BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers’ comments and the authors’ responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (http://bmjopen.bmj.com).

If you have any questions on BMJ Open’s open peer review process please email info.bmjopen@bmj.com
Long Covid and the mental and physical health of children and young people; a national matched cohort study (CLoCk).

<table>
<thead>
<tr>
<th>Journal</th>
<th>BMJ Open</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>bmjopen-2021-052838</td>
</tr>
<tr>
<td>Article Type</td>
<td>Protocol</td>
</tr>
<tr>
<td>Date Submitted by the Author</td>
<td>06-May-2021</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Stephenson, Terence; UCL Great Ormond Street Institute of Child Health Population Policy and Practice Shafran, Roz; UCL Great Ormond Street Institute of Child Health Population Policy and Practice De Stavola, Bianca; UCL Great Ormond Street Institute of Child Health Population Policy and Practice Rojas, Natalia; UCL Great Ormond Street Institute of Child Health Population Policy and Practice Aiano, Felicity; Public Health England Amin-Chowdhury, Zahin; Public Health England McOwat, Kelsey; Public Health England Simmons, Ruth; Public Health England Zavala, Maria; Public Health England CLoCk, Consortium; UCL Great Ormond Street Institute of Child Health Population Policy and Practice Ladhani, Shamez N.; Public Health England</td>
</tr>
<tr>
<td>Keywords:</td>
<td>COVID-19, PAEDIATRICS, INFECTIOUS DISEASES, PUBLIC HEALTH, VIROLOGY</td>
</tr>
</tbody>
</table>
I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd (“BMJ”) its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge (“APC”) for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author’s Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.
Long Covid and the mental and physical health of children and young people; a national matched cohort study (CLoCk).

Abstract 250 words. Main paper 2294 words.

Stephenson T\textsuperscript{1}, Shafran R\textsuperscript{1}, De Stavola B\textsuperscript{1}, Rojas N\textsuperscript{1}, Aiano F\textsuperscript{2}, Amin-Chowdhury Z\textsuperscript{2}, McOwat K\textsuperscript{2}, Simmons R\textsuperscript{2}, Zavala M\textsuperscript{2}, CLoCk Consortium\textsuperscript{1}, Ladhani S\textsuperscript{2}.

1 UCL Great Ormond Street Institute of Child Health
2 Public Health England, Colindale

Corresponding author: Professor Sir Terence Stephenson; UCL Great Ormond Street Institute of Child Health, 30 Guilford Street, London, WC1N 1EH; Email: t.stephenson@ucl.ac.uk; Telephone: 020 7242 9789; Fax Number: N/A

Five Keywords/Phrases: Long COVID, children, young people, cohort study, physical and mental health
Abstract

Introduction: There is uncertainty surrounding the diagnosis, prevalence, phenotype, duration and treatment of Long COVID. This study aims to (a) describe the clinical phenotype of post-COVID symptomatology in CYP with laboratory-confirmed SARS-CoV-2 infection compared to test-negative controls (b) produce an operational definition of Long COVID in CYP (c) establish the prevalence of Long COVID in CYP.

Methods and analysis: A cohort study of SARS-CoV-2 positive CYP aged 11-17 years compared with age-, sex- and geographically-matched SARS-CoV-2 test-negative CYP. CYP aged 11-17 testing positive and negative for SARS-CoV-2 infection will be identified and contacted 3, 6, 12 and 24 months after the test date. Consenting CYP will complete an online questionnaire. We will recruit 3,000 test positives and 3,000 test negatives. Data visualisation techniques will be used to examine trajectories over time for symptoms and variables measured repeatedly, separately by original test status. Summary measures of fatigue and mental health dimensions will be generated using dimension reduction methods such as latent variables/latent class/PCA methods. Cross-tabulation of collected and derived variables against test status and discriminant analysis will help operationalise preliminary definitions of Long COVID.

Ethics and Dissemination: Research Ethics Committee Approval granted. Data will be stored in secure PHE servers or UCL’s Data Safe Haven. Risks of harm will be minimised by providing information on where to seek support. Results will be published on a pre-print server followed by journal publication, with reuse of articles under a CC BY licence. Data will be published with protection against identification when small frequencies.

Study registration: Study registered with ISRCTN: Trial ID: ISRCTN34804192 and granted Urgent Public Health prioritisation.

Article Summary

STRENGTHS

- This study does not start with an arbitrary definition of a new condition – ‘opinion based medicine’
- Rather it represents ‘evidence based medicine’ - we will seek the views of 6,000 children and young people on what they have experienced physically and mentally in the months following COVID-19
- The study incorporates a comparator matched cohort of children and young people who have experienced a pandemic, school closure and social isolation but who had a negative COVID test

WEAKNESSES

- The comparator matched cohort of children and young people must also have had a reason to seek a COVID test: symptoms, anxiety, a contact or a bereavement
- Ideally, we would test all children for antibodies to confirm COVID-19 positive and negative status throughout the study period but we thought would significantly reduce participation
Introduction

Currently, there is huge uncertainty surrounding the diagnosis, prevalence, phenotype, duration and treatment of Long COVID. There is currently no diagnostic test or code for Long COVID. Consequently, cases of Long COVID are not captured in routine National Health Service (NHS) administrative datasets. Instead, Long COVID may be coded as a variety of different conditions in hospital or symptom clusters in non-hospitalised CYP. However, it is possible the symptoms associated with Long COVID are in fact a mixture of factors relating to the pandemic and lockdown as a whole rather than the viral infection itself. For example, factors such as social isolation, anxiety, depression or educational concerns may be the root cause of these symptoms in CYP both with and without SARS-CoV-2 infection. The effects on the developing brain and behaviour of adolescents could be far reaching.\(^1\)

Despite acute COVID-19 illness being milder in CYP, it should not be assumed that those CYP at low risk of life-threatening acute infections\(^2\)\(^3\) do not suffer the longer-term consequences of SARS-CoV-2 infection. It is important to study Long COVID within this population given that the psychological and social impact of Long COVID could have major consequences for transition to adulthood. There is a clear need to define the clinical phenotype of Long COVID in order to understand those most at risk, the illness trajectory and to provide accurate information on the natural course of the condition. It is equally important to understand both the physical and mental health impact on CYP with Long COVID, given the rise in mental health problems among CYP since the start of the pandemic.\(^4\) Obtaining a better understanding of which CYP are affected will help with targeting potential interventions.

The greatest risk factor for severe acute SAR-CoV-2 illness has been old age, with other risk factors at all ages including obesity, comorbid long-term conditions, learning and neurological disabilities, mental health problems and ethnic minority status. It is plausible that such CYP may be most at risk of suffering from Long COVID, given that much is still unknown in CYP about the immunological susceptibility and underlying biology of Long COVID. Of all CYP, those likely to be most at risk of Long COVID are teenagers, with existing literature showing they make up the majority of CYP with chronic fatigue, post viral syndromes and persistent symptoms.\(^5\)

In adults, there is emerging evidence that gender is a risk factor for Long COVID, with middle-aged females more susceptible than men.\(^6\) In terms of symptoms, two main groupings have been identified (i) respiratory symptoms (e.g. cough, shortness of breath) as well as fatigue and headaches, and (ii) multi-system, including the brain, gut and heart. Among adults with Long COVID, heart symptoms such as palpitations or fast heartbeat, pins and needles or numbness and problems concentrating (‘brain fog’) have been reported. Experiencing a greater number of symptoms during the first week of infection, as well as older age, have also been linked to Long COVID. It should be noted mental health conditions were not reported in this cohort. Four syndromes have been described in adults: Post-viral fatigue; Fluctuating multi-system symptoms; Post-intensive care syndrome; and Lasting organ damage. It is known in both adults and young people that a wide range of long-term physical conditions increase the risk of mental ill-health, particularly if the condition involves the central nervous system, with several studies...
reporting increased rates of all common mental health conditions in children. Moreover, fatigue has also been studied in paediatric long-term conditions including CYP with multiple sclerosis. Emerging clinical observations and preliminary research indicate that COVID-19 can have a long-term impact on CYP in a range of domains but well conducted, methodologically robust studies are lacking.

This study primarily aims to (a) describe the clinical phenotype and prevalence of post-COVID symptoms (e.g. pain or physical symptoms, fatigue, sleep problems, mental health problems) among test positive and test negative CYP, (b) produce an operational definition of Long COVID in CYP, and (c) establish the prevalence of Long COVID in CYP testing positive for SARS-CoV-2 infection. The second aim is of particular importance given that this is a prerequisite for any future epidemiological or interventional study. This will be achieved by using the empirical data we will acquire from 6,000 CYP and a national Delphi consensus process.

**Methods**

**Study design and setting**

Public Health England (PHE) has been conducting national surveillance of SARS-CoV-2 since the start of the pandemic in England. PHE receives daily electronic notifications of all SARS-CoV-2 RT-PCR tests performed in healthcare settings (Pillar 1 tests) and in the community (Pillar 2 tests), which are reported through the Second Generation Surveillance System (SGSS). Information within the SGSS reports includes NHS number, name, age, sex, postcode, date of sample, reporting laboratory and test result. PHE also has access to the electronic Patient Demographic Service (PDS), which contains the names, addresses and status (alive/dead) of all patients registered with the NHS.

A sample of CYP aged 11-17 years old when testing positive for SARS-CoV-2 between September 2020 and March 2021 and twice as many COVID test-negative CYP matched for age, sex and region identified through SGSS will be linked to PDS using available identifiers and postal addresses. A letter will be posted to them, informing them about the study and inviting them to take part using an online link. This link will provide them with information about the study, with an option to consent online and complete a short recruitment questionnaire (or both as paper options).

This will be a cohort study of SARS-CoV-2 test-positive CYP aged 11-17 years matched on the test date, age, sex and geography to SARS-CoV-2 test-negative controls identified by PHE.

**Participants**

Between 1st September 2020 and 31st March 2021, 234,803 CYP aged between 11 and 17 years were tested positive for SARS-CoV-2 in England, and 1,245,288 CYP tested negative. Among those who tested negative, 80,365 were excluded as they had a positive result pre- and/or post their negative test and we wish to compare test-positive with test-negative CYP. The positive and
negative cohorts were matched using age at test, gender, month of test, and lower super output area resulting in 102,402 positive individuals and 147,561 negative individuals. Matched individuals were checked against PDS to exclude any individuals who have died, and to extract participants addresses to send invitation letters. 33 individuals were excluded because of death since their COVID-19 test (6 among those positive and 27 among those negative), and 11,982 individuals of the positive cohort and 16,635 individuals of the negative cohort were excluded as an address was not available. Finally, 246 CYP were excluded as they were included in a previous pilot study which helped inform our methodology. Following exclusions, 90,227 positive CYP and 130,243 negative CYP form the population from which we will randomly sampled individuals for inclusion within the study.

Participants will be contacted in a ratio of 1:2 positives:negatives on the expectation of a lower participation rate in the latter. We anticipate 20% to be recruited in the positive group, as estimated by CYP recruitment rates in other PHE studies on COVID-19 (sKIDs)\textsuperscript{12}. We also oversampled CYP who had a test in the months of January-March 2021 because only for these groups we will collect information at 3 months of follow-up.

This is shown schematically in Figure 1.

Data collection

Participating families of CYP will be contacted 3, 6, 12 and 24 months after the CYP’s SARS-CoV-2 test (depending on recruitment month) and invited to take part in the study. Following online informed consent, the CYP will self-complete an online (or paper) questionnaire about their mental and physical health. CYP towards the lower end of the 11-18 age band and CYP with special educational needs or disability may require the help of a carer to complete the questionnaire.

Outcomes

The questionnaires given to CYP include elements of the ISARIC Paediatric COVID-19 follow-up questionnaire\textsuperscript{13} and the recent Mental Health of Children and Young people in England surveys\textsuperscript{14} to facilitate international comparisons regarding the risk factors and profile of Long COVID in CYP. The questionnaires given to CYP were designed together with the International Severe Acute Respiratory and emerging Infection Consortium (ISARIC) Paediatric Working Group to produce a harmonised data collection tool, allowing for standardisation and metanalysis by design.

The elements taken from the ISARIC Paediatric COVID-19 follow-up questionnaire\textsuperscript{13} include questions about demographics and physical symptoms, with particular questions enquiring about cough and fever, given that these are the main acute symptoms in non-hospitalised CYP\textsuperscript{15} and gastrointestinal symptoms, as a study has shown such symptoms were common in seropositive CYP\textsuperscript{16}. It is likely other symptoms will manifest later in Long COVID (e.g., tiredness, headaches, myalgia etc.), with the skin rashes and cardiac problems in hospitalised children.
with Paediatric multisystem inflammatory syndrome being borne in mind for Long COVID\textsuperscript{5}. Those older CYP could very well overlap with adult symptomatology (e.g., a neuropsychiatric-encephalitic sub-group).

In terms of the other questions, the Strengths and Difficulties Questionnaire (SDQ)\textsuperscript{17} will be used to assess emotional and mental health, consisting of 25 items. The EQ-5D-Y\textsuperscript{18} will be used as a measure of quality of life/functioning, with fatigue measured using the 11-item Chalder Fatigue Questionnaire (CFQ)\textsuperscript{19}. Additionally, the short, 7-item version of The Warwick Edinburgh Mental Wellbeing Scale (WEMWBS)\textsuperscript{20} will be included. Loneliness will be measured using the 4-item UCLA Loneliness Scale for Children\textsuperscript{21} as well as self-reported school attendance.

The length of the online self-assessment has an estimated completion time of approximately 20 minutes for CYP. The CYP do not have to complete all of the questions at once but can pause and save the answers at any point.

**Statistical analyses**

Response rates for positive and negative CYP will inform expectations of final study size and possibly refinement of the sampling percentages. Descriptive statistics will be used for all dimensions collected at baseline and the follow-up questionnaires, separately by test status, wave and month of test. Appropriate data visualisation techniques will be used to examine trajectories over time for those variables that are measured repeatedly, separately by test status and month of test. Summary measures of the fatigue and mental health dimensions will be generated using dimension reduction methods such as latent variables/latent class/PCA methods. Cross-tabulation of latent variables against test status as well as discriminant analysis will help operationalise preliminary definitions of Long COVID. These will be compared with definitions based on categorisation of prolonged symptoms supplemented by a Delphi consensus process.

**Sample size**

A total of 5,000 participants (2,500 test-positive and 2,500 test-negative) would give 80% power to detect at least a 4% difference in symptoms at 5% significance, if the COVID-19 negative participants had a prevalence similar to that reported for adults (i.e., 34%). Greater baseline prevalences (e.g., 40%) would require slightly greater numbers for the same power and confidence; smaller baseline prevalences (e.g., 20%) smaller numbers. However, identifying risk factors for long COVID-19 would require studying demographic, social, family and clinical data and thus we would require larger numbers. For this reason, we are targeting the recruitment of 6,000 participants.
Patient and Public Involvement

Resources have been allocated for any Patient and Public Involvement (PPI) activities at INVOLVE rates including a nationally representative PPI research advisory group. PPI members will be offered training and support in accordance with the INVOLVE\textsuperscript{22} report and other guidance\textsuperscript{23} including online training. PPI meetings will take place quarterly.

Ethics and Dissemination

Ethical considerations

PHE has legal permission, provided by Regulation 3 of The Health Service (Control of Patient Information) Regulations 2002, to process patient confidential information for national surveillance of communicable diseases. Individual patient consent is not required for initial invitation to the study. Parents/carers and young people will be sent an invitation with a link to the website with the relevant information sheets and consent forms. They will have the opportunity to ask any questions about the study. Parents/carers of CYP under 16 years of age will be asked to complete a Parent/carer consent form. The young person will also be asked to sign an assent form to indicate their agreement. 16-18-year-olds will be asked for consent (using the Young Person Consent Form); but their parents will not. This is in line with HRA recommended processes\textsuperscript{24}.

Study approved by Yorkshire & The Humber - South Yorkshire Research Ethics Committee (REC reference: 21/YH/0060; IRAS project ID:293495).

It is possible that the questionnaires may make some vulnerable participants feel fatigued or distressed from completing questionnaires relating to their mental health and/or report serious symptoms that put them at immediate risk. The research team will provide information on where to seek support and provide self-help information. Unfortunately, the researchers are unable to provide medical advice. However, existing national surveys of children's mental health also follow this risk protocol.

Dissemination

Peer reviewed publications, briefings for policy makers and lay summaries for participants and CYP and carers. Results will be made public, initially on a pre-print server and we will allow reuse of articles under a CC BY licence. The dataset will be made publicly available.
Author contributions:
Terence Stephenson t.stephenson@ucl.ac.uk conceived the idea for the study, submitted the successful grant application and undertook the first revision of the manuscript.
Roz Shafran r.shafran@ucl.ac.uk provided ideas on mental health follow up to the original grant application, submitted the ethics and R&D applications and revised the manuscript.
Bianca De Stavola b.destavola@ucl.ac.uk provided statistical input to the design and analysis, including sample size calculations.
Natalia Rojas n.rojas@ucl.ac.uk wrote the first draft of the manuscript based on the grant application, protocol and ethics form and compiled the references.
Felicity Aiano Felicity.Aiano@phe.gov.uk helped develop the study design
Zahin Amin-Chowdhury Zahin.Amin@phe.gov.uk helped develop of the study design
Kelsey McOwat Kelsey.Mcowat@phe.gov.uk adapted the questionnaire for the online SNAP survey platform.
Maria Zavala Maria.Zavala@phe.gov.uk undertook a separate pilot study which informed the CLoCk consent process and the online questionnaire.
Ruth Simmons Ruth.Simmons@phe.gov.uk designed the participant sampling and dataflow.
Shamez Ladhani shamez.ladhani@phe.gov.uk operationalised the regulatory and recruitment ideas for the study and revised the manuscript.
Olivia Swann and Elizabeth Whittaker designed the elements of the ISARIC Paediatric COVID-19 follow-up questionnaire which were incorporated into the online questionnaire used in this study to which all the CLoCk Consortium members contributed.

All members of the CLoCk Consortium made substantial contributions to the conception or design of the work; were involved in drafting both the funding application and this manuscript; approved the version to be published; and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Co-Applicants on the grant application and CLoCk Consortium members (alphabetical):
Marta Buszewicz, University College London
Trudie Chalder, King’s College London, Chalder, trudie.chalder@kcl.ac.uk
Esther Crawley, University of Bristol, Esther.Crawley@bristol.ac.uk
Emma Dalrymple, PPI representative, emmajdalrymple@gmail.com
Tamsin Jane Ford, University of Cambridge, tjf52@medschl.cam.ac.uk
Shruti Garg, University of Manchester, Shruti.Garg@mft.nhs.uk
Anthony Harnden, University of Oxford, anthony.harnden@phc.ox.ac.uk
Dougal Hargreaves, Imperial College London, Dougal.HARGREAVES@education.gov.uk
Isobel Heyman, Great Ormond Street Hospital for Children NHS Foundation Trust Heyman, i.heyman@ucl.ac.uk
Michael Levin, Imperial College London, m.levin@imperial.ac.uk
Vanessa Poutie, University of Liverpool, v.poutie@liverpool.ac.uk
Terry Segal, University College London Hospitals NHS Foundation Trust, terry.segal@nhs.net
Malcolm Semple, University of Liverpool, M.G.Semple@liverpool.ac.uk
Kishan Sharma, Manchester University NHS Foundation Trust, Kishan.Sharma@mft.nhs.uk
Olivia Swann, University of Edinburgh, olivia.swann@ed.ac.uk
Elizabeth Whittaker, Imperial College London, e.whittaker@imperial.ac.uk

Acknowledgement: Michael Lattimore, Public Health England, as Project Officer for the CLoCk study

Competing interests statement: Terence Stephenson is Chair of the Health Research Authority and therefore recused himself from the Research Ethics Application

Funding statement: Funded by The Department of Health and Social Care, in their capacity as the National Institute for Health Research (NIHR), and by UK Research & Innovation (UKRI) who have awarded funding grant number COVLT0022. All research at Great Ormond Street Hospital NHS Foundation Trust and UCL Great Ormond Street Institute of Child Health is made possible by the NIHR Great Ormond Street Hospital Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.
References


Page 13 of 14

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml
Figure 1: Flowchart of the selection process.
**Figure 1**: Flowchart of the selection process.

- CYP (11-17yrs) Positive for SARS-CoV-2 infection 234,803
- CYP (11-17yrs) Negative for SARS-CoV-2 infection 1,245,288
- Excluded due to a pre / post positive SARS-CoV-2 infection result 80,365
- CYP (11-17yrs) Negative for SARS-CoV-2 infection 1,164,923
- Linked on age at test, gender, lower super output area (LSOA) and test Month
- Randomly Sampled Positives:
  - September 3611
  - October 2943
  - November 2970
  - December 2970
  - January 6012
  - February 2619
  - March 4356
- Randomly Sampled Negatives:
  - September 6633
  - October 5537
  - November 5914
  - December 5911
  - January 11,628
  - February 3109
  - March 5509
- Exclusions:
  - Died = 6
  - No link to DBS = 11,982
  - Previous study = 187
- CYP (11-17yrs) Positive for SARS-CoV-2 infection 102,402
- CYP (11-17yrs) Negative for SARS-CoV-2 infection 147,561
- Exclusions:
  - Died = 27
  - No link to DBS = 16,635
  - Previous study = 59
# Long Covid and the mental and physical health of children and young people: national matched cohort (CLoCk) study protocol

<table>
<thead>
<tr>
<th>Journal:</th>
<th>BMJ Open</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>bmjopen-2021-052838.R1</td>
</tr>
<tr>
<td>Article Type:</td>
<td>Protocol</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>09-Jul-2021</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Stephenson, Terence; UCL Great Ormond Street Institute of Child Health Population Policy and Practice  
Shafran, Roz; UCL Great Ormond Street Institute of Child Health Population Policy and Practice  
De Stavola, Bianca; UCL Great Ormond Street Institute of Child Health Population Policy and Practice  
Rojas, Natalia; UCL Great Ormond Street Institute of Child Health Population Policy and Practice  
Aiano, Felicity; Public Health England  
Amin-Chowdhury, Zahin; Public Health England  
McOwat, Kelsey; Public Health England  
Simmons, Ruth; Public Health England  
Zavala, Maria; Public Health England  
CLoCk, Consortium; UCL Great Ormond Street Institute of Child Health Population Policy and Practice  
Ladhani, Shamez N.; Public Health England |
| Primary Subject Heading: | Paediatrics |
| Secondary Subject Heading: | Public health, Infectious diseases |
| Keywords: | COVID-19, PAEDIATRICS, INFECTIOUS DISEASES, PUBLIC HEALTH, VIROLOGY |
I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author’s Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.
Long Covid and the mental and physical health of children and young people: national matched cohort (CLoCk) study protocol

Abstract 280 words. Main paper 2284 words.

Stephenson T¹, Shafran R¹, De Stavola B¹, Rojas N¹, Aiano F², Amin-Chowdhury Z², McOwat K², Simmons R², Zavala M², CLoCk Consortium¹, Ladhani S².

1 UCL Great Ormond Street Institute of Child Health
2 Public Health England, Colindale

Corresponding author: Professor Sir Terence Stephenson; UCL Great Ormond Street Institute of Child Health, 30 Guilford Street, London, WC1N 1EH; Email: t.stephenson@ucl.ac.uk; Telephone: 020 7242 9789; Fax Number: N/A

Five Keywords/phrases: Long COVID, children, young people, cohort study, physical and mental health
Abstract

Introduction: There is uncertainty surrounding the diagnosis, prevalence, phenotype, duration and treatment of Long COVID. This study aims to (a) describe the clinical phenotype of post-COVID symptomatology in children and young people (CYP) with laboratory-confirmed SARS-CoV-2 infection compared to test-negative controls, (b) produce an operational definition of Long COVID in CYP, and (c) establish its prevalence in CYP.

Methods and analysis: A cohort study of SARS-CoV-2 positive CYP aged 11-17 years compared with age-, sex- and geographically-matched SARS-CoV-2 test-negative CYP. CYP aged 11-17 testing positive and negative for SARS-CoV-2 infection will be identified and contacted 3, 6, 12 and 24 months after the test date. Consenting CYP will complete an online questionnaire. We initially planned to recruit 3,000 test positives and 3,000 test negatives but have since extended our target. Data visualisation techniques will be used to examine trajectories over time for symptoms and variables measured repeatedly, separately by original test status. Summary measures of fatigue and mental health dimensions will be generated using dimension reduction methods such as latent variables/latent class/PCA methods. Cross-tabulation of collected and derived variables against test status and discriminant analysis will help operationalise preliminary definitions of Long COVID.

Ethics and Dissemination: Research Ethics Committee Approval granted. Data will be stored in secure Public Health England (PHE) servers or University College London’s Data Safe Haven. Risks of harm will be minimised by providing information on where to seek support. Results will be published on a pre-print server followed by journal publication, with reuse of articles under a CC BY licence. Data will be published with protection against identification when small frequencies.

Study registration: Study registered with ISRCTN: Trial ID: ISRCTN34804192 and granted Urgent Public Health prioritisation.
Article Summary

STRENGTHS
- This study does not start with an arbitrary definition of a new condition – ‘opinion-based medicine’
- Rather it represents ‘evidence-based medicine’ - we will seek the views of children and young people on what they have experienced physically and mentally in the months following COVID-19
- The study incorporates a comparator matched cohort of children and young people who have experienced a pandemic, school closure and social isolation but who had a negative COVID test

WEAKNESSES
- The comparator matched cohort of children and young people must also have had a reason to seek a COVID test: symptoms, anxiety, a contact or a bereavement
- Ideally, we would test all children for antibodies to confirm COVID-19 positive and negative status throughout the study period but we thought would significantly reduce participation
Introduction

Currently, there is huge uncertainty surrounding the diagnosis, prevalence, phenotype, duration and treatment of Long COVID. There is currently no diagnostic test or code for Long COVID. Consequently, cases of Long COVID are not captured in routine National Health Service (NHS) administrative datasets. Instead, Long COVID may be coded as a variety of different conditions in hospital or symptom clusters in non-hospitalised children and young people (CYP). However, it is possible the symptoms associated with Long COVID are in fact a mixture of factors relating to the pandemic and lockdown as a whole rather than the viral infection itself. For example, factors such as social isolation, anxiety, depression or educational concerns may be the root cause of these symptoms in CYP both with and without SARS-CoV-2 infection. The effects on the developing brain and behaviour of adolescents could be far reaching\(^1\).

Despite acute COVID-19 illness being milder in CYP, it should not be assumed that those CYP at low risk of life-threatening acute infections\(^2\)\(^3\) do not suffer the longer-term consequences of SARS-CoV-2 infection. It is important to study Long COVID within this population given that the psychological and social impact of Long COVID could have major consequences for transition to adulthood. There is a clear need to define the clinical phenotype of Long COVID in order to understand those most at risk, the illness trajectory and to provide accurate information on the natural course of the condition. It is equally important to understand both the physical and mental health impact on CYP with Long COVID, given the rise in mental health problems among CYP since the start of the pandemic\(^4\). Obtaining a better understanding of which CYP are affected will help with targeting potential interventions.

The greatest risk factor for severe acute SAR-CoV-2 illness has been old age, with other risk factors at all ages including obesity, comorbid long-term conditions, learning and neurological disabilities, mental health problems and ethnic minority status. It is plausible that such CYP may be most at risk of suffering from Long COVID, given that much is still unknown in CYP about the immunological susceptibility and underlying biology of Long COVID. Of all CYP, those likely to be most at risk of Long COVID are teenagers, with existing literature showing they make up the majority of CYP with chronic fatigue, post viral syndromes and persistent symptoms\(^5\).

In adults, there is emerging evidence that gender is a risk factor for Long COVID, with middle-aged females more susceptible than men\(^6\). In terms of symptoms, two main groupings have been identified (i) respiratory symptoms (e.g., cough, shortness of breath) as well as fatigue and headaches, and (ii) multi-system, including the brain, gut and heart. Among adults with Long COVID, heart symptoms such as palpitations or fast heartbeat, pins and needles or numbness and problems concentrating (‘brain fog’) have been reported. Experiencing a greater number of symptoms during the first week of infection, as well as older age, have also been linked to Long COVID. It should be noted mental health conditions were not reported in this cohort. Four syndromes have been described in adults: Post-viral fatigue; Fluctuating multi-system symptoms; Post-intensive care syndrome; and Lasting organ damage. It is known in both adults and young people that a wide range of long-term physical conditions increase the risk of mental ill-health, particularly if the condition involves the central nervous system, with several studies...
reporting increased rates of all common mental health conditions in children. Moreover, fatigue has also been studied in paediatric long-term conditions including CYP with multiple sclerosis. Emerging clinical observations and preliminary research indicate that COVID-19 can have a long-term impact on CYP in a range of domains but well conducted, methodologically robust studies are lacking.

This study primarily aims to (a) describe the clinical phenotype and prevalence of post-COVID symptoms (e.g., pain or physical symptoms, fatigue, sleep problems, mental health problems) among test positive and test negative CYP, (b) produce an operational definition of Long COVID in CYP, and (c) establish the prevalence of Long COVID in CYP testing positive for SARS-CoV-2 infection. The second aim is of particular importance given that this is a prerequisite for any future epidemiological or interventional study. This will be achieved by using the empirical data we will acquire from 17,000 CYP and a national Delphi consensus process.

Methods

Study design and setting

Public Health England (PHE) has been conducting national surveillance of SARS-CoV-2 since the start of the pandemic in England. PHE receives daily electronic notifications of all SARS-CoV-2 RT-PCR tests performed in healthcare settings (Pillar 1 tests) and in the community (Pillar 2 tests), which are reported through the Second Generation Surveillance System (SGSS). Information within the SGSS reports includes NHS number, name, age, sex, postcode, date of sample, reporting laboratory and test result. PHE also has access to the electronic Patient Demographic Service (PDS), which contains the names, addresses and status (alive/dead) of all patients registered with the NHS.

A sample of CYP aged 11-17 years old when testing positive for SARS-CoV-2 between September 2020 and March 2021 and COVID test-negative CYP matched for age, sex and region identified through SGSS will be linked to PDS using available identifiers and postal addresses. A letter will be posted to them, informing them about the study and inviting them to take part using an online link. This link will provide them with information about the study, with an option to consent online and complete a short recruitment questionnaire (or both as paper options). Recruitment was planned to start in March 2021 and all data collection anticipated to end in April 2023.

This will be a cohort study of SARS-CoV-2 test-positive CYP aged 11-17 years matched on the test date, age, sex and geography to SARS-CoV-2 test-negative controls identified by PHE.

Participants

Between September 2020 and March 2021, 234,803 CYP aged between 11 and 17 years were diagnosed with COVID-19 in England. During the same period there were 1,481,154 negatives tests among this age group equivalent to 1,203,996 CYP. Among those who tested negative
76,689 individuals (100,154 tests) were excluded as they had a positive result pre and/or post their negative test. The positive and negative cohorts were linked using age at test, gender, month of test, and lower super output area resulting in 102,402 positive individuals and 147,561 negative individuals. Matched individuals were checked against the NHS spine to exclude any individuals who have died, and to extract participants addresses to send invitation letters. 37 individuals were excluded because they had died since their COVID-19 test (6 among those positive and 31 among those negative), and 11,193 individuals of the positive cohort and 19,251 individuals of the negative cohort were excluded as an address was not available. Finally, 246 CYP were excluded as they were included in a previous study. Following exclusions 91,016 positive CYP and 128,220 negative CYP will be available to be invited to participate in the study. For all the months except December, all eligible individuals will be sent an invitation. For those tested in December, 2,970 positive CYP and 5,911 negative CYP will be invited as the eligible number was so large that additional funding would be needed to invite all those eligible.

This is shown schematically in Figure 1.

Data collection

Participating families of CYP will be contacted 3, 6, 12 and 24 months after the CYP’s SARS-CoV-2 test (depending on recruitment month) and invited to take part in the study. Following online informed consent, the CYP will self-complete an online (or paper) questionnaire about their mental and physical health. CYP towards the lower end of the 11-17 age band and CYP with special educational needs or disability may require the help of a carer to complete the questionnaire.

Outcomes

The questionnaires given to CYP include elements of the ISARIC Paediatric COVID-19 follow-up questionnaire\(^\text{12}\) and the recent Mental Health of Children and Young people in England surveys\(^\text{13}\) to facilitate international comparisons regarding the risk factors and profile of Long COVID in CYP. The questionnaires given to CYP were designed together with the International Severe Acute Respiratory and emerging Infection Consortium (ISARIC) Paediatric Working Group to produce a harmonised data collection tool, allowing for standardisation and metanalysis by design.

The elements taken from the ISARIC Paediatric COVID-19 follow-up questionnaire\(^\text{12}\) include questions about demographics and physical symptoms, with particular questions enquiring about cough and fever, given that these are the main acute symptoms in non-hospitalised CYP\(^\text{14}\) and gastrointestinal symptoms, as a study has shown such symptoms were common in seropositive CYP\(^\text{15}\). It is likely other symptoms will manifest later in Long COVID (e.g., tiredness, headaches, myalgia etc.), with the skin rashes and cardiac problems in hospitalised children with paediatric multisystem inflammatory syndrome being borne in mind for Long COVID\(^\text{5}\). Those older CYP could very well overlap with adult symptomatology (e.g., a neuropsychiatric-encephalitic sub-group).
In terms of the other questions, the Strengths and Difficulties Questionnaire (SDQ)\(^6\) will be used to assess emotional and mental health, consisting of 25 items. The EQ-5D-Y\(^7\) will be used as a measure of quality of life/functioning, with fatigue measured using the 11-item Chalder Fatigue Questionnaire (CFQ)\(^8\). Additionally, the short, 7-item version of The Warwick Edinburgh Mental Wellbeing Scale (WEMWBS)\(^9\) will be included. Loneliness will be measured using the 4-item UCLA Loneliness Scale for Children\(^10\) as well as self-reported school attendance.

The length of the online self-assessment has an estimated completion time of approximately 20 minutes for CYP. The CYP do not have to complete all of the questions at once but can pause and save the answers at any point.

**Statistical analyses**

Response rates for positive and negative CYP will inform expectations of final study size and possibly refinement of the sampling percentages. Descriptive statistics will be used for all dimensions collected at baseline and the follow-up questionnaires, separately by test status, wave and month of test. Appropriate data visualisation techniques will be used to examine trajectories over time for those variables that are measured repeatedly, separately by test status and month of test. Summary measures of the fatigue and mental health dimensions will be generated using dimension reduction methods such as latent variables/latent class/PCA methods. Cross-tabulation of latent variables against test status as well as discriminant analysis will help operationalise preliminary definitions of Long COVID. These will be compared with definitions based on categorisation of prolonged symptoms supplemented by a Delphi consensus process.

**Sample size**

A total of 5,000 participants (2,500 test-positive and 2,500 test-negative) would give 80% power to detect at least a 4% difference in symptoms at 5% significance, if the COVID-19 negative participants had a prevalence similar to that reported for adults (i.e., 34%). Greater baseline prevalences (e.g., 40%) would require slightly greater numbers for the same power and confidence; smaller baseline prevalences (e.g., 20%) smaller numbers. However, identifying risk factors for long COVID-19 would require studying demographic, social, family and clinical data and thus we would require larger numbers. For this reason, we are planning to invite all available participants.

**Patient and Public Involvement**

Resources have been allocated for any Patient and Public Involvement (PPI) activities at INVOLVE rates including a nationally representative PPI research advisory group. PPI members will be offered training and support in accordance with the INVOLVE\(^{21}\) report and other guidance\(^{22}\) including online training. PPI meetings will take place quarterly.
Ethics and Dissemination

**Ethical considerations**

PHE has legal permission, provided by Regulation 3 of The Health Service (Control of Patient Information) Regulations 2002, to process patient confidential information for national surveillance of communicable diseases. Individual patient consent is not required for initial invitation to the study. Parents/carers and young people will be sent an invitation with a link to the website with the relevant information sheets and consent forms. They will have the opportunity to ask any questions about the study. Parents/carers of CYP under 16 years of age will be asked to complete a Parent/carer consent form. The young person will also be asked to sign an assent form to indicate their agreement. 16-17-year-olds will be asked for consent (using the Young Person Consent Form); but their parents will not. This is in line with Health Research Authority recommended processes\(^2^3\).

Study approved by Yorkshire & The Humber - South Yorkshire Research Ethics Committee (REC reference: 21/YH/0060; IRAS project ID:293495).

It is possible that the questionnaires may make some vulnerable participants feel fatigued or distressed from completing questionnaires relating to their mental health and/or report serious symptoms that put them at immediate risk. The research team will provide information on where to seek support and provide self-help information. Unfortunately, the researchers are unable to provide medical advice. However, existing national surveys of children's mental health also follow this risk protocol.

**Dissemination**

Peer reviewed publications, briefings for policy makers and lay summaries for participants and CYP and carers. Results will be made public, initially on a pre-print server and we will allow reuse of articles under a CC BY licence. The dataset will be made publicly available.
Figure 1: Flowchart of the selection process.
Author contributions:
Terence Stephenson t.stephenson@ucl.ac.uk conceived the idea for the study, submitted the
successful grant application and undertook the first revision of the manuscript.
Roz Shafran r.shafran@ucl.ac.uk provided ideas on mental health follow up to the original grant
application, submitted the ethics and R&D applications and revised the manuscript.
Bianca De Stavola b.destavola@ucl.ac.uk provided statistical input to the design and analysis,
including sample size calculations.
Natalia Rojas n.rojas@ucl.ac.uk wrote the first draft of the manuscript based on the grant
application, protocol and ethics form and compiled the references.
Felicity Aiano Felicity.Aiano@phe.gov.uk helped develop the study design.
Zahin Amin-Chowdhury Zahin.Amin@phe.gov.uk helped develop the study design.
Kelsey McOwat Kelsey.Mcowat@phe.gov.uk adapted the questionnaire for the online SNAP
survey platform.
Maria Zavala Maria.Zavala@phe.gov.uk undertook a separate pilot study which informed the
CLoCk consent process and the online questionnaire.
Ruth Simmons Ruth.Simmons@phe.gov.uk designed the participant sampling and dataflow.
Shamez Ladhani shamez.ladhani@phe.gov.uk operationalised the regulatory and recruitment
ideas for the study and revised the manuscript.

Olivia Swann and Elizabeth Whittaker designed the elements of the ISARIC Paediatric COVID-
19 follow-up questionnaire which were incorporated into the online questionnaire used in this
study to which all the CLoCk Consortium members contributed.

All members of the CLoCk Consortium made substantial contributions to the conception or
design of the work; were involved in drafting both the funding application and this manuscript;
approved the version to be published; and agree to be accountable for all aspects of the work in
ensuring that questions related to the accuracy or integrity of any part of the work are
appropriately investigated and resolved.

Co-Applicants on the grant application and CLoCk Consortium members (alphabetical):
Marta Buszewicz, University College London
Trudie Chalder, King’s College London, trudie.chalder@kcl.ac.uk
Esther Crawley, University of Bristol, Esther.Crawley@bristol.ac.uk
Emma Dalrymple, PPI representative, emmajdalrymple@gmail.com
Tamsin Jane Ford, University of Cambridge, tjf52@medschl.cam.ac.uk
Shruti Garg, University of Manchester, Shruti.Garg@mft.nhs.uk
Anthony Harnden, University of Oxford, anthony.harnden@phc.ox.ac.uk
Dougal Hargreaves, Imperial College London, Dougal.HARGREAVES@education.gov.uk
Isobel Heyman, Great Ormond Street Hospital for Children NHS Foundation Trust, i.heyman@ucl.ac.uk
Michael Levin, Imperial College London, m.levin@imperial.ac.uk
Vanessa Poustie, University of Liverpool, v.poustie@liverpool.ac.uk
Terry Segal, University College London Hospitals NHS Foundation Trust, terry.segal@nhs.net
Malcolm Semple, University of Liverpool, M.G.Semple@liverpool.ac.uk
Acknowledgement: Michael Lattimore, Public Health England, as Project Officer for the CLoCk study

Competing interests statement: Terence Stephenson is Chair of the Health Research Authority and therefore recused himself from the Research Ethics Application

Funding statement: Funded by The Department of Health and Social Care, in their capacity as the National Institute for Health Research (NIHR), and by UK Research & Innovation (UKRI) who have awarded funding grant number COVLT0022. All research at Great Ormond Street Hospital NHS Foundation Trust and UCL Great Ormond Street Institute of Child Health is made possible by the NIHR Great Ormond Street Hospital Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health

References


Figure 1: Flowchart of the selection process.

CYP (11-17yrs) first Positive for COVID
- Tests = 234,803
- Individuals = 234,803

CYP (11-17yrs) Negative for COVID
- Tests = 1,481,154
- Individuals = 1,203,996

Excluded due to a pre/post positive COVID result
- Tests = 100,154
- Individuals = 76,689

CYP (11-17yrs) Negative for COVID
- Tests = 1,381,000
- Individuals = 1,127,307

Linked on age at test, gender, lower super output area (LSOA) and Month of test

CYP (11-17yrs) Positive for COVID
- Tests and Individuals = 102,402

CYP (11-17yrs) Negative for COVID
- Tests and Individuals = 147,561

Exclusions: N=11,386
- Died = 6
- Address not available = 11,193
- Previous study = 187

Positives: N=91,016
- September 3611
- October 11,134
- November 18,295
- December 34,928
(2970 invited)
- January 16,074
- February 2619
- March 4355

Negatives: N=128,220
- September 6777
- October 15,277
- November 23,463
- December 54,905
(5911 invited)
- January 19,107
- February 3185
- March 5506