

**Long-term symptom profiles after COVID-19 versus other acute respiratory infections (COVIDENCE UK): a population-based observational study**

*Giulia Vivaldi MSc<sup>1,2</sup>, Paul E Pfeffer PhD<sup>3,4</sup>, Mohammad Talaei PhD<sup>2</sup>, Jayson Basera MSc<sup>1</sup>, Seif O Shaheen PhD<sup>2</sup>, Adrian R Martineau PhD<sup>1,2,5</sup>*

<sup>1</sup> Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK

<sup>2</sup> Wolfson Institute of Population Health, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK

<sup>3</sup> Barts Health NHS Trust, London, UK

<sup>4</sup> Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK

<sup>5</sup> Asthma UK Centre for Applied Research, Queen Mary University of London, London, UK

Correspondence to:

Giulia Vivaldi, Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London E1 2AT, UK

**[g.vivaldi@qmul.ac.uk](mailto:g.vivaldi@qmul.ac.uk)**

Word count: 295 abstract, 4994 main paper

## Summary

**Background** Long COVID is a well recognised, if heterogeneous, entity. Acute respiratory infections (ARIs) due to other pathogens may cause long-term symptoms, but few studies compare post-acute sequelae between SARS-CoV-2 and other ARIs. We aimed to compare symptom profiles between people with previous SARS-CoV-2 infection, people with previous non-COVID-19 ARIs, and contemporaneous controls, and to identify clusters of long-term symptoms.

**Methods** COVIDENCE UK is a prospective, population-based UK study of ARIs in adults. We analysed data for 16 potential long COVID symptoms and health-related quality of life (HRQoL), reported between Jan 21 and Feb 15, 2021, by participants unvaccinated against SARS-CoV-2. We classified participants as having previous SARS-CoV-2 infection or previous non-COVID-19 ARI ( $\geq 4$  weeks prior) or no reported ARI. We compared symptoms by infection status using logistic and fractional regression, and identified symptom clusters using latent class analysis (LCA). This study is registered with ClinicalTrials.gov, NCT04330599.

**Findings** We included 10,203 participants (1343 [13.2%] with SARS-CoV-2 infection, 472 [4.6%] with non-COVID-19 ARI). Both types of infection were associated with increased prevalence/severity of most symptoms and decreased HRQoL compared with no infection. Participants with SARS-CoV-2 infection had increased odds of taste/smell problems (odds ratio 1.74, 95% CI 1.05–3.70) and lightheadedness or dizziness (1.74, 1.18–2.56) compared with participants with non-COVID-19 ARIs. Separate LCA models identified three symptom severity groups for each infection type. In the most severe groups (including 23% of participants with SARS-CoV-2, and 21% with non-COVID-19 ARI), SARS-CoV-2 infection presented with a higher probability of taste/smell problems (probability 0.41 vs 0.04), memory problems (0.70 vs 0.55), and hair loss (0.25 vs 0.16) than non-COVID-19 ARI.

**Interpretation** Both SARS-CoV-2 and non-COVID-19 ARIs are associated with a wide range of long-term symptoms. Research on post-acute sequelae of ARIs should extend from SARS-CoV-2 to include other pathogens.

**Funding** Barts Charity.

## **Research in context**

### **Evidence before this study**

We searched PubMed and Google Scholar for studies on post-acute sequelae of COVID-19 and other acute respiratory infections (ARIs), published up to May 24, 2023. We used search terms relating to COVID-19 and other ARIs (“COVID-19”, “SARS”, “severe acute respiratory syndrome”, “Middle East respiratory”, “MERS”, “respiratory infection”, “influenza”, “flu”) and post-acute symptoms (“long COVID”, “post-acute”, “PACS”, “sequelae”, “long-term”). Previous studies have shown a wide range of post-acute sequelae for COVID-19, affecting people with all severities of the acute disease. The few studies comparing long-term symptoms between people with COVID-19 and non-COVID-19 ARIs have generally found a higher symptom burden among the former; however, these studies have been restricted to hospitalised patients or electronic health record data, and thus do not capture the full picture in the community.

### **Added value of this study**

In this population-based study of ARIs in the community, we observed high symptom burden among people with previous SARS-CoV-2 infection when compared with controls, highlighting the extensive reach of long COVID. Our finding of a similar symptom burden among people with non-COVID-19 ARIs suggests that post-acute sequelae of other ARIs may be going unrecognised. For both types of infection, latent class analyses grouped participants according to overall symptom severity, rather than symptom types, suggesting that overall symptom burden may best characterise the experience of people with post-acute sequelae.

### **Implications of all the available evidence**

The long-term symptoms experienced by some people with previous ARIs, including SARS-CoV-2, highlights the need for improved understanding, diagnosis, and treatment of post-acute infection syndromes. As much-needed research into long COVID continues, we must take the opportunity to investigate and consider the post-acute burden of ARIs due to other pathogens.

## Introduction

Long COVID—the post-acute sequelae of SARS-CoV-2 infection—is a highly heterogeneous condition. It can be broadly defined as new or ongoing symptoms more than 4 weeks after the acute infection,<sup>1</sup> with some of the most common being fatigue, dyspnoea, and cognitive impairment.<sup>2-5</sup> However, more than 200 symptoms have been investigated,<sup>4</sup> making it difficult to precisely characterise, diagnose, or treat the condition.<sup>6</sup> Its high heterogeneity and multisystem nature has encouraged the search for different long COVID phenotypes, with the aim of improving understanding of the condition and targeting treatment options for different phenotypes.

Long COVID is estimated to affect at least 10% of people infected with SARS-CoV-2, with far higher incidence among those hospitalised.<sup>6</sup> With more than 670 million people infected to date,<sup>7</sup> the number of people globally affected by long COVID may be estimated at 67 million, and will almost certainly be higher.<sup>6</sup> Importantly, this high incidence of long COVID has taken place over a short period of time—the 3 years of the COVID-19 pandemic—meaning that awareness of these post-acute sequelae has grown much more rapidly than for other coronavirus pandemics with smaller reach and a lower percentage of survivors.<sup>8,9</sup>

Post-acute infection syndromes are not a new phenomenon; indeed, many cases of chronic fatigue syndrome are reported to follow an infection-like episode.<sup>10</sup> Nonetheless, these syndromes often go undiagnosed owing to the wide range of symptoms and lack of diagnostic tests.<sup>10</sup> Studies into Middle East respiratory syndrome, SARS-CoV-1, and influenza have identified chronic pulmonary and extrapulmonary sequelae a year after the acute infection,<sup>8</sup> but research has largely been restricted to people with severe disease. By contrast, long COVID has been studied in people with all levels of severity of the acute infection,<sup>2,6,11</sup> and has been found to affect even those with mild initial symptoms.<sup>11</sup>

This spotlight on the long-term effects of SARS-CoV-2 infection leads to the question of whether there are post-acute sequelae of other acute respiratory infections (ARIs) in the community, including those that present with mild initial symptoms, and how these sequelae compare with long COVID. Importantly, given the non-specific nature of symptoms such as fatigue, any comparison needs to take into account the symptom prevalence in a contemporaneous, uninfected population—particularly given the background of a disruptive pandemic, with increased stressors, limitations on movements, and elevated levels of population fatigue.<sup>12</sup> We therefore aimed to compare symptom profiles between people with previous SARS-CoV-2 infection, people with previous non-COVID-19 ARIs, and those with no reported infection, and then to identify distinct symptom clusters among people with previous SARS-CoV-2 infections or non-COVID-19 ARIs.

## **Methods**

### **Study participants**

COVIDENCE UK is a prospective, longitudinal, population-based observational study of COVID-19 in the UK population (<https://www.qmul.ac.uk/covidence>).<sup>13</sup> Inclusion criteria were age 16 years or older and UK residence at enrolment, with no exclusion criteria. Participants were invited via a national media campaign via print and online newspapers, radio, television, social media, and online advertising. Further details on the recruitment strategy can be found in the appendix (p 2). Enrolled participants completed an online baseline questionnaire and monthly follow-up questionnaires to capture information on potential symptoms of COVID-19, results of nose or throat swab tests for SARS-CoV-2, SARS-CoV-2 vaccine status, and details of a wide range of potential long COVID symptoms. The study was launched on May 1, 2020, and closed to enrolment on Oct 6, 2021. The final COVIDENCE UK cohort was majority female (70.2%) and White (93.7%), with under-representation of people younger than 50 years, men, and minoritised ethnicities.<sup>13</sup> This study is registered with ClinicalTrials.gov, NCT04330599. COVIDENCE UK was approved by Leicester South Research Ethics Committee (ref 20/EM/0117). All participants gave informed consent to take part in the study before enrolment.

For this analysis, we included all COVIDENCE UK participants responding to the January, 2021, follow-up questionnaire who had not been vaccinated against SARS-CoV-2. We focused on this date for several reasons. During this period, all COVIDENCE UK participants were asked a panel of questions about potential long COVID symptoms, which allowed us to assess prevalence of potential long COVID symptoms regardless of reported previous infections or self-reported long COVID. Additionally, January, 2021, was sufficiently early in the vaccine rollout that most of our participants were not vaccinated, and sufficiently late into the pandemic that we were able to record a substantial number of infections.

Participants were categorised as having had a previous SARS-CoV-2 infection, a previous non-COVID-19 ARI, or no previous recorded infection. Participants who reported both a previous SARS-CoV-2 infection and a previous non-COVID-19 ARI were excluded. To minimise infection misclassification, we also excluded participants who reported symptom-defined non-COVID-19 ARIs that were not accompanied by a negative SARS-CoV-2 swab test (lateral flow or RT-PCR). Finally, we excluded any participants who had either infection fewer than 4 weeks before the survey date, in order to focus on long-term symptoms.

### **Previous infections**

We defined previous SARS-CoV-2 infection as any of the following: a positive swab test for SARS-CoV-2, a positive antibody test for SARS-CoV-2, or symptom-defined probable COVID-19, based on the algorithm described by Menni et al.<sup>14</sup> Infection dates were defined as the date of the test, for swab tests, and the date of symptom onset, for participants with symptom-defined COVID-19. Participants reporting positive antibody tests were asked to recall a suspected infection date (based on symptoms, swab test results, or close proximity to a COVID-19 case); participants unable to provide a date who nonetheless reported probable COVID-19 symptoms<sup>14</sup> before their antibody test were assigned the symptom onset date as their infection date. Participants with positive swab or antibody tests and no reported symptoms were defined as asymptomatic cases.

We defined a previous non-COVID-19 ARI as self-report of a general practitioner or hospital diagnosis of pneumonia, influenza, bronchitis, tonsillitis, pharyngitis, ear infection, common cold, or other upper or lower respiratory infection not caused by SARS-CoV-2; or self-report of a symptom-defined ARI accompanied by a negative SARS-CoV-2 swab test (lateral flow or RT-PCR). Symptom-defined episodes of ARI were identified using modified Jackson criteria for upper respiratory infection,<sup>15</sup> modified Macfarlane criteria for lower respiratory infection,<sup>16</sup> and a triad of cough, fever, and myalgia for influenza-like illness.<sup>17</sup> Infection dates were defined as the date of symptom onset.

## **Symptoms**

We assessed the prevalence or severity of 16 potential COVID-19 symptoms: coughing, problems with sleep, memory problems, difficulty concentrating, muscle or joint pain, problems with sense of taste or smell, diarrhoea, stomach problems (abdominal pains), changes to voice, hair loss, unusual racing of the heart, lightheadedness or dizziness, unusual sweating, breathlessness, anxiety or depression, and fatigue.<sup>18</sup> We measured breathlessness with the Medical Research Council (MRC) Dyspnoea Scale,<sup>19</sup> anxiety or depression with the Patient Health Questionnaire-4 (PHQ-4),<sup>20</sup> and fatigue with the Functional Assessment of Chronic Illness Therapy (FACIT) Fatigue Score.<sup>21</sup> For symptoms measured on a scale (dyspnoea, anxiety or depression, and fatigue), we refer to the severity of symptoms; for all other symptoms, we refer to the prevalence (ie, whether they were present or absent).

We also assessed health-related quality of life (HRQoL) using the EQ-5D-3L,<sup>22</sup> which covers five dimensions of health—mobility, self-care, usual activities, pain or discomfort, and anxiety or depression—and includes a visual analogue scale (VAS) to record the respondent's self-rated health status on a scale from 0 to 100.

## Statistical analysis

All analyses considered symptoms at a single timepoint. For our first analysis, we compared prevalence or severity of potential long COVID symptoms and HRQoL between participants with previous SARS-CoV-2 infection, those with non-COVID-19 ARIs, and those with no reported infections. To explore the long-term effects of COVID-19, we additionally compared symptoms between participants with a SARS-CoV-2 infection more than 12 weeks prior and those with no infection. When comparing previous SARS-CoV-2 infection with non-COVID-19 ARIs, we restricted the analysis to participants with symptomatic SARS-CoV-2 infection, as participants had no means of reporting asymptomatic non-COVID-19 ARIs. Finally, we explored how symptoms differed among participants with previous SARS-CoV-2 infection according to time since infection (4–12 weeks vs >12 weeks) and self-reported severity of the initial infection, classed as either asymptomatic (no symptoms reported), mild (able to do most of usual activities), moderate (unable to do usual activities, but without requiring bedrest), severe (requiring bedrest), and hospitalised.

Regressions were carried out in all participants, adjusted for potential confounders that could influence both risk or severity of infection and likelihood of experiencing or reporting symptoms: age, sex, highest educational attainment, baseline self-reported general health, body-mass index, asthma or chronic obstructive pulmonary disease, and having received an influenza vaccination in the previous 6 months. Analyses were additionally adjusted for time since enrolment, to account for differential follow-up. For analyses restricted to participants with infections, we instead adjusted for time since infection to take into account the changing or waning of symptoms over time. Categorical variables were analysed using logistic regression or ordered logistic regression, as appropriate. Logistic regression models were checked for misspecification and goodness of fit. The proportional odds assumption for ordinal variables (MRC Dyspnoea, PHQ-4, EQ-5D Activities, and EQ-5D Pain) was assessed using the Brant test.<sup>23</sup> Models with significant deviations were refitted using a partial proportional odds model;<sup>24</sup> as deviations are common with large sample sizes, and proportional odds models are less parsimonious, model fit was compared using Lacy's adjusted ordinal explained variation measure.<sup>25</sup> As our continuous variables (FACIT-13 and the EQ-5D Visual Analogue Scale) were bounded, which can create problems with linear regression, we rescaled them to the [0,1] interval and analysed them using fractional regression.<sup>26</sup> Regression results are presented as odds ratios for both types of logistic regression and as predicted percentage point change in outcome for fractional regressions, to aid interpretation. Further details on these analyses can be found in the appendix (pp 3–4). We adjusted for multiple testing using the Benjamini–Hochberg procedure,<sup>27</sup> applying corrections to the 21 outcomes being simultaneously assessed in each subgroup.

We carried out four sensitivity analyses: first, we repeated the analyses in participants who had swab-test or antibody-test confirmed COVID-19 (ie, excluding any participants with symptom-defined potential COVID-19). Second, to better account for baseline health differences that may have influenced susceptibility to infection, we repeated analyses by infection status excluding those whose reported infections occurred before cohort enrolment, so that the cohort baseline general health would represent health status before infection. Third, to compare the long-term effects of SARS-CoV-2 and non-COVID-19 ARIs, we repeated the comparison between participants with previous infections restricted to participants whose infections occurred more than 12 weeks prior. Finally, to better account for the longer time since infection among participants with SARS-CoV-2 compared with non-COVID-19 ARIs, we repeated the comparison between participants with previous infections restricted to the period of overlap between the two groups (ie, time since infection between 28 and 260 days).

For our second analysis, we carried out separate latent class analyses (LCA) in participants who had previous SARS-CoV-2 infection or a previous non-COVID-19 ARI, to explore whether there was any clustering of symptoms. LCA is a statistical method of identifying subgroups in a given population who share measured characteristics, such as certain symptoms. The underlying assumption is that different patterns of symptoms can be explained by membership to a certain group or class. With this approach, we hoped to identify distinct symptom phenotypes among people who had a previous SARS-CoV-2 infection or a non-COVID-19 ARI.

We fitted latent class models with up to ten classes, and identified the optimal number of classes by considering changes in the Bayesian Information Criterion (BIC), the Vuong-Lo-Mendell-Rubin (VLMR) adjusted likelihood ratio test, and the bootstrapped likelihood ratio test.<sup>28</sup> We additionally considered the interpretability and separation of any identified classes, as well as examining classification diagnostics such as model entropy and average latent class posterior probabilities.<sup>28</sup> Once the number of classes had been chosen, we attempted to improve model fit by considering direct effects between pairs of indicators that showed violations in the independence assumption. Finally, we adjusted for age, sex, and time since infection in the final model by including them as covariates.<sup>29</sup> Owing to problems with convergence, we converted responses to the five dimensions of the EQ-5D-3L into a weighted health state index, using EQ-5D preference weights obtained from the UK general population.<sup>30</sup> The upper bound of 1 represents full health, whereas a value of 0 represents death. Negative values are allowed (ie, states considered to be worse than death).

Participants with missing data were excluded from regression analyses. In LCA, missing data were not imputed; classifications for each individual were done with the data available.



Analyses were done in Stata (version 17.0) and LatentGOLD (version 6.0).

**Role of the funding source**

The funder of the study had no role in study design, data collection, data analysis, data interpretation, writing of the report, or the decision to submit the paper for publication.

## Results

15,668 (89.0%) of 17,612 participants enrolled in COVIDENCE UK completed a questionnaire between Jan 21 and Feb 15, 2021 (appendix p 1; appendix figure S1). 10,171 of these participants were included in the analysis, with a median age of 62.8 years (IQR 53.7–68.9). 6994 (68.8%) participants were female and the vast majority were White ([table 1](#)). 1311 (12.9%) had a previous SARS-CoV-2 infection and 472 (4.6%) had a previous non-COVID-19 ARI ([table 1](#)). 243 (18.5%) of participants with previous SARS-CoV-2 infection reported a positive swab test and 304 (23.2%) a positive antibody test; the remaining participants had symptom-defined infections. On average, participants with SARS-CoV-2 had been infected earlier than participants with non-COVID-19 ARIs, with the majority of participants with previous SARS-CoV-2 infection infected more than 12 weeks earlier ([table 1](#)). Participants with SARS-CoV-2 infection were also more likely to have been hospitalised with their acute infection ([table 1](#)). Participants with no infection were the least to be frontline workers, whereas participants with non-COVID-19 ARIs were the most likely to have asthma ([table 1](#)). Symptom prevalence by infection status, timing, and severity is shown in the appendix (figures S2–S4).

We observed increased prevalence or severity of all symptoms, accompanied by decreased HRQoL, among participants with previous SARS-CoV-2 infection compared with those without any infection, regardless of whether their infection had been more than 12 weeks prior ([table 2](#)). The greatest difference observed was in the prevalence of problems with taste or smell, although this was slightly reduced among participants who had been infected more than 12 weeks prior ([table 2](#)). When compared with participants with no infection, those with non-COVID-19 ARIs also showed increased prevalence or severity of most symptoms considered, with the exception of muscle or joint pain, problems with sense of taste or smell, and hair loss ([table 2](#)).

When compared with participants with non-COVID-19 ARI, those with SARS-CoV-2 infection had greater odds of reporting problems with sense of taste or smell and lightheadedness or dizziness, but few other differences were observed ([table 2](#)).

Among participants with SARS-CoV-2 infection, those who had been infected more than 12 weeks prior were less likely to report coughing or problems with taste or smell, and had lower levels of dyspnoea, than those infected 4–12 weeks prior ([table 3](#)). Additionally, participants with a more severe acute infection were more likely to report increased prevalence or severity of ongoing symptoms (appendix table S4).

We obtained similar results in our sensitivity analyses, although some analyses were affected by a lack of power (appendix tables S5–S10). We additionally found little evidence that our findings were driven by more severe SARS-CoV-2 infections: an exploratory analysis comparing symptoms between participants with asymptomatic or mild SARS-CoV-2 infection and those with no infection continued to show significant differences between the two groups for all symptoms and measures, except for self-care and mobility (data not shown).

For the LCA in participants with previous SARS-CoV-2 infection, our final model was a three-class model, with an entropy of 0.948 and a strong distinction between classes (appendix table S12); further details on model selection are given in the appendix (pp 14–16). The three classes can be interpreted as different levels of ongoing symptom severity: mild (comprising 45% of participants), moderate (32% of participants), and severe (22% of participants). [Figure 1A](#) shows the symptom profile for these three classes. The mild class presents with generally low prevalence and severity of the symptoms considered, with the most frequent symptom being sleep problems ([figure 1](#)). Participants assigned to the moderate class present with a slightly increased probability of reporting all symptoms, but with a larger increase in muscle or joint pain, sleep problems, difficulty concentrating, and PHQ-4 score ([figure 1A](#)). The final class, comprising nearly a quarter of participants with previous SARS-CoV-2 infection, shows a marked increase in the prevalence and severity of all symptoms, with the greatest increases among memory problems, difficulty concentrating, and lightheadedness or dizziness ([figure 1A](#)). When comparing participant characteristics across the three classes, we observed that with increasing symptom severity, participants were more likely to be female, socioeconomically deprived, frontline workers, overweight or obese, and to have comorbidities (appendix table S14). Ongoing symptom severity also appeared to be linked to the severity of the initial infection: the proportion of participants who were either very unwell or hospitalised with their initial infection increased from 33% in the mild class to 61% in the severe class (appendix table S14); notably, however, 39% of participants in the severe class had an initial infection of mild-to-moderate severity (appendix table S14).

Finally, as ongoing symptom severity increased, participants were more likely to have reported having long COVID in the current questionnaire, ranging from 5.9% in the mild class to nearly half of participants in the severe class (appendix table S14). However, even in the severe class, many participants did not report themselves as having long COVID (appendix table S14).

Our final model for non-COVID-19 ARIs was also a three-class model representing different levels of severity (40% mild, 38% moderate, and 22% severe). This model had an entropy of 0.950 and showed strong distinction between classes (appendix table S12). To enable visual comparisons

between participants by infection status, we additionally carried out an LCA in participants with no infection, resulting in a two-class model representing two levels of symptom severity (56% mild, 44% severe), and overlaid the symptom profiles of the severe class from each LCA ([figure 1B](#)). In this comparison, participants with non-COVID-19 ARI or SARS-CoV-2 infection showed a higher probability of reporting almost all symptoms than participants with no infection, particularly cognitive problems ([figure 1B](#)). However, symptom profiles seem to differ by ARI among those most severely affected: participants with SARS-CoV-2 infection showed a higher probability of reporting memory problems, difficulty concentrating, unusual sweating, hair loss, and problems with sense of taste or smell compared with non-COVID-19 ARI ([figure 1B](#)). Notably, a tenth of participants with previous non-COVID-19 ARI and severe symptoms attributed those symptoms to long COVID (appendix table S15). When comparing participant characteristics across the three groups with the most severe symptoms, participants with no reported infection were more likely to be older than those with previous respiratory infections and presented with fewer symptoms overall, whereas participants with non-COVID-19 ARIs had a much higher prevalence of heart disease than the other groups (appendix table S15). Given this finding, we did an exploratory analysis comparing the prevalence and severity of all symptoms between participants with SARS-CoV-2 infection and those with non-COVID-19 ARI, additionally adjusting for heart disease, but this did not materially affect the results (data not shown).

## Discussion

In this large, observational study, we found that previous SARS-CoV-2 infection was associated with increased prevalence and severity of a wide range of symptoms—covering gastrointestinal, neurological, musculoskeletal, and cardiopulmonary problems—as well as lower HRQoL, and that this increased symptom burden persisted more than 12 weeks after the acute infection. Participants with SARS-CoV-2 infection were more likely to report problems with taste or smell and lightheadedness or dizziness than those with non-COVID-19 ARI, but we observed little difference in other symptoms or HRQoL measures. The lower burden of coughing, problems with taste or smell, and dyspnoea among participants who had been infected with SARS-CoV-2 more than 12 weeks prior compared with those with more recent infections suggests these symptoms may be the first to show improvement; however, other symptoms considered showed little difference between remote and more recent infections. A severe initial SARS-CoV-2 infection—either requiring bedrest or hospitalisation—was found to be associated with a greater prevalence of ongoing symptoms, and reduction in HRQoL. Finally, as ongoing symptom severity increased, participants were more likely to report having long COVID, with nearly half of participants with severe symptoms reporting suspected long COVID.

The scale and fast spread of the COVID-19 pandemic, alongside a lower case–fatality ratio than previous coronavirus pandemics,<sup>31</sup> has resulted in hundreds of millions of survivors globally over just a few years, focusing attention on their post-COVID-19 experience. While post-acute sequelae of other viral respiratory infections have been observed,<sup>32</sup> they have not been as well characterised. Similar to our findings for previous SARS-CoV-2 infection, we observed increased burden of many symptoms among participants with previous non-COVID-19 ARI when compared with no infection. However, long-term symptom profiles differed slightly between SARS-CoV-2 infection and non-COVID-19 ARI, with the former showing greater increases in problems with taste or smell and lightheadedness or dizziness. Our findings suggest that there may be long-lasting health impacts from other respiratory infections that are going unrecognised, although we do not yet have evidence that these symptoms have a similar duration to long COVID. Our cohort of community infections will not represent those worst affected by long COVID, and so the elevated symptom burden seen for both SARS-CoV-2 infections and non-COVID-19 ARIs is likely to represent a milder phenotype of post-acute sequelae. Indeed, retrospective cohort studies have generally found worse post-acute outcomes among patients with SARS-CoV-2 compared with influenza<sup>33,34</sup> and other viral infections;<sup>35</sup> these studies have largely been restricted to hospitalised patients or electronic health record data, thus representing people with either a severe acute infection or sufficiently severe post-acute symptoms to seek medical help. Given our findings of post-acute sequelae in people with milder

disease, longitudinal studies are needed to investigate the distinct pathogens responsible and trajectories of recovery.

We found increased risk of ongoing symptoms across different levels of acute disease severity, highlighting the wide reach of long COVID burden. While long COVID has been well documented in previously hospitalised patients,<sup>36-38</sup> and risk of long COVID has been found to be associated with severe COVID-19,<sup>3,39</sup> an increasing number of studies have found long-term sequelae in people with mild or asymptomatic infections.<sup>11,40</sup> Importantly, less than a quarter of our participants with previous SARS-CoV-2 infection had ever reported long COVID—including half of participants with the most severe symptoms—suggesting that some people with ongoing symptoms may not be ascribing these symptoms to the infection, or may not consider their symptoms serious enough to qualify as long COVID. This could exacerbate under-reporting of the condition,<sup>41</sup> likely impacting the healthcare resources made available to people with long-term sequelae. A lack of awareness of post-acute sequelae of other ARIs—or even the lack of a common term, like "long COVID"—is likely to contribute to under-reporting as well. Among participants with the most severe symptoms, 12% with previous non-COVID-19 ARI attributed their symptoms to long COVID compared with 48% with previous SARS-CoV-2 infection, despite their symptom profiles being largely similar, highlighting their need for an alternative diagnosis. Given the few diagnostic tests available, future research needs to focus on enabling diagnosis of long COVID and other post-acute sequelae, to ensure all people with ongoing symptoms can access the support they need.

While exploring long COVID phenotypes, we found that people with previous SARS-CoV-2 infections could largely be classified into three groups representing different overall severity of ongoing symptoms. The most severely affected were characterised by neurocognitive symptoms such as memory problems and difficulty concentrating: these symptoms distinguished them not only from participants with milder ongoing symptoms after SARS-CoV-2 infection, but also from the most severely affected participants with non-COVID-19 ARI. Neurological and cognitive symptoms in long COVID are well documented,<sup>2,42</sup> and some studies have identified neurological symptom clusters as a potential long COVID phenotype.<sup>3,43</sup> However, research so far into phenotypes has been inconclusive, with other studies instead identifying clusters reflecting lower or higher overall symptom severity.<sup>2,44-46</sup> While our findings support the importance of neurocognitive symptoms in long COVID, the clearest distinction between our participants was the overall severity of their symptoms, rather than physiological systems affected. Such severity-based clusters may offer less insight into the underlying mechanisms of long COVID, but it will nonetheless be important to evaluate whether trajectories of recovery differ between these severity groups, in order to better predict outcomes for patients and support them in their recovery.

Our study has several strengths. By comparing symptom burden across three groups, we can paint a clearer picture of how post-acute sequelae differ between SARS-CoV-2 and other ARIs, and how these symptoms differ from those experienced by uninfected population controls. The importance of a control population can be seen clearly when considering muscle and joint pain and sleep problems: while more frequent in people with previous ARI, these symptoms are also widespread in the uninfected population, and so may not be the most characteristic of previous ARIs. Our large sample includes a mix of asymptomatic, symptomatic, and hospitalised cases, allowing us to examine post-acute sequelae across all severities of the acute infection. We recorded the presence and severity of potential long COVID symptoms from all participants in our cohort, regardless of previous SARS-CoV-2 infection status or long COVID reports, which reduces potential reporting bias. This also enabled us to compare responses across all participants at the same point in the pandemic, effectively allowing us to adjust for baseline levels of symptoms such as fatigue and depression, which were likely elevated in the general population owing to the stress of the pandemic.<sup>12</sup> Access to pre-vaccination antibody test data for more than half of our participants, both from COVIDENCE UK serological sub-studies<sup>47</sup> and as reported by our participants, allowed us to more accurately identify previous infections and thus more robustly classify our controls.

This study also has some limitations. First, we focused on symptoms for each individual at a single timepoint, and thus could not map the change in each participant's symptoms over time with repeated measures; however, our aim for this study was to provide a descriptive snapshot of the post-COVID-19 symptom burden, and longitudinal analyses are planned for future work. Second, we did not adjust for pre-COVID-19 symptom burden, as few participants had these data available, meaning that we cannot know whether participants with previous infections had a higher baseline symptom burden than those without. However, our inclusion of all participants with previous SARS-CoV-2 infection, rather than only symptomatic or severe COVID-19, reduces the risk of baseline differences, as many determinants of SARS-CoV-2 infection reflect lifestyle or behavioural factors rather than biological ones.<sup>47,48</sup> Additionally, sensitivity analyses adjusting for pre-infection general health did not substantially affect our results. Third, we do not have details on the type of respiratory infections experienced by our participants reporting non-COVID-19 ARIs, as these are not routinely tested for in the community, preventing us from determining which ARIs are most likely to cause long-term symptoms. However, the fact that we consistently observed differences between our three infection groups, despite the grouping together of different non-COVID-19 ARIs, shows not only that long-term symptoms can occur after non-COVID-19 ARIs, but that these symptoms differ from long COVID. Importantly, our requirement of a negative SARS-CoV-2 swab test for non-COVID-19 ARIs will reduce information bias, particularly as the timing of our study means reported tests are

likely to have been RT-PCR tests, reducing the risk of false negatives.<sup>49</sup> Additionally, a COVIDENCE UK serological sub-study was carried out in December, 2020,<sup>47</sup> meaning we had recent serology data for approximately 40% of our participants, helping to reduce misclassification. Fourth, as COVIDENCE UK questionnaires are issued monthly and completed online, the study may have limitations commonly found in web-based studies,<sup>50</sup> such as participant disinterest and recall bias. In particular, participants may have forgotten to report symptoms from minor ARIs that had resolved several weeks earlier. However, more than 60% of participants with ARIs reported being mildly unwell, suggesting that our findings have not been skewed towards people with the most severe infections, and the questionnaire response rate of 89% shows a high level of engagement among COVIDENCE UK participants. Fifth, the final models chosen for the clustering analysis had evidence of residual covariance between the included variables (appendix p 15), suggesting that the latent class assignments did not fully explain the symptom distribution between participants; nonetheless, the models showed strong distinction between classes, and our findings of clusters defined by symptom severity mirror those of several other studies,<sup>44-46</sup> lending strength to our results. Sixth, our findings are restricted to unvaccinated patients infected with either the wild-type or alpha strains of SARS-CoV-2. However, 29% of people in the UK reporting long COVID were infected with the wild-type strain,<sup>5</sup> which dominated before the vaccination rollout in the UK, showing that our participants represent an important and substantial subgroup with long-lasting symptoms. Seventh, many of our symptoms were recorded as binary measures, requiring participants to classify themselves as either having them or not; more use of scales, such as MRC Dyspnoea or FACIT-13, would have added further nuance to our findings. Finally, COVIDENCE UK is a self-selected cohort,<sup>13</sup> and so certain groups—such as women, older age groups, and White ethnicity—are over-represented in our study, potentially limiting the generalisability of our results.

The COVID-19 pandemic has cast a much-needed spotlight on post-acute infection syndromes, highlighting the need for improved understanding, diagnosis, and treatment of these conditions. While the high symptom burden we observed in participants with previous SARS-CoV-2 infection illustrates the extensive reach of long COVID, the similar burden observed among people with previous non-COVID-19 ARI suggests that the lasting impacts of these infections may be underestimated. As research into long COVID continues, we must take the opportunity to investigate and consider the post-acute burden of other ARIs, to ensure all people with post-acute sequelae can access the treatment and care they deserve.



## **Contributors**

GV, ARM, and PEP conceived the analysis. GV and MT contributed to data management, and have directly accessed and verified the data. Statistical analyses were done by GV, with input from PEP, ARM, MT, SOS, and JB. GV wrote the first draft of the report. All authors critically revised the manuscript for intellectually important content. All authors provided critical conceptual input, interpreted the data analysis, and read and approved the final draft. All authors had full access to all the data in the study and had final responsibility for the decision to submit the manuscript for publication.

## **Data sharing statement**

De-identified participant data will be made available upon reasonable request to the corresponding author.

## **Declaration of interests**

PEP declares grants paid to their institution from the National Institute for Health and Care Research (NIHR) and UK Research and Innovation (UKRI). The remaining authors declare no competing interests.

## **Acknowledgments**

COVIDENCE UK has received support from Barts Charity (MGU0459, MGU0466), Pharma Nord, the Fischer Family Foundation, DSM Nutritional Products, the Exilarch's Foundation, the Karl R Pflieger Foundation, the AIM Foundation, Synergy Biologics, Cytoplan, the UK NIHR Clinical Research Network (52255; 52257), the Health Data Research UK BREATHE Hub, the UKRI Industrial Strategy Challenge Fund (MC\_PC\_19004), Thornton & Ross, Warburtons, Matthew Isaacs (personal donation), Barbara Boucher (personal donation), and Hyphens Pharma. MT was supported by Barts Charity (MGU0570). The views expressed are those of the authors and not necessarily those of the funders. We thank all participants of COVIDENCE UK, and the following organisations who supported study recruitment: Asthma UK/British Lung Foundation, the British Heart Foundation, the British Obesity Society, Cancer Research UK, Diabetes UK, Future Publishing, Kidney Care UK, Kidney Wales, Mumsnet, the National Kidney Federation, the National Rheumatoid Arthritis Society, the North West London Health Research Register (DISCOVER), Primary Immunodeficiency UK, the Race Equality Foundation, SWM Health, the Terence Higgins Trust, and Vasculitis UK.

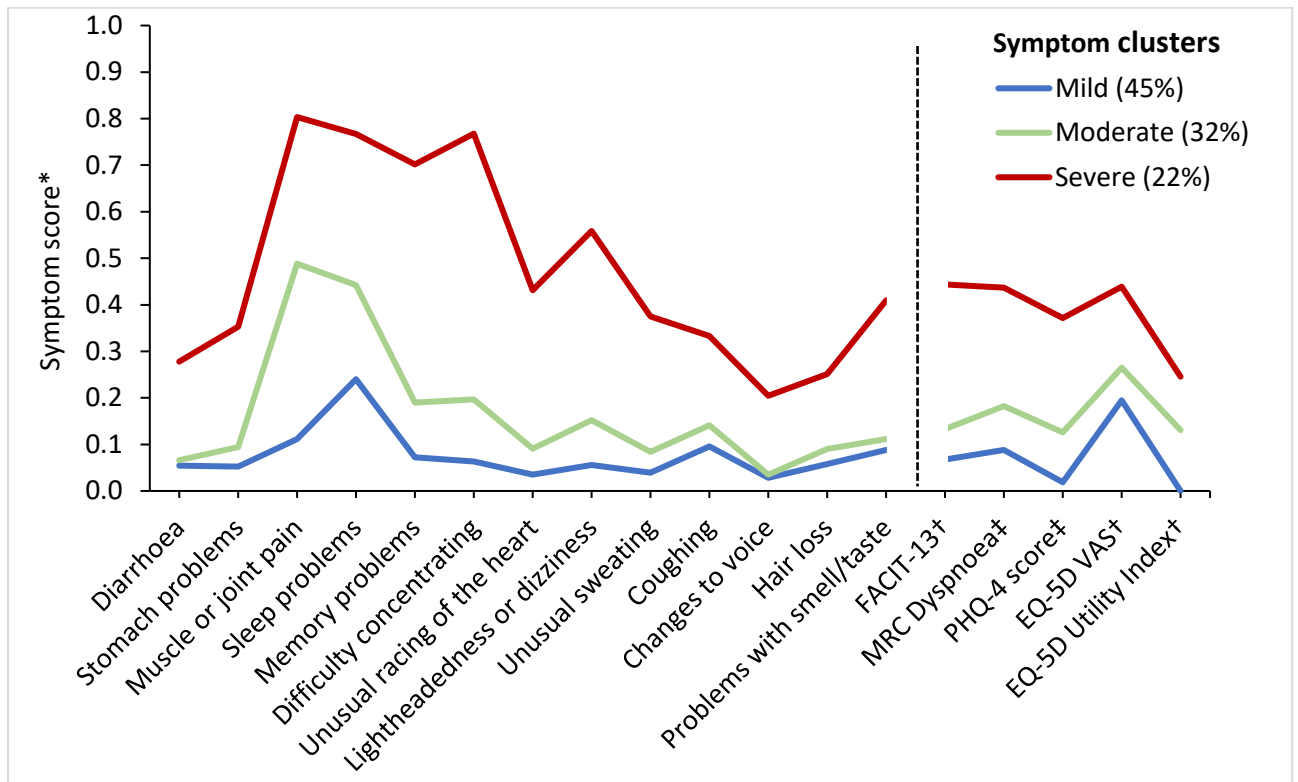
## References

- 1 NICE. National Institute for Health and Care Excellence: Clinical Guidelines. COVID-19 rapid guideline: managing the long-term effects of COVID-19. London: National Institute for Health and Care Excellence, 2022.
- 2 Evans RA, Leavy OC, Richardson M, et al. Clinical characteristics with inflammation profiling of long COVID and association with 1-year recovery following hospitalisation in the UK: a prospective observational study. *Lancet Respir Med* 2022; **10**: 761–75.
- 3 Goldhaber NH, Kohn JN, Ogan WS, et al. Deep Dive into the Long Haul: Analysis of Symptom Clusters and Risk Factors for Post-Acute Sequelae of COVID-19 to Inform Clinical Care. *Int J Environ Res Public Health* 2022; **19**: 16841.
- 4 Davis HE, Assaf GS, McCorkell L, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *EClinicalMedicine* 2021; **38**: 101019.
- 5 Office for National Statistics. Prevalence of ongoing symptoms following coronavirus (COVID-19) infection in the UK: 30 March 2023. London: Office for National Statistics, 2023.
- 6 Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. *Nat Rev Microbiol* 2023; **21**: 133–46.
- 7 Johns Hopkins University. Coronavirus resource center. 2021. <https://coronavirus.jhu.edu/> (accessed Oct 11, 2022).
- 8 Narasimhan H, Wu Y, Goplen NP, Sun J. Immune determinants of chronic sequelae after respiratory viral infection. *Sci Immunol* 2022; **7**: eabm7996.
- 9 O'Sullivan O. Long-term sequelae following previous coronavirus epidemics. *Clin Med (Lond)* 2021; **21**: e68–70.
- 10 Choutka J, Jansari V, Hornig M, Iwasaki A. Unexplained post-acute infection syndromes. *Nat Med* 2022; **28**: 911–23.
- 11 Mizrahi B, Sudry T, Flaks-Manov N, et al. Long covid outcomes at one year after mild SARS-CoV-2 infection: nationwide cohort study. *BMJ* 2023; **380**: e072529.
- 12 Office for Health Improvement and Disparities. COVID-19 mental health and wellbeing surveillance: report. London: Office for Health Improvement and Disparities, 2022.
- 13 Holt H, Relton C, Talaei M, et al. Cohort Profile: Longitudinal population-based study of COVID-19 in UK adults (COVIDENCE UK). *Int J Epidemiol* 2022; published online Sept 29. DOI: 10.1093/ije/dyac189.
- 14 Menni C, Valdes AM, Freidin MB, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. *Nat Med* 2020; **26**: 1037–40.
- 15 Jackson GG, Dowling HF, Spiesman IG, Boand AV. Transmission of the common cold to volunteers under controlled conditions. I. The common cold as a clinical entity. *AMA Arch Intern Med* 1958; **101**: 267–78.
- 16 Macfarlane J, Holmes W, Gard P, et al. Prospective study of the incidence, aetiology and outcome of adult lower respiratory tract illness in the community. *Thorax* 2001; **56**: 109–14.
- 17 Pillonel J, Laperche S. Trends in risk of transfusion-transmitted viral infections (HIV, HCV, HBV) in France between 1992 and 2003 and impact of nucleic acid testing (NAT). *Euro Surveill* 2005; **10**: 5–8.
- 18 Jolliffe DA, Holt H, Greenig M, et al. Effect of a test-and-treat approach to vitamin D supplementation on risk of all cause acute respiratory tract infection and covid-19: phase 3 randomised controlled trial (CORONAVIT). *BMJ* 2022; **378**: e071230.
- 19 UK Research and Innovation. MRC Dyspnoea Scale. 2022. <https://www.ukri.org/councils/mrc/facilities-and-resources/find-an-mrc-facility-or-resource/mrc-dyspnoea-scale/> (accessed June 16, 2023).
- 20 Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics* 2009; **50**: 613–21.

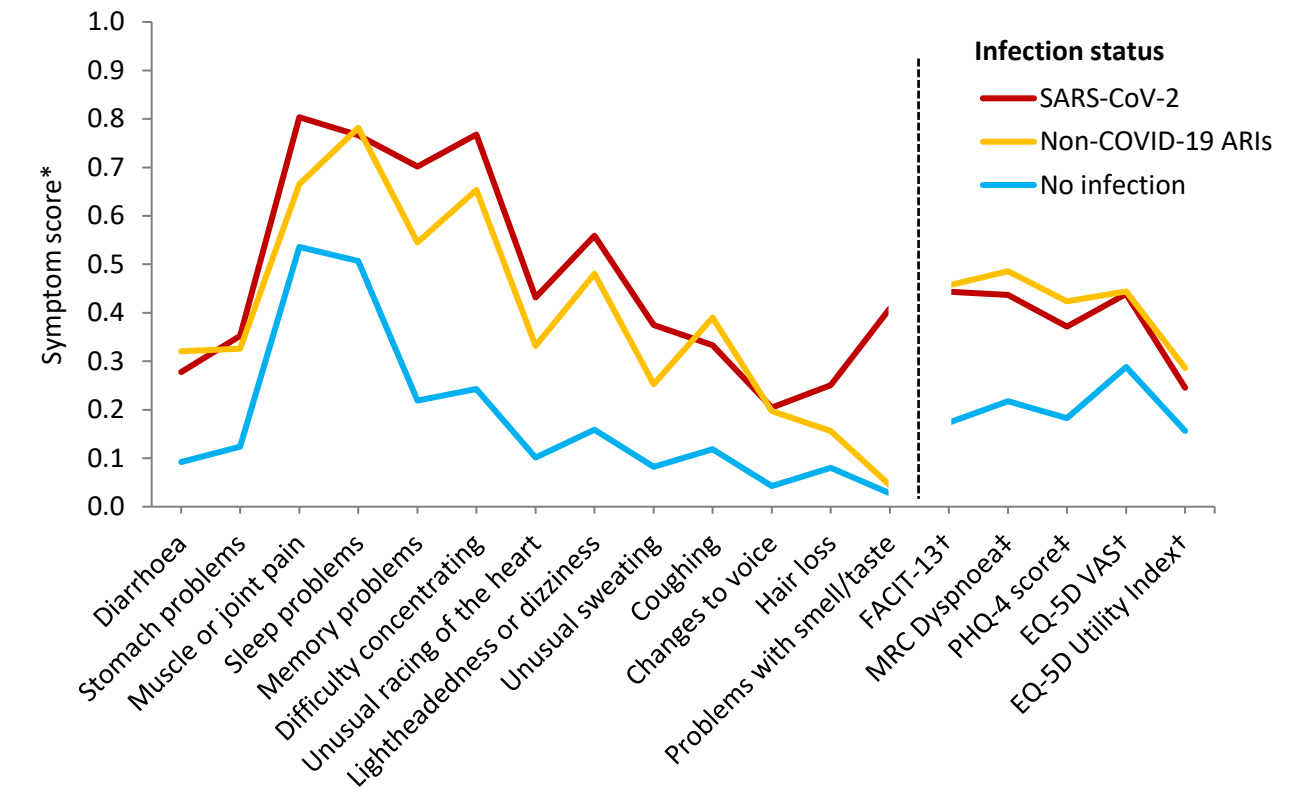
- 21 Yellen SB, Cella DF, Webster K, Blendowski C, Kaplan E. Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. *J Pain Symptom Manage* 1997; **13**: 63–74.
- 22 Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001; **33**: 337–43.
- 23 Brant R. Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. *Biometrics* 1990; **46**: 1171–78.
- 24 Williams R. Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata J* 2006; **6**: 58–82.
- 25 Lacy MG. An Explained Variation Measure for Ordinal Response Models With Comparisons to Other Ordinal R<sup>2</sup> Measures. *Sociol Methods Res* 2006; **34**: 469–520.
- 26 Papke LE, Wooldridge JM. Econometric methods for fractional response variables with an application to 401: plan participation rates. *J Appl Econom* 1996; **11**: 619–32.
- 27 Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc* 1995; **57**: 289–300.
- 28 Nylund-Gibson K, Choi AY. Ten frequently asked questions about latent class analysis. *Transl Issues Psychol Sci* 2018; **4**: 440–61.
- 29 Vermunt JK, Magidson J. How to Perform Three-Step Latent Class Analysis in the Presence of Measurement Non-Invariance or Differential Item Functioning. *Struct Equ Model* 2021; **28**: 356–64.
- 30 Szende A, Oppe M, Devlin N, eds. EQ-5D value sets: inventory, comparative review and user guide. Dordrecht, The Netherlands: Springer, 2007.
- 31 Pustake M, Tambolkar I, Giri P, Gandhi C. SARS, MERS and CoVID-19: An overview and comparison of clinical, laboratory and radiological features. *J Family Med Prim Care* 2022; **11**: 10–17.
- 32 Hirschenberger M, Hunszinger V, Sparrer KMJ. Implications of Innate Immunity in Post-Acute Sequelae of Non-Persistent Viral Infections. *Cells* 2021; **10**: 2134.
- 33 Taquet M, Dercon Q, Luciano S, Geddes JR, Husain M, Harrison PJ. Incidence, co-occurrence, and evolution of long-COVID features: A 6-month retrospective cohort study of 273,618 survivors of COVID-19. *PLoS Med* 2021; **18**: e1003773.
- 34 Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. *Nature* 2021; **594**: 259–64.
- 35 Baskett WI, Qureshi AI, Shyu D, Armer JM, Shyu CR. COVID-Specific Long-term Sequelae in Comparison to Common Viral Respiratory Infections: An Analysis of 17 487 Infected Adult Patients. *Open Forum Infect Dis* 2023; **10**: ofac683.
- 36 Wu X, Liu X, Zhou Y, et al. 3-month, 6-month, 9-month, and 12-month respiratory outcomes in patients following COVID-19-related hospitalisation: a prospective study. *Lancet Respir Med* 2021; **9**: 747–54.
- 37 Huang L, Yao Q, Gu X, et al. 1-year outcomes in hospital survivors with COVID-19: a longitudinal cohort study. *Lancet* 2021; **398**: 747–58.
- 38 Evans RA, McAuley H, Harrison EM, et al. Physical, cognitive, and mental health impacts of COVID-19 after hospitalisation (PHOSP-COVID): a UK multicentre, prospective cohort study. *Lancet Respir Med* 2021; **9**: 1275–87.
- 39 Whitaker M, Elliott J, Chadeau-Hyam M, et al. Persistent COVID-19 symptoms in a community study of 606,434 people in England. *Nat Commun* 2022; **13**: 1957.
- 40 Ma Y, Deng J, Liu Q, Du M, Liu M, Liu J. Long-Term Consequences of Asymptomatic SARS-CoV-2 Infection: A Systematic Review and Meta-Analysis. *Int J Environ Res Public Health* 2023; **20**: 1613.
- 41 Walker AJ, MacKenna B, Inglesby P, et al. Clinical coding of long COVID in English primary care: a federated analysis of 58 million patient records in situ using OpenSAFELY. *Br J Gen Pract* 2021; **71**: e806–14.

- 42 Taquet M, Geddes JR, Husain M, Luciano S, Harrison PJ. 6-month neurological and psychiatric outcomes in 236 379 survivors of COVID-19: a retrospective cohort study using electronic health records. *Lancet Psychiatry* 2021; **8**: 416–27.
- 43 Reese JT, Blau H, Casiraghi E, et al. Generalisable long COVID subtypes: findings from the NIH N3C and RECOVER programmes. *EBioMedicine* 2023; **87**: 104413.
- 44 Sivan M, Parkin A, Makower S, Greenwood DC. Post-COVID syndrome symptoms, functional disability, and clinical severity phenotypes in hospitalized and nonhospitalized individuals: A cross-sectional evaluation from a community COVID rehabilitation service. *J Med Virol* 2022; **94**: 1419–27.
- 45 Bowyer RCE, Huggins C, Toms R, et al. Characterising patterns of COVID-19 and long COVID symptoms: evidence from nine UK longitudinal studies. *Eur J Epidemiol* 2023; **38**: 199–210.
- 46 Peluso MJ, Kelly JD, Lu S, et al. Persistence, Magnitude, and Patterns of Postacute Symptoms and Quality of Life Following Onset of SARS-CoV-2 Infection: Cohort Description and Approaches for Measurement. *Open Forum Infect Dis* 2021; **9**: ofab640.
- 47 Talaei M, Faustini S, Holt H, et al. Determinants of pre-vaccination antibody responses to SARS-CoV-2: a population-based longitudinal study (COVIDENCE UK). *BMC Med* 2022; **20**: 87.
- 48 Thomas DR, Fina LH, Adamson JP, et al. Social, demographic and behavioural determinants of SARS-CoV-2 infection: a case-control study carried out during mass community testing of asymptomatic individuals in South Wales, December 2020. *Epidemiol Infect* 2022; **150**: e115.
- 49 Dinnes J, Sharma P, Berhane S, et al. Rapid, point-of-care antigen tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database Syst Rev* 2022; **7**: CD013705..
- 50 Singh S, Sagar R. A critical look at online survey or questionnaire-based research studies during COVID-19. *Asian J Psychiatr* 2021; **65**: 102850.

**A**



**B**



**Figure 1: Symptom profiles among all participant with previous SARS-CoV-2 infection (A) and among participants with the most severe symptoms, by infection status (B)**

Figure shows the conditional probability (left of the dashed line) or mean severity (right of the dashed line) for all symptoms considered, adjusted for age and sex. Parallel lines between the classes indicate a fairly even increase in the probability or severity of the symptoms considered; disproportionate changes in symptom probability or severity are shown by deviations from the parallel. (A) Includes all participants with previous SARS-CoV-2. Displayed probabilities or scores are shown in appendix table S13. (B) Symptom profiles for participants with the most severe symptoms from the three separate latent class analyses are overlaid. MRC=Medical Research Council. VAS=visual analogue scale. PHQ=Patient Health Questionnaire. \*Symptom score represents conditional probability for binary variables and mean severity for ordinal and continuous variables. †Continuous variables have been reversed to aid with interpretation, so that higher values indicate worse severity or health state. ‡Ordinal variable.

	Previous SARS-CoV-2 infection (N=1311)	Previous non-COVID-19 ARI (N=472)	No infection (N=8388)
<b>Sociodemographics</b>			
Age, years	59.3 (50.7–66.1)	57.2 (47.4–65.0)	63.5 (54.8–69.3)
<30	42 (3.2%)	30 (6.4%)	242 (2.9%)
30 to <40	84 (6.4%)	44 (9.3%)	411 (4.9%)
40 to <50	181 (13.8%)	74 (15.7%)	792 (9.4%)
50 to <60	380 (29.0%)	137 (29.0%)	1716 (20.5%)
60 to <70	458 (34.9%)	136 (28.8%)	3,361 (40.1%)
≥70	166 (12.7%)	51 (10.8%)	1866 (22.2%)
Sex			
Female	908 (69.3%)	385 (81.6%)	5685 (67.8%)
Male	403 (30.7%)	87 (18.4%)	2703 (32.2%)
Ethnicity			
White	1229 (93.7%)	453 (96.0%)	8010 (95.5%)
Mixed/multiple/other ethnic groups	44 (3.4%)	10 (2.1%)	206 (2.5%)
South Asian	27 (2.1%)	7 (1.5%)	121 (1.4%)
Black/African/Caribbean/Black British	11 (0.8%)	2 (0.4%)	51 (0.6%)
Country			
England	1176 (89.7%)	430 (91.1%)	7400 (88.3%)
Northern Ireland	20 (1.5%)	11 (2.3%)	144 (1.7%)
Scotland	72 (5.5%)	23 (4.9%)	533 (6.4%)
Wales	43 (3.3%)	8 (1.7%)	306 (3.7%)
No. people per bedroom			
<1	808 (62.1%)	278 (59.1%)	5849 (70.1%)
1 to <2	461 (35.4%)	173 (36.8%)	2341 (28.1%)
2 to <3	30 (2.3%)	17 (3.6%)	145 (1.7%)
3 to <4	3 (0.2%)	2 (0.4%)	4 (<0.1%)
Quartiles of IMD decile			
Q4 (least deprived)	420 (32.1%)	145 (30.7%)	2737 (32.7%)
Q3	289 (22.1%)	103 (21.8%)	2259 (27.0%)
Q2	271 (20.7%)	122 (25.8%)	1699 (20.3%)
Q1 (most deprived)	330 (25.2%)	102 (21.6%)	1683 (20.1%)
Frontline worker			
No	1067 (81.4%)	374 (79.2%)	7367 (87.9%)
Non-health	185 (14.1%)	84 (17.8%)	838 (10.0%)
Health	59 (4.5%)	14 (3.0%)	173 (2.1%)
Highest educational level attained			
Primary or secondary	152 (11.6%)	54 (11.5%)	950 (11.3%)
Higher or further (A levels)	184 (14.0%)	71 (15.1%)	1279 (15.3%)
College or university	589 (44.9%)	190 (40.3%)	3682 (43.9%)
Post-graduate	386 (29.4%)	156 (33.1%)	2467 (29.4%)
<b>Clinical characteristics</b>			
BMI, kg/m <sup>2</sup>	25.5 (22.9–29.2)	25.5 (23.3–29.3)	24.9 (22.5–28.3)
<25	590 (45.0%)	213 (45.3%)	4256 (50.8%)

25 to <30	443 (33.8%)	155 (33.0%)	2632 (31.4%)
≥30	277 (21.1%)	102 (21.7%)	1487 (17.8%)
Asthma	263 (20.1%)	114 (24.2%)	1239 (14.8%)
Atopy	351 (26.8%)	136 (28.8%)	1958 (23.3%)
Autoimmune disease	110 (8.4%)	48 (10.2%)	708 (8.4%)
Cancer			
Past (cured or in remission)	106 (8.1%)	34 (7.2%)	739 (8.8%)
Active treatment	10 (0.8%)	5 (1.1%)	73 (0.9%)
COPD	37 (2.8%)	8 (1.7%)	156 (1.9%)
Diabetes	39 (3.0%)	21 (4.4%)	393 (4.7%)
Heart disease	33 (2.5%)	22 (4.7%)	293 (3.5%)
Hypertension	252 (19.2%)	73 (15.5%)	1803 (21.5%)
Immunodeficiency	12 (0.9%)	2 (0.4%)	43 (0.5%)
Kidney disease	24 (1.8%)	12 (2.5%)	142 (1.7%)
Major neurological conditions	40 (3.1%)	9 (1.9%)	202 (2.4%)
<b>Infections</b>			
SARS-CoV-2 infection type			
Swab test	243 (18.5%)	..	..
Antibody test	304 (23.2%)	..	..
Symptom defined	764 (58.3%)	..	..
Non-COVID-19 ARI infection type			
Influenza-like illness*	..	24 (5.4%)	..
Upper respiratory infection*	..	146 (32.6%)	..
Lower respiratory infection*	..	361 (80.6%)	..
Diagnosis			
Hospital or GP diagnosis	..	55 (11.7%)	..
Symptom defined	..	417 (88.3%)	..
Weeks since infection	44 (34–47)	11 (7–17)	
4–12 weeks	165 (12.6%)	274 (58.1%)	..
>12 weeks	1146 (87.4%)	198 (41.9%)	..
Infection severity†			
Asymptomatic	89/752 (11.8%)	..	..
Mildly unwell	160/752 (21.3%)	317 (67.2%)	..
Moderately unwell	178/752 (23.7%)	89 (18.9%)	..
Very unwell	285/752 (37.9%)	58 (12.3%)	..
Hospitalised	40/752 (5.3%)	8 (1.7%)	..

ARI=acute respiratory infection. BMI=body-mass index. COPD=chronic obstructive pulmonary disease. IMD=Index of Multiple Deprivation. \*Participants could have more than one non-COVID-19 ARI infection type. †For symptomatic and non-hospitalised participants, severity was self-reported with the following statements: “Mildly unwell – I could do most of my usual activities”, “Moderately unwell – I couldn’t do usual activities but didn’t need to go to bed in the daytime”, and “Very unwell – I had to go to bed in the daytime”.

**Table 1: Participant characteristics**



	Previous SARS-CoV-2 infection vs no infection		SARS-CoV-2 infection >12 weeks prior vs no infection		Previous SARS-CoV-2 infection vs previous non-COVID-19 ARI		Previous non-COVID-19 ARI vs no infection	
	Estimate* (95% CI)	p value†	Estimate* (95% CI)	p value†	Estimate* (95% CI)	p value‡	Estimate* (95% CI)	p value§
Coughing	2.02 (1.70–2.40)	<0.001	1.80 (1.49–2.17)	<0.001	1.30 (0.90–1.88)	0.169	2.91 (2.28–3.72)	<0.001
Problems with sleep	1.27 (1.13–1.44)	<0.001	1.27 (1.12–1.45)	<0.001	0.91 (0.67–1.24)	0.568	1.53 (1.26–1.86)	<0.001
Memory problems	2.11 (1.82–2.46)	<0.001	2.08 (1.78–2.44)	<0.001	1.60 (1.12–2.30)	0.011	1.69 (1.33–2.15)	<0.001
Difficulty concentrating	1.94 (1.67–2.25)	<0.001	1.91 (1.63–2.23)	<0.001	1.46 (1.03–2.07)	0.032	1.59 (1.26–2.00)	<0.001
Muscle or joint pain	1.38 (1.21–1.57)	<0.001	1.37 (1.20–1.57)	<0.001	1.16 (0.84–1.61)	0.357	1.24 (1.01–1.53)	0.038
Problems with sense of smell / taste	9.35 (7.51–11.63)	<0.001	8.23 (6.54–10.37)	<0.001	19.74 (10.53–37.00)	<0.001	1.30 (0.72–2.33)	0.382
Diarrhoea	1.75 (1.42–2.15)	<0.001	1.61 (1.29–2.01)	<0.001	1.27 (0.82–1.97)	0.284	2.12 (1.59–2.84)	<0.001
Stomach (abdominal pains)	1.72 (1.42–2.08)	<0.001	1.72 (1.41–2.09)	<0.001	0.96 (0.63–1.45)	0.835	2.14 (1.64–2.80)	<0.001
Changes to voice	2.68 (2.05–3.50)	<0.001	2.46 (1.85–3.26)	<0.001	1.34 (0.77–2.31)	0.296	3.26 (2.25–4.72)	<0.001
Hair loss	2.09 (1.70–2.57)	<0.001	2.10 (1.69–2.61)	<0.001	2.07 (1.20–3.60)	0.009	1.08 (0.73–1.58)	0.712
Unusual racing of the heart	2.23 (1.84–2.69)	<0.001	2.14 (1.75–2.62)	<0.001	1.51 (0.98–2.33)	0.064	1.81 (1.35–2.43)	<0.001
Lightheadedness or dizziness	2.17 (1.85–2.56)	<0.001	2.12 (1.79–2.52)	<0.001	1.74 (1.18–2.56)	0.005	1.59 (1.22–2.08)	<0.001
Unusual sweating	2.46 (2.02–3.00)	<0.001	2.42 (1.96–2.98)	<0.001	1.45 (0.91–2.31)	0.115	1.92 (1.40–2.63)	<0.001
MRC Dyspnoea	1.29 (1.12–1.47)	<0.001	1.18 (1.02–1.37)	0.027	1.47 (1.07–2.03)	0.017	1.40 (1.13–1.72)	0.002
PHQ-4 grade	1.29 (1.12–1.48)	<0.001	1.24 (1.07–1.44)	0.004	0.97 (0.70–1.34)	0.851	1.34 (1.09–1.66)	0.006
FACIT-13 score (reversed) ¶	3.4% (2.5–4.3)	<0.001	2.9% (2.0–3.8)	<0.001	16.5% (–1.6–34.6)	0.074	3.6% (2.3–4.9)	<0.001
EQ-5D Activities	2.02 (1.69–2.43)	<0.001	1.90 (1.57–2.31)	<0.001	1.48 (0.99–2.21)	0.054	1.88 (1.42–2.48)	<0.001
EQ-5D Pain	1.40 (1.22–1.60)	<0.001	1.39 (1.21–1.61)	<0.001	0.95 (0.68–1.32)	0.752	1.39 (1.12–1.72)	0.003
EQ-5D Self-care (two groups)	1.15 (0.81–1.61)	0.433	1.24 (0.87–1.75)	0.236	0.97 (0.41–2.27)	0.936	0.97 (0.57–1.63)	0.896
EQ-5D Mobility (two groups)	1.07 (0.87–1.32)	0.495	1.02 (0.82–1.26)	0.884	1.05 (0.66–1.69)	0.826	1.21 (0.88–1.64)	0.238
EQ-5D VAS ¶	–1.9% (–2.9 to –1.0)	<0.001	–1.6% (–2.7 to –0.6)	0.001	–8.7% (–21.7 to 4.2)	0.187	–1.5% (–3.1 to 0.0)	0.047

Comparison between non-COVID-19 ARI and no infection is shown in the main paper. ARI=acute respiratory infection. PHQ=Patient Health Questionnaire. MRC=Medical Research Council. VAS=visual analogue scale. \*Estimates are odds ratios for binary and ordinal outcomes, and predicted percentage point changes for continuous outcomes. †After adjustment for multiple corrections, p<0.045 is the threshold for statistical significance. ‡After adjustment for multiple corrections, p<0.005 is the threshold for statistical significance. §After adjustment for multiple testing, p<0.036 is the threshold for statistical significance. || Ordinal outcome. ¶ Continuous outcome. Raw coefficients are available in the appendix (table S1).

**Table 2: Symptom associations among participants with previous SARS-CoV-2 infection vs no infection or non-COVID-19 ARI**

	SARS-CoV-2 infection >12 weeks prior vs 4–12 weeks prior	
	Estimate* (95% CI)	p value†
Coughing	0.44 (0.29–0.67)	<0.001
Problems with sleep	1.02 (0.72–1.44)	0.925
Memory problems	0.91 (0.61–1.36)	0.652
Difficulty concentrating	0.90 (0.60–1.34)	0.597
Muscle or joint pain	0.97 (0.68–1.40)	0.886
Problems with sense of smell / taste	0.41 (0.28–0.62)	<0.001
Diarrhoea	0.56 (0.34–0.92)	0.021
Stomach (abdominal pains)	1.05 (0.62–1.78)	0.851
Changes to voice	0.50 (0.27–0.93)	0.029
Hair loss	1.03 (0.58–1.83)	0.919
Unusual racing of the heart	0.74 (0.45–1.19)	0.215
Lightheadedness or dizziness	0.83 (0.54–1.27)	0.396
Unusual sweating	0.82 (0.49–1.37)	0.452
MRC Dyspnoea‡	0.57 (0.40–0.82)	0.002
PHQ-4 grade‡	0.81 (0.56–1.17)	0.263
FACIT-13 score (reversed)§	-4.1% (-7.2 to -1.0)	0.009
EQ-5D Activities‡	0.72 (0.46–1.13)	0.157
EQ-5D Pain‡	1.03 (0.71–1.50)	0.867
EQ-5D Self-care (two groups)	3.13 (0.72–13.69)	0.130
EQ-5D Mobility (two groups)	0.72 (0.42–1.23)	0.226
EQ-5D VAS§	1.9% (-1.0 to 4.9)	0.192

PHQ=Patient Health Questionnaire. MRC=Medical Research Council. VAS=visual analogue scale. \*Estimates are odds ratios for binary and ordinal outcomes, and predicted percentage point changes for continuous outcomes. †After adjustment for multiple corrections,  $p < 0.007$  is the threshold for statistical significance. ‡Ordinal outcome. §Continuous outcome. Raw coefficients are available in the appendix (table S1).

**Table 3: Symptom associations by time since infection, among participants with previous SARS-CoV-2 infection**