

Cohort profile: OpenPROMPT

Alasdair D Henderson¹, Oliver Carlile¹, Iain Dillingham², Ben FC Butler-Cole², Keith Tomlin¹, Mark Jit¹, Laurie A Tomlinson¹, Michael Marks^{1, 4, 5}, Andrew Briggs¹, Liang-Yu Lin¹, Chris Bates³, John Parry³, Sebastian CJ Bacon², Ben Goldacre², Amir Mehrkar², The OpenSAFELY Collaborative, Emily Herrett¹⁺, Rosalind M Eggo^{1+*}

¹ London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

² Bennett Institute for Applied Data Science, Nuffield Department of Primary Care Health Sciences, University of Oxford, OX2 6GG, UK

³ TPP, TPP House, 129 Low Lane, Horsforth, Leeds, LS18 5PX, UK

⁴ Hospital for Tropical Diseases, University College London Hospital, London, UK, WC1E 6JD

⁵ Division of Infection and Immunity, University College London, London, London WC1E 6BT

+ Contributed equally

Word count: 3,395

Key features

- OpenPROMPT is a cohort of individuals with longitudinal patient reported questionnaire data and linked to routinely collected health data from primary and secondary care.
- With the approval of NHS England we collected responses from 7,574 individuals, with detailed questionnaire responses from 6,337 individuals who responded using a smartphone app.
- Data were collected from each participant over 90 days at 30-day intervals using questionnaires to ask about HRQoL, productivity and symptoms of long COVID.
- Responses from the majority of OpenPROMPT (6,006; 79.3%) were linked to participants' existing health records from primary care, secondary care, COVID-19 testing and vaccination data.
- OpenPROMPT can currently be used to estimate the impact of long COVID on HRQoL, and because of the linkage within OpenSAFELY, the data from OpenPROMPT can be used to enrich routinely collected records in further research by approved researchers on behalf of NHS England.

Abstract

OpenPROMPT is a cohort of individuals with longitudinal patient reported questionnaire data and linked to routinely collected health data from primary and secondary care. Data were collected between November 2022 and October 2023 in England. OpenPROMPT was designed to measure the impact of long COVID on health-related quality-of-life (HRQoL). With the approval of NHS England we collected responses from 7,574 individuals, with detailed questionnaire responses from 6,337 individuals who responded using a smartphone app. Data were collected from each participant over 90 days at 30-day intervals using questionnaires to ask about HRQoL, productivity and symptoms of long COVID. Responses from the majority of OpenPROMPT (6,006; 79.3%) were linked to participants' existing health records from primary care, secondary care, COVID-19 testing and vaccination data. Analysis takes place using the OpenSAFELY data analysis platform which provides a secure software interface allowing the analysis of pseudonymized primary care patient records from England. OpenPROMPT can currently be used to estimate the impact of long COVID on HRQoL, and because of the linkage within OpenSAFELY, the data from OpenPROMPT can be used to enrich routinely collected records in further research by approved researchers on behalf of NHS England.

Lay summary

OpenPROMPT is a study which used a phone app to conduct a longitudinal survey aimed at measuring the health related quality of life of people living with long COVID. The study

recruited participants between November 2022 and July 2023 and followed them up for 90 days. The key advantage of this study is that the responses are linked to the individual's personal health records, so we have access to much more data than the questionnaire responses alone.

Here, we summarised who has used the app, how much data has been collected and the quality of the data. We also provide details to document how and why the data were collected so that the data can be used by other researchers in the future. This will maximise the benefit of this study, and ensure that the time invested by participants is put to best use.

In this study we aimed to provide lots of important information about how many people are involved, how much information we have about them, their age, where they live, and how healthy they are. Finally, for certain variables we compared the responses people recorded in the app with what is kept on their electronic record to see if they agree or disagree.

Why was the cohort set up?

OpenPROMPT is a longitudinal prospective cohort of participant-recorded healthcare outcomes among UK adults, collected during the COVID-19 pandemic (November 2022 - July 2023). OpenPROMPT was funded by the National Institute for Healthcare research (NIHR) to better understand the impact of long COVID on health-related quality-of-life (HRQoL). Long COVID, here defined as symptoms consistent with SARS-COV-2 that persist for 12 weeks following an acute infection (1,2), has emerged as one of the most costly impacts of the pandemic in the UK (3) and globally (4,5). In the UK, the Office for National Statistics regularly collected information on the number and impact of persistent Covid-19 symptoms and as of November 2022 there were an estimated 2.1 million people with long COVID, 73% of whom reported that the symptoms adversely affected day-to-day activities (6).

Electronic health records (EHRs) are a valuable research tool to better understand the impact of health conditions. They have contributed to better understanding of COVID-19 fatalities (7,8), infection risk (9,10), the effectiveness of treatments (11,12) and vaccinations (13–15). However, long COVID in EHRs is poorly captured (16), and even in those that have a record this binary categorisation does not reflect the variety of experiences in those with long COVID, which may be better captured by patient-reported outcomes. Studies have previously used EHRs to study long COVID (16–19), and others have collected patient-reported outcome measurements (20–24), however OpenPROMPT is the first study of its kind to merge these two data sources in a trusted research environment (25).

The gold standard metric to quantify the impact of a disease is the quality adjusted life year (QALY), which takes into account impacts on morbidity as well as mortality. QALYs are used in National Health Service (NHS) resource planning to make healthcare decisions, evaluate and prioritise interventions, such as vaccination boosters for COVID-19. We therefore established OpenPROMPT to measure long COVID impacts using (1) standardised measures of changes in patient HRQoL (including physical and mental health effects) using the EuroQoL EQ-5D instrument (26) and (2) formal quantification of the healthcare burden of the condition and its treatment. We embedded patient and public involvement throughout OpenPROMPT to ensure that patient input on the effects of long COVID were included, and that possible concerns about data security were respected.

Who is in the cohort

Study design and setting

The protocol for data collection for OpenPROMPT was pre-specified and details have been previously published (27). Briefly, OpenPROMPT is a cohort study among adults in the UK, which recruited between November 2022 and July 2023. This study was conducted in collaboration with TPP, a general practitioner (GP) software provider. TPP's software system is called SystemOne and holds patient electronic health records for approximately 40% of the population of England. Their in-house smartphone app, *Airmid* (28), has been developed as a patient-facing mobile application, and links directly to SystemOne records. It was created to allow appointment booking, for patients to access their own medical records, and for patients

to input their own data to feed into their medical record and personal health record. We used Airmid's Research Module to host bespoke questionnaires to assess quality of life in respondents. Any individual can download the Airmid app, even if their GP does not use TPP software, and access to Airmid and SystemOne was provided by TPP for no charge.

Inclusion and exclusion criteria

Any adult with an NHS number was eligible to self-enrol in the study, irrespective of history of COVID-19 infection. There was an implicit exclusion criteria that participants needed to own a smartphone and be able to download and use the Airmid app. All responses to OpenPROMPT questionnaires are available for analysis, however access to linked primary care records is only available from those registered with a TPP practice.

Linked healthcare records

OpenPROMPT responses are stored within the EHR vendor's highly secure data centre and can be securely accessed through OpenSAFELY, a data analytics platform created by our team on behalf of NHS England to address urgent COVID-19 research questions (<https://opensafely.org>). If the respondent is registered with a TPP practice then their responses can be linked to other digital healthcare records within OpenSAFELY, including hospitalisations, prescriptions, GP appointments, diagnoses and COVID-19 test results (**Figure 1**).

OpenSAFELY provides a secure software interface allowing the analysis of pseudonymized primary care patient records from England, avoiding the need for large volumes of potentially disclosive pseudonymized patient data to be transferred off-site. This, in addition to other technical and organisational controls, minimises any risk of re-identification. Similarly pseudonymized datasets from other data providers are securely provided to the EHR vendor and linked to the primary care data. The dataset analysed within OpenSAFELY is based on 24 million people currently registered with GP surgeries using TPP SystemOne software. It includes pseudonymized data such as coded diagnoses, medications and physiological parameters. No free text data are included. Further details on our information governance can be found under "[Information governance](#)".

To ensure protection of anonymity and in line with OpenSAFELY practice, we have censored any counts less than seven and have rounded all frequencies to the nearest 7 in this article. All code used for this report is freely available online at <https://github.com/opensafely/openprompt-cohort-profile>.

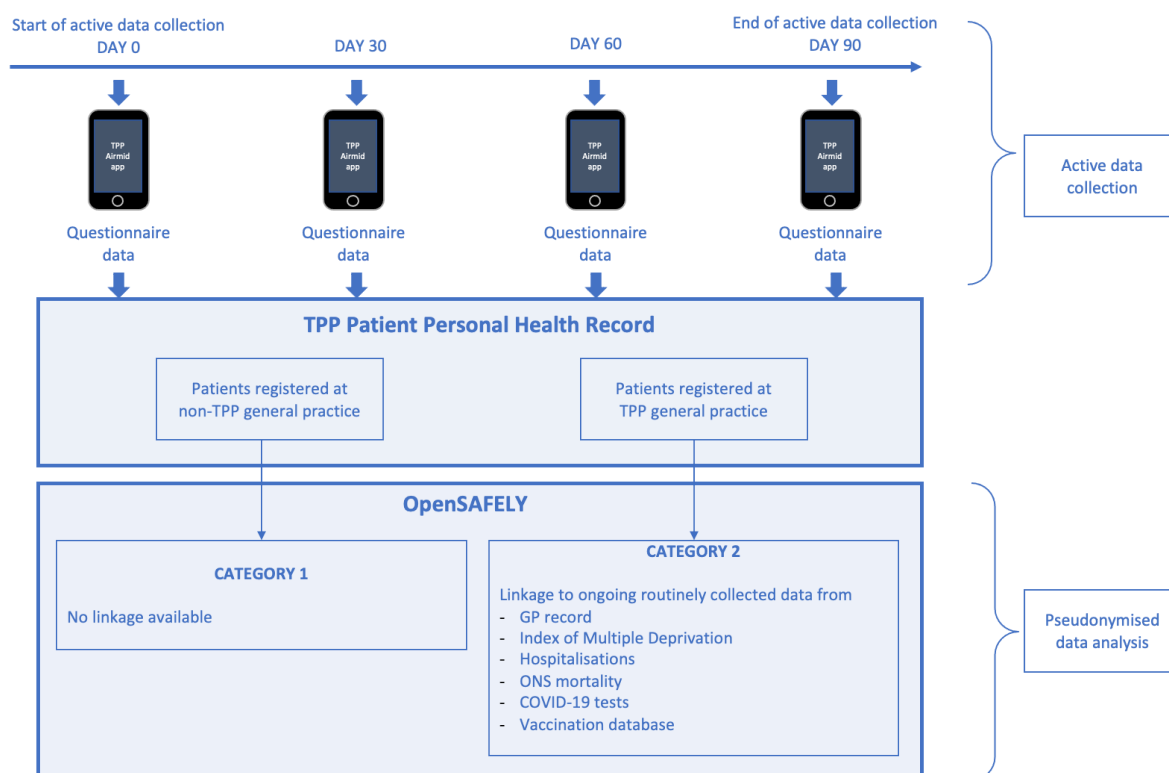


Figure 1: *OpenPROMPT study design*

OpenPROMPT cohort

The OpenPROMPT cohort comprises 7,574 individuals who have consented to data collection and responded to the baseline questionnaire, 6,337 of which proceeded to respond to our research questionnaire. We enrolled participants between November 2022 and July 2023 and each respondent was followed up for 90 days from baseline (data collection up to October 2023). We used the Airmid app to send notification reminders to respondents to complete the research questionnaire every 30 days. In total, 299,240 responses to our research questionnaire were collected over the study period (**Figure S1**).

Overall 4,599/7,574 (60.7%) of the baseline OpenPROMPT cohort identified as female and 7,203/7,574 (95.1%) reported their ethnicity as white. We collected information on the highest education level, relationship status, and household income at baseline. Amongst these categories approximately half of the cohort had attended college or university (4,004/7,574; 52.9%), were married or in a civil partnership (4,172/7,574; 55.1%), and had a household income between £26,000 and £63,999 (3,052/7,574; 40.2%) (**Figure 2B & Table 1**).

We did not collect information on age in the app, but date of birth information is available in GP records linked with OpenSAFELY. Amongst the population with this linkage available (6,006; 79.3%) the median age was 51 (interquartile range: 41-61). We were able to analyse a wider range of demographic and clinical information in this subset with OpenSAFELY linkage available (**Figure 2C-D**). The largest representation of respondents was from the East of England (1,407; 23.5%) and East Midlands (1,217; 20.3%), but linked data was only available for respondents from England (**Figure 2C**). The majority of respondents were from

urban areas (4,424/6,006; 73.7%) compared to rural areas 1,288 (21.4%) and information are available on household composition (collected in 2020) which showed that most OpenPROMPT respondents lived in household size of at least 2, although 903/6,006 (15%) respondents were living alone as of 2020.

We were also able to characterise the clinical status of participants with linked OpenSAFELY by investigating the clinical diagnoses in linked GP records. We counted the presence of previous morbidity codes for a range of common chronic conditions as in previous OpenSAFELY research including asthma, cancer, diabetes, cardiovascular disease and rheumatoid disease and other immunosuppressive conditions (full details in **Supplementary Methods**). We found that the majority did not have a comorbidity (3,409/6,006; 56.8%), but 2,009 (33.4%) had one recorded comorbidity and 588 (9.8%) had at least two comorbidities at baseline. Finally, we analysed linked records of the index of multiple deprivation and found that the most common quintile were the least deprived (1,358/6,006; 22.6%) and second least deprived (1,197; 19.9%) quintile, consistent with the reported household income information collected in the app (**Figure 2D**).

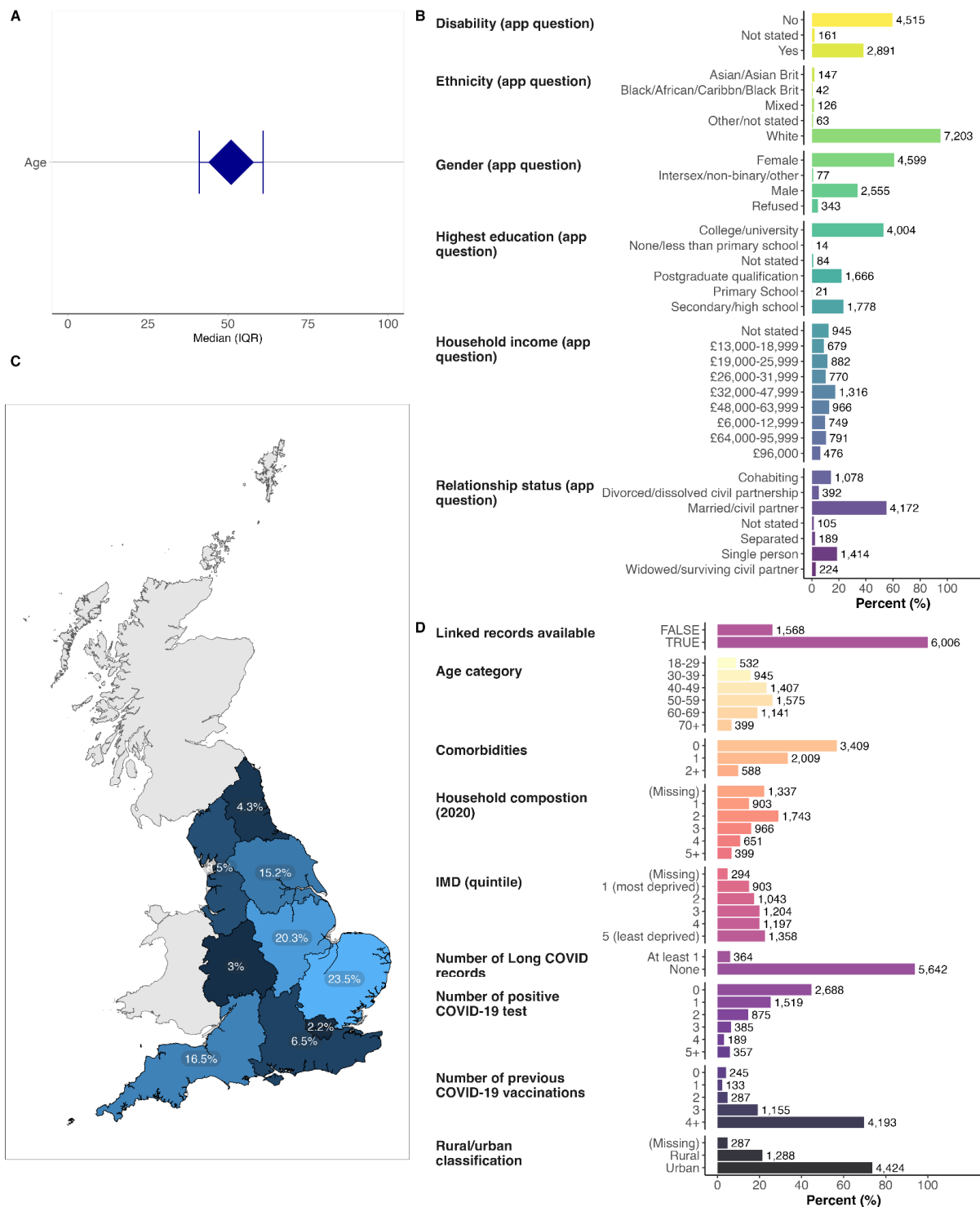


Figure 2: Baseline characteristics of the OpenPROMPT cohort. (A) median age and interquartile range for OpenPROMPT respondents with linked TPP data available (n=6,006). (B) Responses to the baseline OpenPROMPT questionnaire (n=7,574). (C) geographic distribution of respondents with linked TPP data available. It was possible for people to contribute to OpenPROMPT outside of England but linked data is available for English residents only. (D) further clinical and demographic data for the OpenPROMPT cohort with linked TPP data available (n=6,006)

The advantage of linking data within OpenSAFELY is that we can directly compare the baseline characteristics of the OpenPROMPT cohort with the wider OpenSAFELY population of 24 million patients registered at TPP practices. We found that our cohort had greater representation of people aged between 40-70 but that there were fewer participants in younger and older age groups (**Figure 3A**). The OpenPROMPT cohort has a greater representation of individuals from less deprived areas, more female participants and more white participants than the overall OpenSAFELY population.

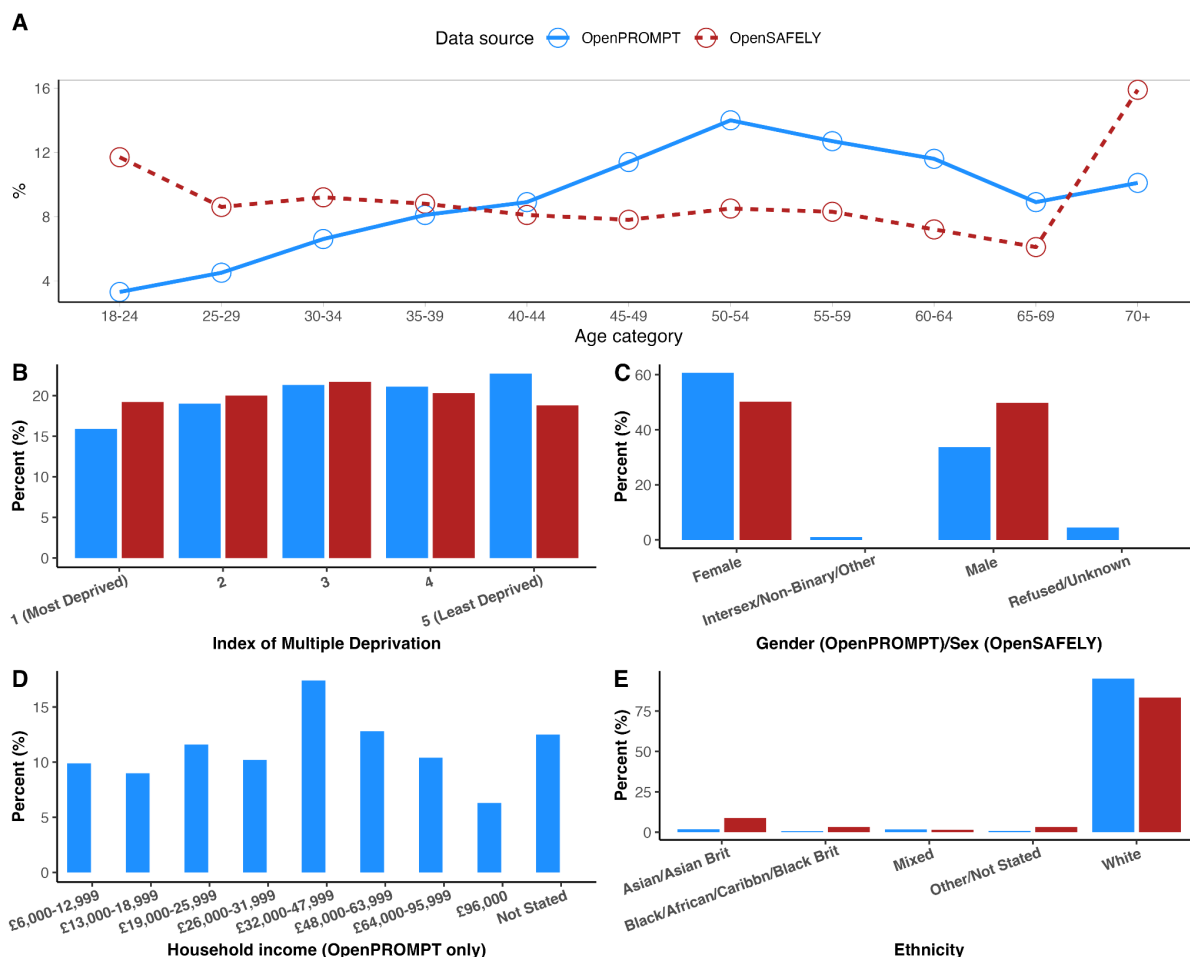


Figure 3: Comparison of OpenPROMPT cohort (blue) with the currently registered OpenSAFELY-TPP (red) population as of November 2022. OpenPROMPT data shown for those with a valid baseline response ($n=7,574$). In panel D, household income data is not collected by GPs so is not available in OpenSAFELY

How often were they followed up and for how long?

The OpenPROMPT study was designed to collect information at baseline from respondents, and then every 30 days for a total of 90 days of follow up. Most OpenPROMPT participants were recruited in early 2023 and the final recruitment was in July 2023 (**Figure 4A**). Over the full 90 days of follow up, we collected a large number of responses to our baseline and research questionnaires, which were mostly from the first round of data collection (**Figure 4B**). We found that over half the cohort did not complete a follow up questionnaire at 30 days, and by 90 days we collected responses from 884 individuals (13.9% of original 6,337 respondents) (**Figure 4C**). It was possible for respondents to input data using the app at any

point during follow up, so to retain consistency with our study protocol, we only included responses that were within a ± 5 day window of the intended 30, 60, 90 days after the baseline questionnaire was completed, which means that some responses were discarded (**Figure 4D**).

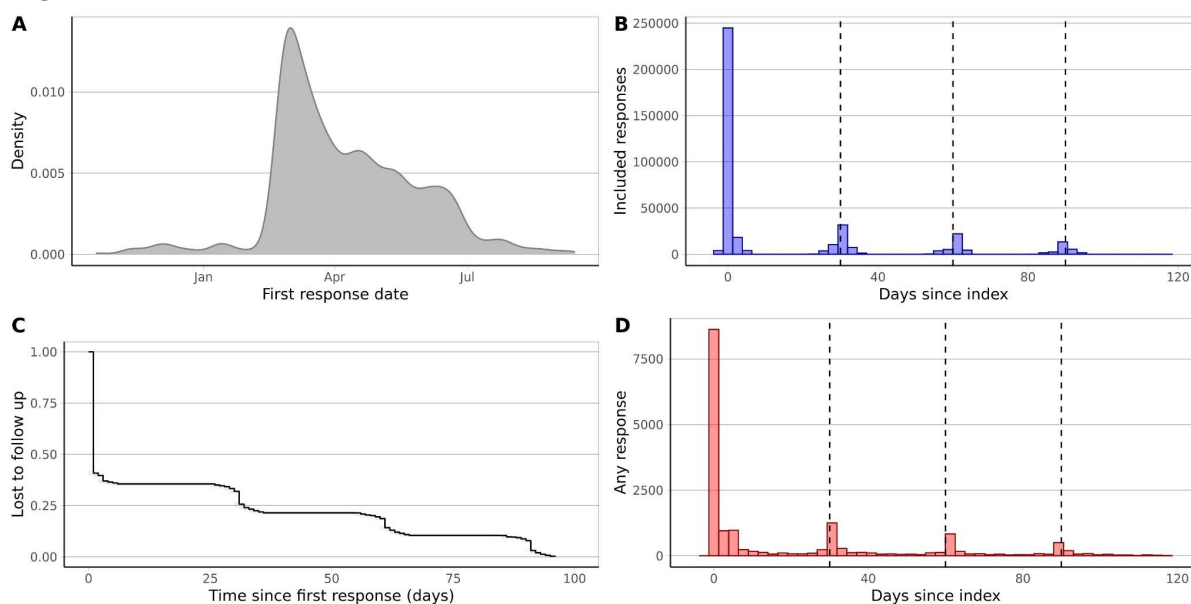


Figure 4: OpenPROMPT data collected over time. (A) density plot of the date of completion of the baseline questionnaire between November 2022 and July 2023. (B) Histogram of the number of valid responses to a question within OpenPROMPT over time. Only responses within a ± 5 day window of the intended 30, 60, 90 days after the baseline questionnaire were included. (C) reverse Kaplan-Meier plot showing loss to follow up over time. The date of the last available data is shown along the x-axis. Time steps are shown once a total of 7 events occur to protect anonymity). (D) Histogram of the number of individuals that submitted a response over time. Responses not within the ± 5 day window are excluded

What has been measured?

We collected detailed data on the quality of life of the OpenPROMPT cohort including people living with and without long COVID. We recorded information from three standardised questionnaires: the Functional Assessment of Chronic Illness Therapy – Fatigue Scale (FACIT-Fatigue) (29), EQ-5D-5L (26), and the MRC breathlessness scale. We also asked questions about historic COVID-19 infections, the number and duration of symptoms, the vaccination status of respondents and whether or not they have recovered from their most recent COVID-19 infection.

Data collected with these questionnaires show a range of health related quality of life in the OpenPROMPT cohort. There was a notable proportion of people who reported severe fatigue using the FACIT questionnaire, especially for being “too tired to eat”, “feel weak” and “need help for usual activities” (**Figure 5A**). Similarly, many respondents had severe problems with anxiety/depression and pain/discomfort in the EQ-5D questionnaire (**Figure 5B**), and the most common MRC breathlessness grade was grade 2 (**Figure C**).

We asked respondents about their experience with COVID-19, recovery and vaccinations. The vast majority of participants had a previous COVID-19 infection, mostly confirmed with a

positive test (**Figure 5D**), and the most common duration of the longest COVID-19 infection was over 12 weeks. In the first survey round, only 1,547/5,390 (28.7%, the denominator is smaller because the question was optional) reported that the symptoms of their initial infection lasted less than 2 weeks (**Figure 5E**). Over 97% of the OpenPROMPT cohort had received a SARS-COV-2 vaccination dose at baseline (**Figure 5F**). In total, 6,337 people responded to the question “How many previous COVID-19 infections have you had?” during the first study round. 903 (14.3%) had not had a previous COVID-19 infection, almost half (2,842; 44.9%) had one previous infection and 784 had 3 or more infections previously (12.3%) (**Figure 5G**).

To assess the experience of long COVID in OpenPROMPT we collected relevant data from two questions. One asked about the duration of the longest infection with COVID-19 as described above. We also asked a more direct question, “have you recovered from your most recent COVID-19 infection?”. Across the study period we found that approximately half of respondents had not recovered from a COVID-19 infection. 2,653 of 5,488 responses at baseline reporting persistent symptoms (48.3%), and this proportion slightly declined over the 90 day study period to 46.8% at 30 days, 42.4% at 60 days and 40.8% at 90 days after baseline (**Figure 5H**).

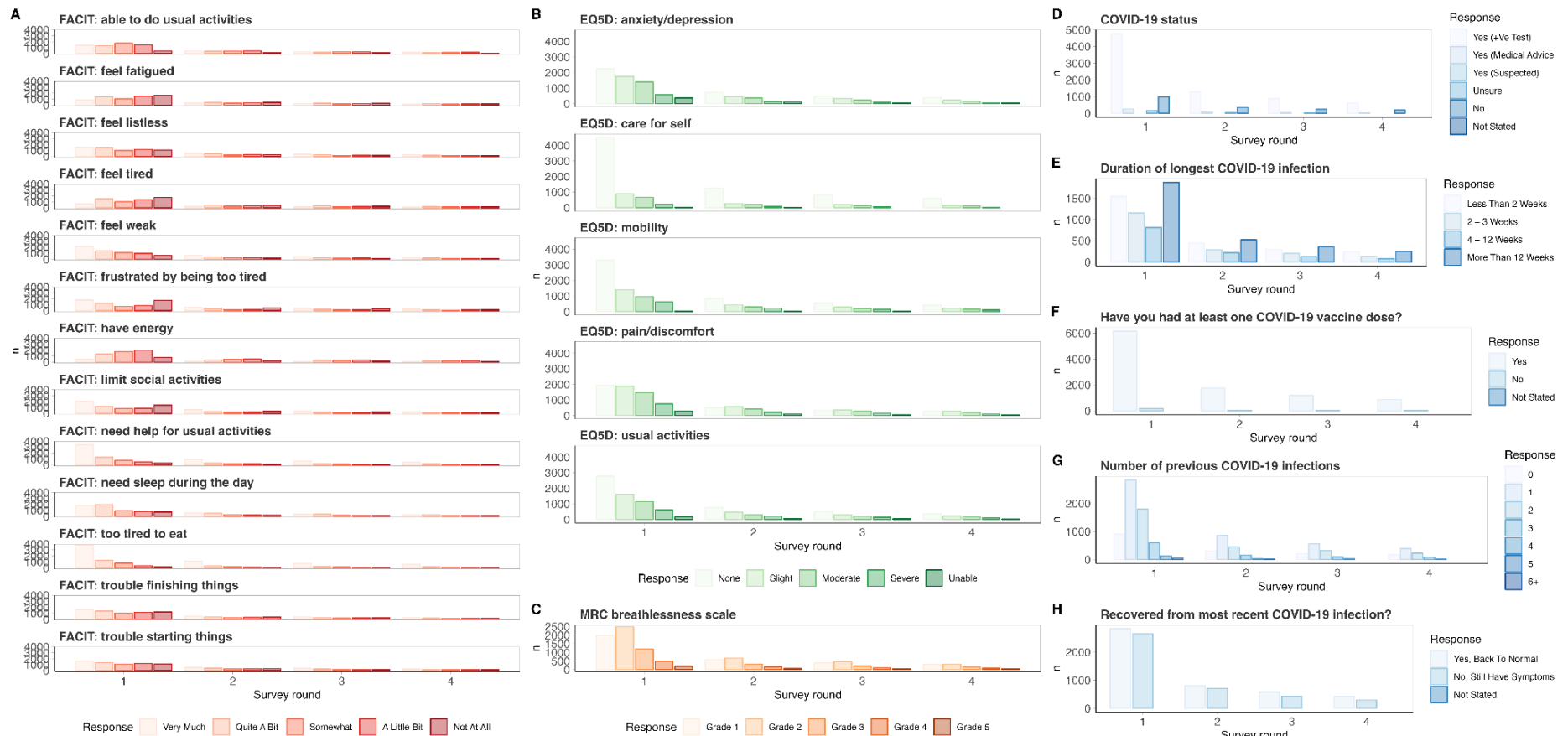
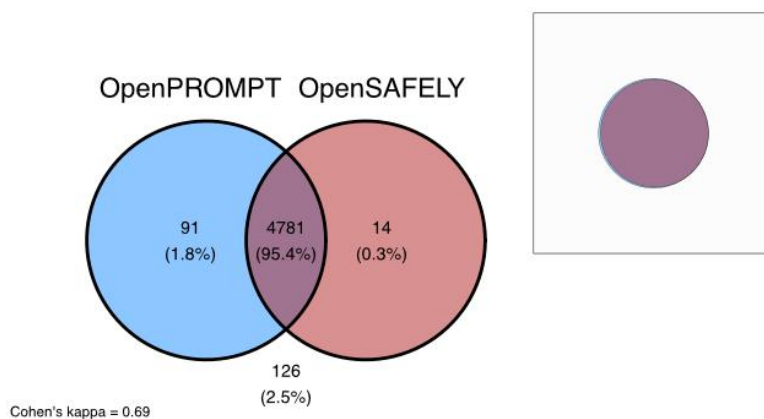


Figure 5: Questionnaire response in OpenPROPMT over 90 days from baseline. (A) Functional Assessment of Chronic Illness Therapy – Fatigue Scale (FACIT-Fatigue), (B) EQ-5D-5L, (C) MRC breathlessness scale, (D-H) responses to specific questions about COVID-19 infection history, duration of infection, SARS-COV-2 vaccinations and whether respondents had recovered from their most recent infection

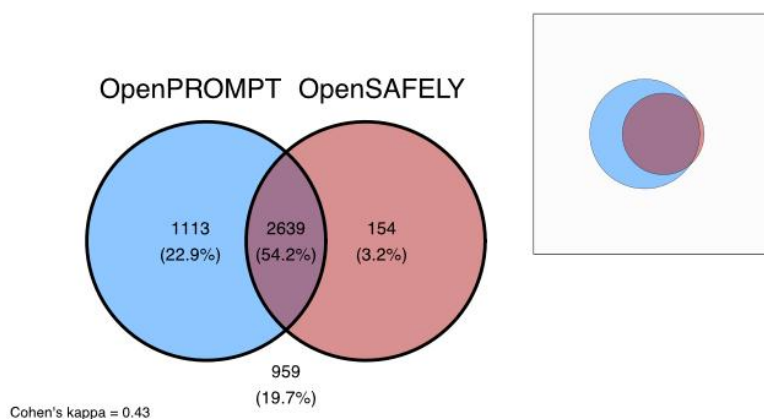
We have demonstrated the advantage of collecting patient reported outcome measures linked to wider healthcare datasets above through direct comparison of the representativeness of our cohort. We can also compare the capture of healthcare outcomes in the two sources of data, since some questions in OpenPROMPT should also be collected in routine electronic health records (EHRs). We therefore compared the assessment of the same variables from different data sources in the same individuals with linked data available. We identified three questions regarding COVID-19 that have comparable data within OpenSAFELY. Firstly, the SARS-COV-2 vaccination status of respondents (“Yes” in **Figure 5F**) which we compared to any SARS-COV-2 vaccination records held within OpenSAFELY. Secondly, whether respondents had a previous COVID-19 infection (“Yes (+ve test)” in **Figure 5D**), which we compared to any positive result in the COVID-19 Second Generation Surveillance System (SGSS). Finally, we compared the proportion of people who were still experiencing symptoms in OpenPROMPT (“No, still have symptoms” in **Figure 5H**) with those with a GP record of long COVID, using a previously validated codelist for long COVID (16).

We found excellent agreement between OpenPROMPT and OpenSAFELY with respect to previous vaccination, demonstrating that the majority of vaccinations are well recorded in OpenSAFELY, and that information from this question is accurately captured in OpenPROMPT (**Figure 6A**). We also found very few examples (154; 3.2%) of people who had a record of previous COVID-19 infection in OpenSAFELY and did not report this in OpenPROMPT. However, a larger proportion of respondents did report a previous COVID-19 infection confirmed with a positive test in OpenPROMPT but had no record in OpenSAFELY (1,113; 22.9%) suggesting under-reporting of infections captured in SGSS, for example when lateral flow test results were not submitted online (**Figure 6B**). The worst agreement (Cohen’s Kappa = 0.15) was for recording of long COVID (**Figure 6C**). In OpenPROMPT 2,079 people reported persistent symptoms at baseline, and had linked data available. However, only 315 (15.1%) also had a GP record for long COVID (7.3% of the total linked data sample).

A: vaccinated (yes/no)



B: COVID-19 infection



C: long COVID

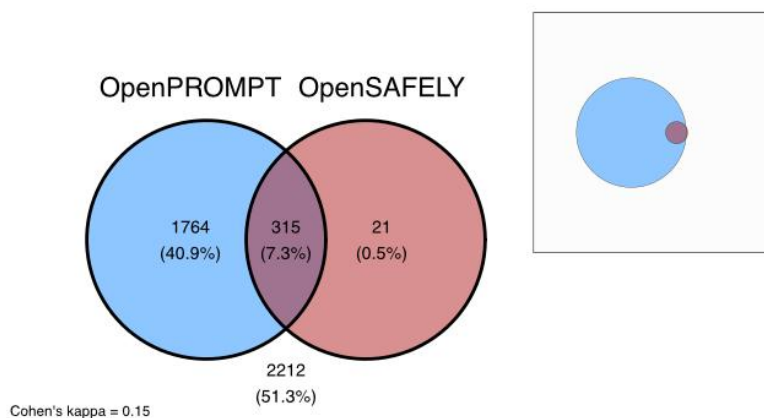


Figure 6: Venn diagrams demonstrating agreement between patient reported (OpenPROMPT, blue) and routinely collected (OpenSAFELY, red) data for the same clinical information in individuals with linked data. Frequency and percentage of total are shown for each variable, the inset shows these proportions on a relative scale. (A) previous vaccination, (B) previous Covid-19 infection, (C) unresolved persistent symptoms following Covid-19 infection

How can the data be accessed?

The OpenPROMPT questionnaire data provided by participants will be made available to allow other researchers to benefit from this work in the future, and for a range of different studies and purposes. Data access will be free, please contact openprompt@lshtm.ac.uk. If accessing through OpenSAFELY then access will be subject to completion of data security requirements specified by the OpenSAFELY platform at the time of application. All other study documentation will be stored for a minimum of 5 years after the follow-up period.

What are the main strengths and weaknesses?

The main strength of OpenPROMPT is that we have collected data on HRQoL using validated instruments, especially the EQ-5D, in people with and without long COVID, and linked these data to a large EHR resource. This facilitates a much wider range of research including information on healthcare history, prescriptions, COVID-19 vaccinations and other healthcare data, without participants having to answer excessive questionnaires, which is especially pertinent in long COVID research.

Linkage between OpenPROMPT and OpenSAFELY is also valuable from an epidemiological perspective as it facilitates direct data comparisons between two separate sources. This means that we can investigate which routinely collected variables are well recorded, for example Covid-19 vaccinations, and which ones are not, such as recovery from long COVID symptoms.

Finally, We embedded patient and public involvement (PPI) throughout the development and delivery of OpenPROMPT through three mechanisms. Firstly, OpenPROMPT recruited a three person steering group, including members with long COVID which met with the research team online every 6 months. Secondly, we held two separate PPI events in January and September 2023 to obtain feedback on the study and respond in real time. Finally, engagement with research using electronic health records was also being undertaken as part of OpenSAFELY. This includes formation of an Oversight Board (which reviewed OpenPROMPT) (<https://www.opensafely.org/governance/>), publishing blogs explaining OpenSAFELY (<https://www.opensafely.org/blog/>), presenting results of studies at public and professional events, soliciting feedback on studies and conducting focus groups.

There are limitations in representation of our OpenPROMPT sample with regards to certain demographic variables. OpenPROMPT includes an over-representation of women, those of white ethnicity and less deprived areas as measured by IMD compared to the average in the wider OpenSAFELY population. This, in part, reflects the overrepresentation of people with long COVID in OpenPROMPT, which has been reported higher in these same groups, however this does need to be considered when analysing OpenPROMPT data. In particular, our study was susceptible to digital exclusion as users required a smartphone and the ability to use it confidently in order to join the study. This digital exclusion will disproportionately impact certain groups in the population from which the sample has been drawn. The generalisability of findings from OpenPROMPT data to the wider UK population are therefore uncertain, for example we are unable to estimate the prevalence of long COVID in the UK from these data.

We also experienced a relatively high loss to follow up during the study. We aimed to reduce the burden of completing these questionnaires, especially since our study focussed on people living with a chronic condition which impacts fatigue. We made the questionnaires available in an app and it was possible to save progress and return at a later time. However, the high rate of loss to follow up does suggest that there was a selection pressure from our data collection method and this should be recognised in analysis of the longitudinal OpenPROMPT data. There are likely to be systematic differences between those that do and do not complete follow up and so future studies should include plans to improve retention over the planned study period, since the convenience of a smartphone app was insufficient in our experience.

Administrative

Acknowledgments

We are very grateful for all the support received from the TPP Technical Operations team throughout this work, and for generous assistance from the information governance and database teams at NHS England and the NHS England Transformation Directorate.

Information governance

LSHTM is the data controller of OpenPROMPT data. NHS England is the data controller of the NHS England OpenSAFELY COVID-19 Service; TPP is the data processor; all study authors using OpenSAFELY have the approval of NHS England (30). This implementation of OpenSAFELY is hosted within the TPP environment which is accredited to the ISO 27001 information security standard and is NHS IG Toolkit compliant (31); Patient data has been pseudonymised for analysis and linkage using industry standard cryptographic hashing techniques; all pseudonymised datasets transmitted for linkage onto OpenSAFELY are encrypted; access to the NHS England OpenSAFELY COVID-19 service is via a virtual private network (VPN) connection; the researchers hold contracts with NHS England and only access the platform to initiate database queries and statistical models; all database activity is logged; only aggregate statistical outputs leave the platform environment following best practice for anonymisation of results such as statistical disclosure control for low cell counts (32).

The service adheres to the obligations of the UK General Data Protection Regulation (UK GDPR) and the Data Protection Act 2018. The service previously operated under notices initially issued in February 2020 by the the Secretary of State under Regulation 3(4) of the Health Service (Control of Patient Information) Regulations 2002 (COPI Regulations), which required organisations to process confidential patient information for COVID-19 purposes; this set aside the requirement for patient consent (33). As of 1 July 2023, the Secretary of State has requested that NHS England continue to operate the Service under the COVID-19 Directions 2020 (34). In some cases of data sharing, the common law duty of confidence is met using, for example, patient consent or support from the Health Research Authority Confidentiality Advisory Group (35).

Taken together, these provide the legal bases to link patient datasets using the service. GP practices, which provide access to the primary care data, are required to share relevant health information to support the public health response to the pandemic, and have been informed of how the service operates.

Ethics

This research is part of the OpenPROMPT study “Quality-of-life in patients with long COVID: harnessing the scale of big data to quantify the health and economic costs” which has ethical approval from HRA and Health and Care Research Wales (HCRW) (IRAS project ID 304354).

Ethical review

The Study Coordination Centre has obtained approval from the LSHTM Research Ethics Committee (ref 28030), as well as a favourable opinion from the South Central—Berkshire B Research Ethics Committee (ref 22/SC/0198).

Dissemination plan

All publications and presentations relating to the study will be authorised by the study management group and also follow the OpenSAFELY publication policy. All publications resulting from the OpenPROMPT study will be open access.

Data sharing

The questionnaire data provided by participants will be made available to allow other researchers to benefit from this work in the future, and for a range of different studies and purposes. Opting-in to the study required an affirmative response to agree to use of data in this way. After completion of the study, a pseudonymised copy of the data will be held according to NHS England retention policy. Data access will be governed by LSHTM and will require researchers to complete a data access form. All other study documentation will be stored for a minimum of 5 years after the follow-up period.

Funding

This research was supported by the National Institute for Health and Care Research (NIHR) (OpenPROMPT: COV-LT2-0073)). The OpenSAFELY Platform is supported by grants from the Wellcome Trust (222097/Z/20/Z) and MRC (MR/V015737/1, MC_PC_20059, MR/W016729/1). In addition, development of OpenSAFELY has been funded by the Longitudinal Health and Wellbeing strand of the National Core Studies programme (MC_PC_20030: MC_PC_20059), the NIHR funded CONVALESCENCE programme (COV-LT-0009), NIHR (NIHR135559, COV-LT2-0073), and the Data and Connectivity National Core Study, led by Health Data Research UK in partnership with the Office for National Statistics and funded by UK Research and Innovation (grant ref MC_PC_20058) and Health Data Research UK (HDRUK2021.000).

The views expressed are those of the authors and not necessarily those of the NIHR, NHS England, UK Health Security Agency (UKHSA) or the Department of Health and Social Care. Funders had no role in the study design, collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

References

1. Sharp S, Boyce S, Karpusheff J, Glen F. Living with Long COVID: Implementing a living approach to the NICE guideline on managing the long-term effects of COVID-19 [Internet]. medRxiv; 2023 [cited 2023 Aug 18]. p. 2023.05.09.23289572. Available from: <https://www.medrxiv.org/content/10.1101/2023.05.09.23289572v1>
2. Overview | COVID-19 rapid guideline: managing the long-term effects of COVID-19 | Guidance | NICE [Internet]. NICE; 2020 [cited 2023 Nov 27]. Available from: <https://www.nice.org.uk/guidance/ng188>
3. Altmann DM, Whettlock EM, Liu S, Arachchillage DJ, Boyton RJ. The immunology of long

- COVID. *Nat Rev Immunol*. 2023 Jul 11;1–17.
4. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. *Nat Rev Microbiol*. 2023 Jan 13;1–14.
 5. Bowe B, Xie Y, Al-Aly Z. Postacute sequelae of COVID-19 at 2 years. *Nat Med*. 2023 Aug 21;1–11.
 6. Prevalence of ongoing symptoms following coronavirus (COVID-19) infection in the UK - Office for National Statistics [Internet]. [cited 2023 Nov 27]. Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/prevalenceofongoingsymptomsfollowingcoronaviruscovid19infectionintheuk/3november2022>
 7. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature*. 2020;584(7821):430–6.
 8. Nab L, Parker EPK, Andrews CD, Hulme WJ, Fisher L, Morley J, et al. Changes in COVID-19-related mortality across key demographic and clinical subgroups in England from 2020 to 2022: a retrospective cohort study using the OpenSAFELY platform. *Lancet Public Health*. 2023 May 1;8(5):e364–77.
 9. Mathur R, Rentsch CT, Morton CE, Hulme WJ, Schultze A, MacKenna B, et al. Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the OpenSAFELY platform. *The Lancet*. 2021 May 8;397(10286):1711–24.
 10. Wing K, Grint DJ, Mathur R, Gibbs HP, Hickman G, Nightingale E, et al. Association between household composition and severe COVID-19 outcomes in older people by ethnicity: an observational cohort study using the OpenSAFELY platform. *Int J Epidemiol*. 2022 Dec 1;51(6):1745–60.
 11. Collaborative TO, Tazare J, Nab L, Zheng B, Hulme WJ, Green ACA, et al. Effectiveness of Sotrovimab and Molnupiravir in community settings in England across the Omicron BA.1 and BA.2 sublineages: emulated target trials using the OpenSAFELY platform [Internet]. medRxiv; 2023 [cited 2023 Nov 27]. p. 2023.05.12.23289914. Available from: <https://www.medrxiv.org/content/10.1101/2023.05.12.23289914v1>
 12. Zheng B, Green ACA, Tazare J, Curtis HJ, Fisher L, Nab L, et al. Comparative effectiveness of sotrovimab and molnupiravir for prevention of severe covid-19 outcomes in patients in the community: observational cohort study with the OpenSAFELY platform. *BMJ*. 2022 Nov 16;379:e071932.
 13. Parker EPK, Horne EMF, Hulme WJ, Tazare J, Zheng B, Carr EJ, et al. Comparative effectiveness of two- and three-dose COVID-19 vaccination schedules involving AZD1222 and BNT162b2 in people with kidney disease: a linked OpenSAFELY and UK Renal Registry cohort study. *Lancet Reg Health – Eur* [Internet]. 2023 May 3 [cited 2023 May 19];0(0). Available from: [https://www.thelancet.com/journals/lanep/article/PIIS2666-7762\(23\)00055-8/fulltext](https://www.thelancet.com/journals/lanep/article/PIIS2666-7762(23)00055-8/fulltext)
 14. Hulme WJ, Horne EMF, Parker EPK, Keogh RH, Williamson EJ, Walker V, et al. Comparative effectiveness of BNT162b2 versus mRNA-1273 covid-19 vaccine boosting in England: matched cohort study in OpenSAFELY-TTP. *BMJ*. 2023 Mar 15;380:e072808.
 15. Curtis HJ, Inglesby P, Morton CE, MacKenna B, Green A, Hulme W, et al. Trends and clinical characteristics of COVID-19 vaccine recipients: a federated analysis of 57.9 million patients' primary care records in situ using OpenSAFELY. *Br J Gen Pract*. 2022 Jan 1;72(714):e51–62.
 16. Walker AJ, MacKenna B, Inglesby P, Tomlinson L, Rentsch CT, Curtis HJ, 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 Nov;71(712):e806–14.
 17. Pfaff ER, Madlock-Brown C, Baratta JM, Bhatia A, Davis H, Girvin A, et al. Coding long COVID: characterizing a new disease through an ICD-10 lens. *BMC Med*. 2023 Feb 16;21(1):58.
 18. Pfaff ER, Girvin AT, Bennett TD, Bhatia A, Brooks IM, Deer RR, et al. Identifying who has long COVID in the USA: a machine learning approach using N3C data. *Lancet Digit*

- Health. 2022 Jul 1;4(7):e532–41.
19. Strasser ZH, Dagliati A, Abad ZSH, Klann JG, Wagholikar KB, Mesa R, et al. A retrospective cohort analysis leveraging augmented intelligence to characterize long COVID in the electronic health record: A precision medicine framework. *PLOS Digit Health*. 2023 Jul 25;2(7):e0000301.
 20. Lloyd-Evans PHI, Baldwin MM, Daynes E, Hong A, Mills G, Goddard ACN, et al. Early experiences of the Your COVID Recovery® digital programme for individuals with long COVID. *BMJ Open Respir Res*. 2022 Sep 1;9(1):e001237.
 21. Hastie CE, Lowe DJ, McAuley A, Winter AJ, Mills NL, Black C, et al. Outcomes among confirmed cases and a matched comparison group in the Long-COVID in Scotland study. *Nat Commun*. 2022 Oct 12;13(1):5663.
 22. O'Brien KK, Brown DA, McDuff K, St Clair-Sullivan N, Solomon P, Chan Carusone S, et al. Conceptualising the episodic nature of disability among adults living with Long COVID: a qualitative study. *BMJ Glob Health*. 2023 Mar;8(3):e011276.
 23. Kerksieck P, Ballouz T, Haile SR, Schumacher C, Lacy J, Domenghino A, et al. Post COVID-19 condition, work ability and occupational changes in a population-based cohort. *Lancet Reg Health - Eur*. 2023 Aug 1;31:100671.
 24. Ayoubkhani D, Bermingham C, Pouwels KB, Glickman M, Nafilyan V, Zaccardi F, et al. Trajectory of long covid symptoms after covid-19 vaccination: community based cohort study. *BMJ*. 2022 May 18;377:e069676.
 25. Alliance UHDR, NHSX. Building Trusted Research Environments - Principles and Best Practices; Towards TRE ecosystems [Internet]. Zenodo; 2021 Dec [cited 2023 Dec 18]. Available from: <https://zenodo.org/records/5767586>
 26. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. *Appl Health Econ Health Policy*. 2017 Apr 1;15(2):127–37.
 27. Herrett E, Tomlin K, Lin LY, Tomlinson LA, Jit M, Briggs A, et al. Protocol for an OpenSAFELY cohort study collecting patient-reported outcome measures using the TPP Airmid smartphone application and linked big data to quantify the health and economic costs of long COVID (OpenPROMPT). *BMJ Open*. 2023 Feb 1;13(2):e071261.
 28. H T. TPP. 2020 [cited 2023 Dec 18]. Airmid. Available from: <https://tpp-uk.com/products/airmid/>
 29. Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System: properties, applications, and interpretation. *Health Qual Life Outcomes*. 2003 Dec 16;1:79.
 30. NHS Digital [Internet]. [cited 2023 Dec 18]. The NHS England OpenSAFELY COVID-19 service - privacy notice. Available from: <https://digital.nhs.uk/coronavirus/coronavirus-covid-19-response-information-governance-hub/the-nhs-england-opensafely-covid-19-service-privacy-notice>
 31. NHS Digital [Internet]. [cited 2023 Dec 18]. Data Security and Protection Toolkit. Available from: <https://digital.nhs.uk/data-and-information/looking-after-information/data-security-and-information-governance/data-security-and-protection-toolkit>
 32. NHS Digital [Internet]. [cited 2023 Dec 18]. ISB1523: Anonymisation Standard for Publishing Health and Social Care Data. Available from: <https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/isb1523-anonymisation-standard-for-publishing-health-and-social-care-data>
 33. GOV.UK [Internet]. 2022 [cited 2023 Dec 18]. [Withdrawn] [withdrawn] Coronavirus (COVID-19): notice under regulation 3(4) of the Health Service (Control of Patient Information) Regulations 2002 – general. Available from: <https://www.gov.uk/government/publications/coronavirus-covid-19-notification-of-data-controllers-to-share-information/coronavirus-covid-19-notice-under-regulation-34-of-the-health-service-control-of-patient-information-regulations-2002-general--2>
 34. NHS Digital [Internet]. [cited 2023 Dec 18]. COVID-19 Public Health Directions 2020. Available from:

<https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notice/secretary-of-state-directions/covid-19-public-health-directions-2020>

35. Health Research Authority [Internet]. [cited 2023 Dec 18]. Confidentiality Advisory Group. Available from: <https://www.hra.nhs.uk/about-us/committees-and-services/confidentiality-advisory-group/>

Appendix

Supplementary Methods

Comorbidities were assessed at study entry. Relevant comorbidities were defined based on previous research of risk factors for Long COVID and COVID-19 in OpenSAFELY (1–3). A previous code for one or more of: diabetes; cancer; haematological cancer; asthma; chronic respiratory disease; chronic cardiac disease; chronic liver disease; stroke or dementia; other neurological condition; organ transplant; dysplasia; rheumatoid arthritis, systemic lupus erythematosus or psoriasis; or other immunosuppressive conditions. Those with no relevant code for a condition will be assumed not to have that condition. Number of conditions were categorised into “0”, “1”, and “2 or more”. All codelists used to define these comorbidities are listed on GitHub

(<https://github.com/opensafely/openprompt-cohort-profile/tree/main/codelists>), and are publicly available on <https://www.opencodelists.org/>.

Supplementary Figures

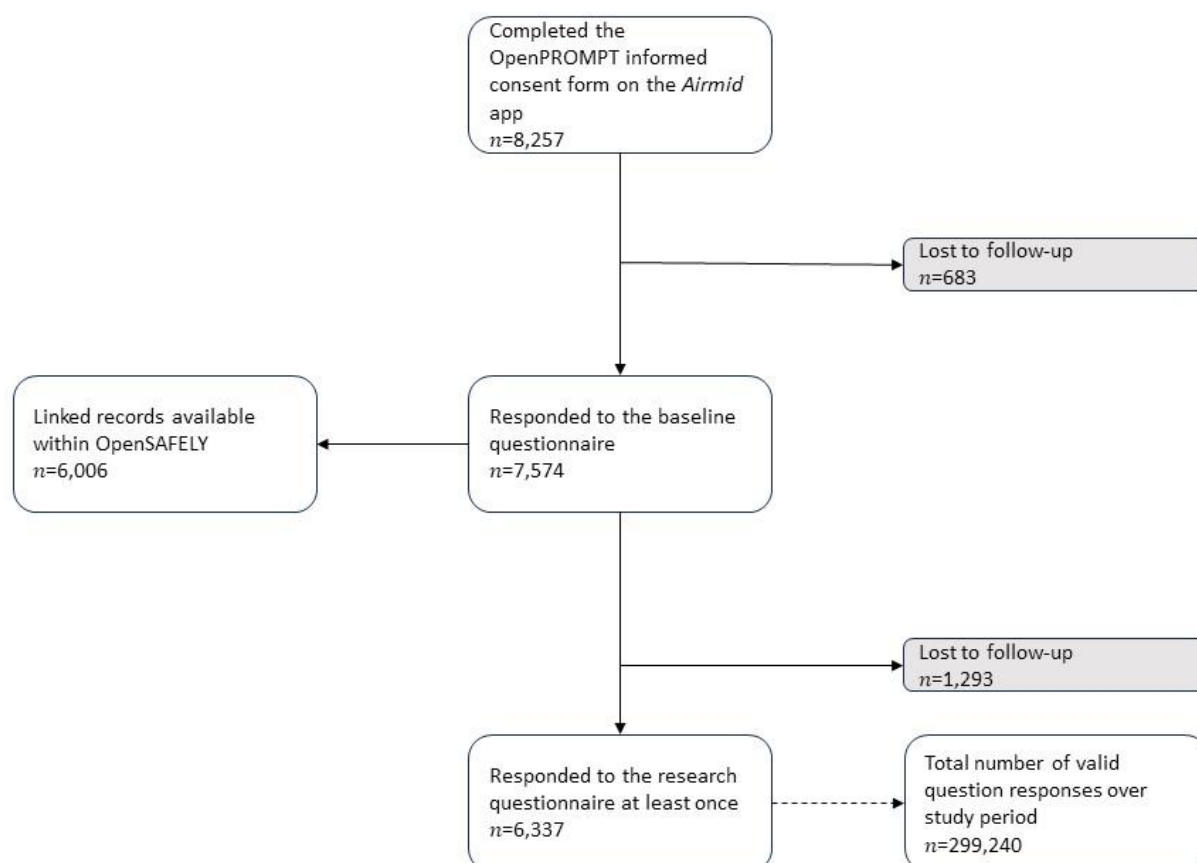


Figure S1: Cohort enrolment using an app-based data collection method for OpenPROMPT

Supplementary Tables

Table S1: characteristics of OpenPROMPT participants at cohort entry and responses to the baseline questionnaire

Variable	Level	Summary at baseline	Percentage of respondents with linked data available
Gender (app question)	Male	2,555 (33.7%)	NA
	Female	4,599 (60.7%)	NA
	Intersex/non-binary/other	77 (1%)	NA
	Refused	343 (4.5%)	NA
Ethnicity (app question)	White	7,203 (95.1%)	NA
	Mixed	126 (1.7%)	NA
	Asian/Asian Brit	147 (1.9%)	NA
	Black/African/Caribbean/Black Brit	42 (0.6%)	NA
	Other/not stated	63 (0.8%)	NA
Highest education (app question)	None/less than primary school	14 (0.2%)	NA
	Primary School	21 (0.3%)	NA
	Secondary/high school	1,778 (23.5%)	NA
	College/university	4,004 (52.9%)	NA
	Postgraduate qualification	1,666 (22%)	NA
	Not stated	84 (1.1%)	NA
Disability (app question)	No	4,515 (59.6%)	NA
	Yes	2,891 (38.2%)	NA
	Not stated	161 (2.1%)	NA
Relationship status (app question)	Single person	1,414 (18.7%)	NA
	Cohabiting	1,078 (14.2%)	NA
	Married/civil partner	4,172 (55.1%)	NA
	Separated	189 (2.5%)	NA
	Divorced/dissolved civil partnership	392 (5.2%)	NA

Variable	Level	Summary at baseline	Percentage of respondents with linked data available
	Widowed/surviving civil partner	224 (3%)	NA
	Not stated	105 (1.4%)	NA
Household income (app question)	£6,000-12,999	749 (9.9%)	NA
	£13,000-18,999	679 (9%)	NA
	£19,000-25,999	882 (11.6%)	NA
	£26,000-31,999	770 (10.2%)	NA
	£32,000-47,999	1,316 (17.4%)	NA
	£48,000-63,999	966 (12.8%)	NA
	£64,000-95,999	791 (10.4%)	NA
	£96,000	476 (6.3%)	NA
	Not stated	945 (12.5%)	NA
Linked records available	FALSE	1,568 (20.7%)	NA
	TRUE	6,006 (79.3%)	NA
Age (median, IQR)		51 (41-61)	NA
Age category	18-29	532 (7%)	8.9%
	30-39	945 (12.5%)	15.7%
	40-49	1,407 (18.6%)	23.4%
	50-59	1,575 (20.8%)	26.2%
	60-69	1,141 (15.1%)	19.0%
	70+	399 (5.3%)	6.6%
	(Missing)	1,568 (20.7%)	-
Region	East	1,407 (23.5%)	23.4%
	East Midlands	1,218 (20.3%)	20.3%
	London	133 (2.2%)	2.2%
	North East	259 (4.3%)	4.3%
	North West	511 (8.5%)	8.5%

Variable	Level	Summary at baseline	Percentage of respondents with linked data available
IMD (quintile)	South East	392 (6.5%)	6.5%
	South West	987 (16.5%)	16.4%
	West Midlands	182 (3%)	3.0%
	Yorkshire and The Humber	910 (15.2%)	15.2%
	1 (most deprived)	903 (11.9%)	15.0%
	2	1,043 (13.8%)	17.4%
	3	1,204 (15.9%)	20.0%
	4	1,197 (15.8%)	19.9%
Rural/urban classification	5 (least deprived)	1,358 (17.9%)	22.6%
	(Missing)	1,862 (24.6%)	-
	Urban	4,424 (58.4%)	73.7%
Household composition (2020)	Rural	1,288 (17%)	21.4%
	(Missing)	1,855 (24.5%)	-
	0	NA (NA%)	NA%
	1	903 (11.9%)	15.0%
	2	1,743 (23%)	29.0%
	3	966 (12.8%)	16.1%
	4	651 (8.6%)	10.8%
	5+	399 (5.3%)	6.6%
Comorbidities	(Missing)	2,905 (38.4%)	-
	0	3,409 (45%)	56.8%
	1	2,009 (26.5%)	33.4%
	2+	588 (7.8%)	9.8%
Number of Long COVID records	(Missing)	1,568 (20.7%)	-
	None	5,642 (74.5%)	93.9%
	At least 1	364 (4.8%)	6.1%

Variable	Level	Summary at baseline	Percentage of respondents with linked data available
Number of positive COVID-19 test	(Missing)	1,568 (20.7%)	-*
	0	2,688 (35.5%)	44.8%
	1	1,519 (20.1%)	25.3%
	2	875 (11.6%)	14.6%
	3	385 (5.1%)	6.4%
	4	189 (2.5%)	3.1%
	5+	357 (4.7%)	5.9%
	(Missing)	1,568 (20.7%)	-
Number of previous COVID-19 vaccinations	0	245 (3.2%)	4.1%
	1	133 (1.8%)	2.2%
	2	287 (3.8%)	4.8%
	3	1,155 (15.2%)	19.2%
	4+	4,193 (55.4%)	69.8%
	(Missing)	1,568 (20.7%)	-

Table S2: Responses to the OpenPROMPT research questionnaire over time since study enrolment (baseline, +30 days, +60 days, +90 days)

Variable	Level	Baseline	Survey 2	Survey 3	Survey 4
N	-	6,337	1,809	1,206	884
Days since baseline	(median, IQR)	0 (0-0)	30 (30-31)	60 (60-61)	90 (90-91)
Quality of Life Score [0-100]	(median, IQR)	70 (50-80)	70 (40-80)	70 (40-85)	70 (50-85)
COVID-19 status	Yes (+ve test)	4,767 (77.4%)	1,295 (73.4%)	868 (73.4%)	623 (72.4%)
	Yes (medical advice)	259 (4.2%)	70 (4%)	42 (3.6%)	35 (4.1%)
	Yes (suspected)	NA (NA%)	NA (NA%)	NA (NA%)	NA (NA%)
	Unsure	147 (2.4%)	42 (2.4%)	21 (1.8%)	NA (NA%)
	No	980 (15.9%)	350 (19.8%)	245 (20.7%)	196 (22.8%)
	Not stated	NA (NA%)	NA (NA%)	NA (NA%)	NA (NA%)
Recovered from most recent COVID-19 infection?	Yes, back to normal	2,835 (51.7%)	805 (53.2%)	581 (57.6%)	434 (60.2%)
	No, still have symptoms	2,653 (48.3%)	707 (46.8%)	427 (42.4%)	294 (40.8%)
	Not stated	NA (NA%)	NA (NA%)	NA (NA%)	NA (NA%)
Have you had at least one COVID-19 vaccine dose?	Yes	6,153 (97.1%)	1,764 (97.7%)	1,183 (98.3%)	861 (97.6%)
	No	175 (2.8%)	42 (2.3%)	28 (2.3%)	28 (3.2%)
	Not stated	NA (NA%)	NA (NA%)	NA (NA%)	NA (NA%)
Currently employed	Employed	3,584 (56.6%)	861 (47.7%)	539 (44.8%)	385 (43.7%)
	Unemployed	2,758 (43.5%)	952 (52.7%)	665 (55.2%)	497 (56.3%)
MRC breathlessness scale	Grade 1: I only get breathless with strenuous exercise	1,988 (31.4%)	581 (32.2%)	392 (32.6%)	301 (34.1%)
	Grade 2: I get short of breath when hurrying on level ground or walking up a slight hill	2,485 (39.2%)	672 (37.2%)	462 (38.4%)	315 (35.7%)
	Grade 3: On level ground, I walk slower than people	1,190 (18.8%)	322 (17.8%)	210 (17.4%)	154 (17.5%)

Variable	Level	Baseline	Survey 2	Survey 3	Survey 4
	of my age because of breathlessness, or I have to stop for breath when walking at my own pace on the level				
	Grade 4: I stop for breath after walking about 100 yards or after a few minutes on level ground	483 (7.6%)	168 (9.3%)	105 (8.7%)	77 (8.7%)
	Grade 5: I am too breathless to leave the house or I am breathless when dressing/undressing	196 (3.1%)	63 (3.5%)	42 (3.5%)	35 (4%)
Number of previous COVID-19 infections	0	903 (14.3%)	301 (16.7%)	210 (17.4%)	168 (19%)
	1	2,842 (44.9%)	861 (47.7%)	560 (46.5%)	392 (44.4%)
	2	1,799 (28.4%)	455 (25.2%)	315 (26.2%)	224 (25.4%)
	3	602 (9.5%)	154 (8.5%)	91 (7.6%)	77 (8.7%)
	4	133 (2.1%)	28 (1.6%)	28 (2.3%)	14 (1.6%)
	5	49 (0.8%)	7 (0.4%)	NA (NA%)	NA (NA%)
	6+	NA (NA%)	NA (NA%)	NA (NA%)	NA (NA%)
Duration of longest COVID-19 infection	Less than 2 weeks	1,547 (28.7%)	448 (30.5%)	301 (30.7%)	245 (35%)
	2 – 3 weeks	1,155 (21.4%)	287 (19.5%)	203 (20.7%)	133 (19%)
	4 – 12 weeks	812 (15.1%)	217 (14.8%)	126 (12.9%)	77 (11%)
	More than 12 weeks	1,876 (34.8%)	525 (35.7%)	357 (36.4%)	245 (35%)
Number of previous vaccine doses	0	NA (NA%)	NA (NA%)	NA (NA%)	NA (NA%)
	1	56 (0.9%)	7 (0.4%)	NA (NA%)	NA (NA%)

Variable	Level	Baseline	Survey 2	Survey 3	Survey 4
	2	539 (8.8%)	105 (6%)	70 (6%)	42 (5%)
	3	2,310 (37.8%)	574 (32.9%)	385 (32.7%)	273 (32.2%)
	4	2,534 (41.5%)	798 (45.8%)	518 (44%)	364 (43%)
	5	483 (7.9%)	168 (9.6%)	112 (9.5%)	98 (11.6%)
	6+	175 (2.9%)	84 (4.8%)	84 (7.1%)	77 (9.1%)
EQ5D: mobility	none	3,297 (52%)	847 (46.9%)	553 (45.9%)	413 (46.8%)
	slight	1,400 (22.1%)	420 (23.3%)	294 (24.4%)	217 (24.6%)
	moderate	966 (15.2%)	301 (16.7%)	196 (16.3%)	154 (17.5%)
	severe	630 (9.9%)	231 (12.8%)	147 (12.2%)	105 (11.9%)
	unable	42 (0.7%)	7 (0.4%)	14 (1.2%)	NA (NA%)
EQ5D: care for self	none	4,522 (71.4%)	1,253 (69.4%)	812 (67.4%)	602 (68.3%)
	slight	903 (14.3%)	259 (14.3%)	196 (16.3%)	140 (15.9%)
	moderate	665 (10.5%)	210 (11.6%)	133 (11%)	112 (12.7%)
	severe	224 (3.5%)	84 (4.7%)	56 (4.7%)	28 (3.2%)
	unable	28 (0.4%)	7 (0.4%)	NA (NA%)	NA (NA%)
EQ5D: usual activities	none	2,779 (43.9%)	763 (42.2%)	518 (43%)	371 (42.1%)
	slight	1,624 (25.6%)	469 (26%)	294 (24.4%)	224 (25.4%)
	moderate	1,141 (18%)	308 (17.1%)	217 (18%)	161 (18.3%)
	severe	616 (9.7%)	217 (12%)	133 (11%)	98 (11.1%)
	unable	175 (2.8%)	56 (3.1%)	42 (3.5%)	28 (3.2%)
EQ5D: pain/discomfort	none	1,939 (30.6%)	504 (27.9%)	343 (28.5%)	259 (29.4%)
	slight	1,890 (29.8%)	574 (31.8%)	357 (29.7%)	280 (31.7%)
	moderate	1,463 (23.1%)	420 (23.3%)	294 (24.4%)	203 (23%)
	severe	749 (11.8%)	217 (12%)	154 (12.8%)	98 (11.1%)
	unable	294 (4.6%)	98 (5.4%)	56 (4.7%)	42 (4.8%)
EQ5D: anxiety/depression	none	2,240 (35.4%)	735 (40.7%)	490 (40.7%)	392 (44.4%)
	slight	1,750 (27.6%)	455 (25.2%)	343 (28.5%)	224 (25.4%)

Variable	Level	Baseline	Survey 2	Survey 3	Survey 4
How much has work been affected	moderate	1,400 (22.1%)	371 (20.5%)	224 (18.6%)	161 (18.3%)
	severe	581 (9.2%)	140 (7.8%)	84 (7%)	63 (7.1%)
	unable	371 (5.9%)	105 (5.8%)	63 (5.2%)	49 (5.6%)
	0 (No effect on my daily activities)	1,106 (37.8%)	315 (44.6%)	189 (45%)	133 (42.2%)
	1	287 (9.8%)	49 (6.9%)	42 (10%)	21 (6.7%)
	2	301 (10.3%)	70 (9.9%)	35 (8.3%)	28 (8.9%)
	3	266 (9.1%)	63 (8.9%)	35 (8.3%)	35 (11.1%)
	4	182 (6.2%)	42 (5.9%)	28 (6.7%)	14 (4.4%)
	5	168 (5.7%)	35 (5%)	21 (5%)	14 (4.4%)
	6	161 (5.5%)	35 (5%)	14 (3.3%)	14 (4.4%)
	7	168 (5.7%)	28 (4%)	21 (5%)	21 (6.7%)
	8	91 (3.1%)	21 (3%)	14 (3.3%)	7 (2.2%)
9	42 (1.4%)	7 (1%)	NA (NA%)	NA (NA%)	
	10 (Completely prevented me from doing my daily activities)	154 (5.3%)	28 (4%)	21 (5%)	21 (6.7%)
How much has life been affected	0 (No effect on my daily activities)	1,792 (28.3%)	546 (30.2%)	364 (30.2%)	280 (31.7%)
	1	595 (9.4%)	175 (9.7%)	126 (10.5%)	77 (8.7%)
	2	602 (9.5%)	154 (8.5%)	98 (8.1%)	91 (10.3%)
	3	532 (8.4%)	147 (8.1%)	91 (7.6%)	70 (7.9%)
	4	378 (6%)	105 (5.8%)	56 (4.7%)	49 (5.6%)
	5	476 (7.5%)	119 (6.6%)	98 (8.1%)	70 (7.9%)
	6	448 (7.1%)	119 (6.6%)	84 (7%)	49 (5.6%)
	7	595 (9.4%)	182 (10.1%)	91 (7.6%)	63 (7.1%)
	8	448 (7.1%)	119 (6.6%)	105 (8.7%)	77 (8.7%)
	9	231 (3.6%)	77 (4.3%)	35 (2.9%)	28 (3.2%)

Variable	Level	Baseline	Survey 2	Survey 3	Survey 4
	10 (Completely prevented me from doing my daily activities)	238 (3.8%)	77 (4.3%)	56 (4.7%)	35 (4%)
FACIT: feel fatigued	Very much	812 (12.8%)	301 (16.7%)	210 (17.4%)	147 (16.7%)
	Quite a bit	1,393 (22%)	406 (22.5%)	294 (24.4%)	210 (23.8%)
	Somewhat	1,043 (16.5%)	308 (17.1%)	189 (15.7%)	154 (17.5%)
	A little bit	1,477 (23.3%)	385 (21.3%)	245 (20.3%)	182 (20.6%)
	Not at all	1,617 (25.5%)	413 (22.9%)	273 (22.7%)	189 (21.4%)
FACIT: feel weak	Very much	2,149 (33.9%)	651 (36%)	434 (36%)	308 (34.9%)
	Quite a bit	1,428 (22.5%)	413 (22.9%)	280 (23.3%)	224 (25.4%)
	Somewhat	1,120 (17.7%)	287 (15.9%)	182 (15.1%)	140 (15.9%)
	A little bit	980 (15.5%)	266 (14.7%)	196 (16.3%)	126 (14.3%)
	Not at all	658 (10.4%)	189 (10.5%)	119 (9.9%)	84 (9.5%)
FACIT: feel listless	Very much	1,568 (24.8%)	497 (27.5%)	364 (30.2%)	266 (30.2%)
	Quite a bit	1,491 (23.5%)	469 (26%)	287 (23.8%)	203 (23%)
	Somewhat	1,015 (16%)	231 (12.8%)	161 (13.4%)	147 (16.7%)
	A little bit	1,183 (18.7%)	315 (17.4%)	203 (16.9%)	140 (15.9%)
	Not at all	1,085 (17.1%)	294 (16.3%)	196 (16.3%)	133 (15.1%)
FACIT: feel tired	Very much	651 (10.3%)	245 (13.6%)	168 (14%)	126 (14.3%)
	Quite a bit	1,533 (24.2%)	469 (26%)	350 (29.1%)	238 (27%)
	Somewhat	1,022 (16.1%)	294 (16.3%)	168 (14%)	147 (16.7%)
	A little bit	1,393 (22%)	364 (20.2%)	245 (20.3%)	175 (19.8%)
	Not at all	1,736 (27.4%)	434 (24%)	273 (22.7%)	196 (22.2%)
FACIT: trouble starting things	Very much	1,603 (25.3%)	518 (28.7%)	364 (30.2%)	266 (30.2%)
	Quite a bit	1,344 (21.2%)	392 (21.7%)	259 (21.5%)	189 (21.4%)
	Somewhat	1,092 (17.2%)	280 (15.5%)	203 (16.9%)	154 (17.5%)
	A little bit	1,190 (18.8%)	287 (15.9%)	182 (15.1%)	147 (16.7%)

Variable	Level	Baseline	Survey 2	Survey 3	Survey 4
FACIT: trouble finishing things	Not at all	1,113 (17.6%)	322 (17.8%)	196 (16.3%)	133 (15.1%)
	Very much	1,582 (25%)	518 (28.7%)	371 (30.8%)	273 (31%)
	Quite a bit	1,372 (21.7%)	371 (20.5%)	259 (21.5%)	175 (19.8%)
	Somewhat	1,029 (16.2%)	273 (15.1%)	175 (14.5%)	133 (15.1%)
	A little bit	1,148 (18.1%)	294 (16.3%)	189 (15.7%)	147 (16.7%)
FACIT: have energy	Not at all	1,204 (19%)	350 (19.4%)	217 (18%)	147 (16.7%)
	Not at all	784 (12.4%)	238 (13.2%)	168 (14%)	112 (12.7%)
	A little bit	2,016 (31.8%)	532 (29.5%)	350 (29.1%)	266 (30.2%)
	Somewhat	1,764 (27.8%)	497 (27.5%)	287 (23.8%)	217 (24.6%)
	Quite a bit	1,344 (21.2%)	399 (22.1%)	287 (23.8%)	210 (23.8%)
FACIT: able to do usual activities	Very much	434 (6.9%)	147 (8.1%)	112 (9.3%)	77 (8.7%)
	Not at all	441 (7%)	140 (7.8%)	112 (9.3%)	63 (7.1%)
	A little bit	1,449 (22.9%)	462 (25.6%)	287 (23.8%)	224 (25.4%)
	Somewhat	1,757 (27.7%)	413 (22.9%)	287 (23.8%)	210 (23.8%)
	Quite a bit	1,316 (20.8%)	364 (20.2%)	224 (18.6%)	161 (18.3%)
FACIT: need sleep during the day	Very much	1,379 (21.8%)	427 (23.6%)	301 (25%)	224 (25.4%)
	Very much	1,827 (28.8%)	567 (31.4%)	378 (31.4%)	273 (31%)
	Quite a bit	1,967 (31%)	553 (30.6%)	385 (32%)	287 (32.5%)
	Somewhat	952 (15%)	287 (15.9%)	189 (15.7%)	112 (12.7%)
	A little bit	847 (13.4%)	210 (11.6%)	126 (10.5%)	112 (12.7%)
FACIT: too tired to eat	Not at all	735 (11.6%)	189 (10.5%)	126 (10.5%)	98 (11.1%)
	Very much	3,822 (60.3%)	1,127 (62.4%)	749 (62.2%)	560 (63.5%)
	Quite a bit	1,253 (19.8%)	336 (18.6%)	231 (19.2%)	140 (15.9%)
	Somewhat	777 (12.3%)	196 (10.9%)	126 (10.5%)	105 (11.9%)
	A little bit	357 (5.6%)	112 (6.2%)	70 (5.8%)	42 (4.8%)
	Not at all	126 (2%)	35 (1.9%)	28 (2.3%)	35 (4%)

Variable	Level	Baseline	Survey 2	Survey 3	Survey 4
FACIT: need help for usual activities	Very much	3,395 (53.6%)	959 (53.1%)	637 (52.9%)	462 (52.4%)
	Quite a bit	1,288 (20.3%)	364 (20.2%)	231 (19.2%)	182 (20.6%)
	Somewhat	777 (12.3%)	196 (10.9%)	140 (11.6%)	98 (11.1%)
	A little bit	518 (8.2%)	161 (8.9%)	112 (9.3%)	84 (9.5%)
	Not at all	357 (5.6%)	126 (7%)	84 (7%)	56 (6.3%)
FACIT: frustrated by being too tired	Very much	1,799 (28.4%)	553 (30.6%)	385 (32%)	280 (31.7%)
	Quite a bit	1,239 (19.6%)	364 (20.2%)	238 (19.8%)	175 (19.8%)
	Somewhat	693 (10.9%)	196 (10.9%)	119 (9.9%)	91 (10.3%)
	A little bit	854 (13.5%)	217 (12%)	154 (12.8%)	126 (14.3%)
	Not at all	1,750 (27.6%)	476 (26.4%)	315 (26.2%)	210 (23.8%)
FACIT: limit social activities	Very much	2,016 (31.8%)	630 (34.9%)	441 (36.6%)	322 (36.5%)
	Quite a bit	1,211 (19.1%)	329 (18.2%)	224 (18.6%)	154 (17.5%)
	Somewhat	847 (13.4%)	217 (12%)	140 (11.6%)	119 (13.5%)
	A little bit	861 (13.6%)	245 (13.6%)	147 (12.2%)	112 (12.7%)
	Not at all	1,407 (22.2%)	385 (21.3%)	259 (21.5%)	175 (19.8%)