is and sometime sometimes a selecting A V., van Zon SKR, olde Hartman TC, Rosmalen JGM. Persistence of somations for a second symptoms after COVID-19 in the Netherlands Ballering A V., van Zon SKR, olde Hartman TC, Rosmalen JGM. Persistence of somatic<br>symptoms after COVID-19 in the Netherlands: an observational cohort study. The Lancet<br>2022;400:452–61. https://doi.org/10.1016/S0140-6736(2 Symptoms after COVID-19 in the Netherlands: an observational cohort study. The Lartman 2022;400:452–61. https://doi.org/10.1016/S0140-6736(22)01214-4.<br>[26] Manai M, Van Middendorp H, Veldhuijzen D, Huizinga T, Evers A. Eff
- 
- symptoms after Covid-19 in the Netherlands and COV22;400:452-61. https://doi.org/10.1016/S0140-6736(22)01214-4.<br>Manai M, Van Middendorp H, Veldhuijzen D, Huizinga T, Evers A. Effects in Pain ?: a Narra<br>Review on Mechanisms Manai M, Van Middendorp H, Veldhuijzen D, Huizinga T, Evers A. Effe<br>Review on Mechanisms , Predictors , and Interventions. Pain Rep 201<br>Haas JW, Bender FL, Ballou S, Kelley JM, Wilhelm M, Miller FG, et al. I<br>Events in the Example 19 Manual M, Van Middelp M, Van Middelp H, Veldhuisen D, Valdendorp 1, 2019;4:1–10.<br>
[27] Maas JW, Bender FL, Ballou S, Kelley JM, Wilhelm M, Miller FG, et al. Frequency of Adverse<br>
Events in the Placebo Arms of CO Review on Mechanisms , Predictors , and Interventions. Pain Rep 2019;4:1–10.<br>
Haas JW, Bender FL, Ballou S, Kelley JM, Wilhelm M, Miller FG, et al. Frequency of Adverse<br>
Events in the Placebo Arms of COVID-19 Vaccine Trial
- Events in the Placebo Arms of COVID-19 Vaccine Trials. JAMA Netw Open 2022;5:e2143955<br>https://doi.org/10.1001/jamanetworkopen.2021.43955.<br>[28] Bardel A, Wallander MA, Wallman T, Rosengren A, Johansson S, Eriksson H, et al. Eventy/doi.org/10.1001/jamanetworkopen.2021.43955.<br>Bardel A, Wallander MA, Wallman T, Rosengren A, Johansson S, Eriksson H, et al. Age and se:<br>related self-reported symptoms in a general population across 30 years: Pattern
- and secular trend. PLoS One 2019;14. https://doi.org/10.1371/journal.pone.0211532.<br>Bardel A, Wallander M-A, Wallman T, Rosengren A, Johansson S, Eriksson H, et al. Age an<br>sex related self-reported symptoms in a general pop related self-reported symptoms in a general population across 30 years: Patterns of reporting<br>and secular trend. PLoS One 2019;14. https://doi.org/10.1371/journal.pone.0211532.<br>[29] Bardel A, Wallander M-A, Wallman T, Rose related self-reported. PLoS One 2019;14. https://doi.org/10.1371/journal.pone.0211532.<br>Bardel A, Wallander M-A, Wallman T, Rosengren A, Johansson S, Eriksson H, et al. Age and<br>sex related self-reported symptoms in a genera and secular trend. Products on the property and secular trend. Parally Sardel A, Wallander M-A, Wallman T, Rosengren A, Johansson S, Eriksson H, et al. Age<br>sex related self-reported symptoms in a general population across sex related self-reported symptoms in a general population across 30 years: Patterns of<br>reporting and secular trend. PLoS One 2019;14:e0211532.<br>https://doi.org/10.1371/journal.pone.0211532.<br>[30] MacKrill K, Groom KM, Petri reporting and secular trend. PLoS One 2019;14:e0211532.<br>https://doi.org/10.1371/journal.pone.0211532.<br>MacKrill K, Groom KM, Petrie KJ. The effect of symptom-tracking apps on symptom repo<br>Br J Health Psychol 2020;25:1074–85
- https://doi.org/10.1371/journal.pone.0211532.<br>MacKrill K, Groom KM, Petrie KJ. The effect of symptom-tra<br>Br J Health Psychol 2020;25:1074–85. https://doi.org/10.1.<br>. MacKrill K, Groom KM, Petrie KJ. The effect of sy<br>Br J Health Psychol 2020;25:1074–85. https://do<br>Thealth Psychol 2020;25:1074–85. https://do  $[30]$  Machine, Groom May, Petric Machine Street of Symptom-tracking appearing appearing.<br>Br J Health Psychol 2020;25:1074-85. https://doi.org/10.1111/bjhp.12459. Br J Health Psychol 2020;25:1074–85. https://doi.org/10.1111/bjhp.12459.

## TABLES

Table 1: Demographic data for study participants.  $\mathbb B$  indicates that the variable was used in the<br>matching process. Continuous variables were compared using the paired Wilcoxon test; categorical<br>variables were compared variables were compared using the chi-squared test. \* indicates significant difference between the<br>two groups. § included here: self-reported categories of: mania, hypomania, bipolar, manic<br>depression, schizophrenia, psych two groups. § included here: self-reported categories of: mania, hypomania, bipolar, manic<br>depression, schizophrenia, psychosis, psychotic illness. Smoking refers to currently smoking. IMD<br>(index of multiple deprivation) i the groups. Schizophrenia, psychosis, psychotic illness. Smoking refers to currently smoking.<br>(index of multiple deprivation) is ordered from 1 (most deprived) to 10 (least deprived). Met<br>illness duration was calculated <u>w</u> (index of multiple deprivation) is ordered from 1 (most deprived) to 10 (least deprived). Median<br>illness duration was calculated <u>without</u> considering the possibility of right-censoring at 12 weeks is<br>the group with long i illness duration was calculated <u>without</u> considering the possibility of right-censoring at 12 weeks<br>the group with long illness duration. Right censoring occurred for 338 individuals who were still<br>logging as unhealthy at illness duration. Right censoring occurred for 338 individuals who were still<br>logging as unhealthy at the time of study cut-off and for 166 individuals who stopped logging while<br>still unhealthy.<br>Matched individuals<br>with sh the group with short illness<br>
dogging as unhealthy at the time of study cut-off and for 166 individuals who stopped logging where still unhealthy.<br>
Matched individuals<br>
with short illness<br>
Individuals with long





[IQR]) 20[16; 23] 20[15; 23] Z20 [16;23]

## FIGURES

Figure 1: Flowchart for Study Participant Selection. LFT - lateral flow test. PCR - polymerase chain reaction test. MH – mental health. Weekly logging – at least one health report logged weekly from 8 weeks before 12 weeks after commencement of COVID-19. Invalid demographics: BMI < 15 or BMI > 55; age >100 years or age < 18 years, no possibility to extract index of multiple deprivation.



Figure 2: Symptom prevalence during the baseline period in individuals with long illness vs. short

illness (descriptive data only, unadjusted for comorbidities, week of testing, prior infection,

vaccination status, smoking or index of multiple deprivation).



Figure 3: Odds ratios from conditional logistic regression of individual symptoms for long illness vs. short illness (reference group), according to baseline symptom reporting. Model 1: no adjustment; Model 2: additionally adjusted for any comorbidity reported at registration; Model 3: additionally adjusted for prior mental health diagnosis. Circle size represents baseline symptom prevalence in individuals with short (gold) vs. long (purple) illness duration; circle scale is shown at the bottom in grey. Odds ratios are shown as dots with 95% confidence intervals as lines: results in red are significant after adjustment for multiple comparisons.



## Figure 4: Concordance of symptoms between baseline and post-COVID periods, in individuals with short (left image) vs. long (right image) illness (n=1350 in each group).



# Figure 5: Odds ratios of symptom concordance (i.e., present in the post-COVID period, if reported at baseline [reference period]) in individuals with long illness. Model 1: adjusted for age, sex, BMI, vaccination number, prior infection, week of testing, smoking and index of multiple deprivation; Model 2: additionally adjusted for comorbidities reported at registration; Model 3: additionally adjusted for prior mental health diagnosis. Circle size refers to symptom prevalence during baseline (gold) and post-COVID (purple) periods; scale is shown at bottom of figure. Symptoms are ordered by decreasing prevalence at baseline. Odds ratios are shown as dots with CI (lines); results in red are significant after adjustment for multiple comparisons.



## Figure 6: Symptom prevalence by sex, considered during baseline (left two panels) and post-COVID (right two panels) periods, in individuals with short (upper two panels) and long (lower two panels) illness.





Figure 7: Symptom prevalence in individuals with long illness according to baseline symptom reporting (any symptom reported vs. no symptom); odds ratios of each individual symptom in the post-COVID period according to baseline symptom reporting (any reported symptom vs. no symptom). Red lines indicate significantly increased odds. Symptoms are ordered by decreasing prevalence at baseline.



Percentage of individuals reporting at least once<br>iven symptom in the post period according to baseline reporting

Figure 8 Symptom prevalence in individuals with long illness duration, comparing individuals with respiratory illness (asthma and/or 'lung disease', allergic rhinitis), mental health diagnosis, heart disease or diabetes to individuals without any comorbidity, during baseline (left panel) and post-COVID (right panel) periods. MH: prior mental health diagnosis. Symptoms are ordered by decreasing prevalence at baseline.

