COVID-19 Antibody Testing in the National Child Development Study, 1970 British Cohort Study, Next Steps and Millennium Cohort Study

User Guide (Version 1)

July 2021

CENTRE FOR LONGITUDINAL STUDIES

MRC Unit for Lifelong Health and Ageing

UKRI Economic and Social Research Council

Contact

Data queries: help@ukdataservice.ac.uk

Authors

Matt Brown, Andrew Peters, George B. Ploubidis, Aida Sanchez, Richard J. Silverwood.

How to cite this guide

Brown, M., Peters, A., Ploubidis, G.B., Sanchez, A., Silverwood, R.J. (2021) *COVID-19 Antibody Testing in the National Child Development Study, 1970 British Cohort Study, Next Steps and Millennium Cohort Study: User Guide (Version 1)*. London: UCL Centre for Longitudinal Studies

This guide was published in July 2021 by the UCL Centre for Longitudinal Studies (CLS).

The UCL Centre for Longitudinal Studies (CLS) is an Economic and Social Research Council (ESRC) Resource Centre based at the UCL Social Research Institute, University College London. It manages four internationally-renowned cohort studies: the 1958 National Child Development Study, the 1970 British Cohort Study, Next Steps, and the Millennium Cohort Study. For more information, visit www.cls.ucl.ac.uk.

This document is available in alternative formats. Please contact the Centre for Longitudinal Studies. For questions and feedback about this user guide:

tel: +44 (0)20 7612 6875
email: clsfeedback@ucl.ac.uk

Contents

1. Introduction

1.1 Background

Between May 2020 and March 2021 the Centre for Longitudinal Studies (CLS) carried out three surveys of the participants of four national longitudinal cohort studies which collected insights into the lives of study participants including their physical and mental health and wellbeing, family and relationships, education, work, and finances during the coronavirus pandemic.

The four studies run by CLS are:

- Millennium Cohort Study (born 2000-02) both cohort members and parents (MCS),
- Next Steps (born 1989-90) (NS),
- 1970 British Cohort Study (BCS70),
- 1958 National Child Development Study (NCDS)

In addition, the surveys were also completed by participants of the MRC National Survey of Health and Development (NSHD, 1946 British birth cohort) run by the MRC Unit for Lifelong Health and Ageing at UCL (MRC LHA). These studies have been following large nationally representative groups of people since birth, and their ages currently range from 19 through to 74.The COVID-19 Surveys are described on the CLS website: https://cls.ucl.ac.uk/covid-19-survey

The Centre for Longitudinal Studies is funded by the Economic and Social Research Council.

1.2 COVID-19 antibody testing

In March 2021, CLS study members who had participated in any of the three COVID-19 Surveys were invited to provide a finger-prick blood sample to be analysed for COVID-19 antibodies. Those who agreed were sent a blood sample collection kit and were asked to post back the sample to a laboratory for analysis.

This study was funded by the Longitudinal Health and Wealth (LH&W) National Core Studies (NCS) for SARS-CoV-2 research which was established by the UK Chief Scientific Advisor. Antibody tests were simultaneously conducted amongst participants from the MRC National Surey of Health and Development (NSHD) and multiple other UK longitudinal studies (Avon Longitudinal Study of Parents and Children (ALSPAC), Understanding Society, English Longitudinal Study of Ageing (ELSA),Twins UK, Extended Cohort for E-health, Environment and DNA (EXCEED) and Southall and Brent Revisited (SABRE). The test results provide evidence of the seroprevalence of COVID-19 antibodies amongst study members and in combination with the data collected in the COVID-19 surveys will be used to identify COVID-19 cases.

The data collection made use of the logistics framework put in place for the Government's at-home antibody testing programme, administered through Thriva, a company who conduct health assessments. The process was approved by the Department of Health and Social Care (DHSC).

CLS is funded by the Economic and Social Research Council.

2. Protocol for data collection

2.1 Eligibility

All participants of the CLS cohorts (National Child Development Study, 1970 British Cohort Study, Next Steps and Millennium Cohort Study) who had taken part in at least one of the three COVID-19 Surveys conducted between May 2020 and March 2021 were invited to participate in the antibody testing project unless:

- They had subsequently opted out of participating in future COVID-19 surveys
- They had fully withdrawn from participating in the cohort studies
- They lived outside of the UK

In the case of the Millennium Cohort Study, parents of cohort members had been invited to take part in the COVID-19 surveys and those that did so were also invited to take part in the antibody testing project.

During the recruitment phase participants were told not to take part if they had a medical condition which means they are at increased risk of bleeding or were currently taking any anti-coagulant medication.

2.2 Recruitment

Eligible participants were emailed by Kantar Public (the survey agency who conducted the COVID-19 Surveys on our behalf) and invited to take part in the antibody testing project. The email invitation included a link to an information sheet which was hosted on the study websites and a detailed set of FAQs which provided full information about the purpose of the study and what taking part would involve. Those who were willing to take part were asked to complete a short web survey in which they recorded their consent to take part and provide the address to which the blood sample collection kit would be posted. Participants consented to take part in the antibody testing project between 22nd March and 5th April.

2.3 Questionnaire

The short web survey also asked those who consented to take part some questions about whether they had been vaccinated and if so the number, dates and type of jab received. The questions are shown in Appendix 1.

2.3 Sample collection

The sample collection process was conducted by Thriva. Blood sample collection kits were collated and sent out to consenting participants. The kits contained full instructions on how to collect the sample and:

- One collection tube
- Three safety retractable lancets
- One sterile wipe
- One cleaning wipe
- Two small plasters
- One mailing return envelope
- One tube label (on return form)
- One clear leak-proof plastic bag
- One cardboard box
- One return form

A link to a video demonstrating how to collect the sample was also provided.

In summary, participants used the lancet to fill a blood collection tube with a ~0.5ml of blood. Once the tube was filled participants were asked to label the tube with a unique bar-coded sticker and return it in a leak proof bag, with absorbent material, in the post. Samples were posted to Pura Diagnostics for analysis, a laboratory which is accredited in accordance with ISO/IEC 17025:2017 and is registered with the Care Quality Commission.

2.4 Antibody testing

Two antibody tests were conducted:

1) Roche Elecsys® Anti-SARS-CoV-2 – an immunoassay qualitative detection of antibodies against SARS-CoV-2 nucleocapsid (N) protein. A positive result is likely to identify exposure to COVID-19. The test is qualitative and provides a binary positive or negative result.

   A full description of the assay can be found here:

   https://diagnostics.roche.com/global/en/products/params/elecsys-anti-sars-cov-2.html

2) Roche Elecsys® Anti-SARS-CoV-2 S – an immunoassay which gives a semi-quantitative assessment of antibodies to the SARS-CoV-2 spike (S) protein. A positive result here may be more likely to result from COVID-19 vaccination.

   A full description of the assay can be found here:

   https://diagnostics.roche.com/gb/en/products/params/elecsys-anti-sars-cov-2-s.html

Further information about the test results included in the data is included in Section 4.

Blood samples were returned to the laboratory between 21st April and 2nd July. Samples were analysed within 72 hours of receipt at the laboratory.

2.5 Results and duty of care

Thriva provided results files on a weekly basis and participants were informed of their results via email. For the purposes of reporting to participants the results of the two assays were combined into a single result – positive,

negative or inconclusive. Participants were provided with guidance on how to interpret the result and encouragement to continue to follow Government advice.

3. Participation

Table 1 below shows the number of participants per cohort who were invited to take part in the antibody testing project, the number who consented and the blood sample return rates.

Table 1: Consent rates and blood sample return rates by cohort

	Invited	Consented	Consent rate (%)	Blood sample returned	% of consented	% of invited
NCDS	6939	4156	60%	3222	78%	46%
BCS70	6594	3741	57%	2547	68%	39%
Next Steps	4826	2090	43%	1267	61%	26%
MCS CM	5266	1397	27%	1140	82%	22%
MCS Parents	7143	3214	45%	2266	71%	32%

4. Research Data

4.1 Licencing and data access

The dataset is available to download from the UK Data Service (UKDS).

All users of the data need to be registered with the UKDS. Details of how to do this are available at https://www.ukdataservice.ac.uk/get-data/how-to-access/registration.

The COVID-19 serology data have been supplied to the UKDS under End User Licence (EUL) for the CLS studies (NCDS, BCS70, NS and MCS).

The data from the four CLS cohort studies are included in the same dataset. This dataset can be downloaded once the EUL access conditions have been accepted by the user.

The CLS data available under EUL exclude detailed data that present a potential risk for disclosivity. Please refer to section 5.8 for information on how these data have been de-identified for inclusion under EUL. This applied to:

1) MCS data for families containing triplets.
2) Exact dates of COVID-19 vaccination

These potentially identifiable CLS data can be accessed securely by applying directly to the CLS Data Access Committee.

The EUL data includes cases who provided consent to take part in the antibody testing project and answered the web survey questions about vaccination but did not return a blood sample.

Study	**Cases in dataset**
NCDS	4157
BCS70	3741
Next Steps	2091
MCS CMs**	1757 (1755)

Study	Cases in dataset
MCS Parents**	3214 (3212)

* The EUL data excludes 2 triplet families (numbers in brackets)

4.2 Identifiers

Individual identifiers

All four CLS-based cohort studies are included in the same dataset, each with their standard research IDs that allow them to be linked to the other study data available at the UKDS.

For NCDS, BCS70 and Next Steps, the data for each cohort member is displayed with one case per row.

MCS data are displayed in long format, where MCSID identifies each family, and an individual identifier identifies each family member: CS_CNUM00 for cohort member and CS_PNUM00 for parent. Therefore, for families with several respondents there will be several rows per family (MCDSID), but one row per family member (CS_CNUM00 / CS_PNUM00). This is the same format as other MCS data deposits at UKDS.

Cohort identifier

Variable CS_COHORT allows the identification of the data by cohort study, and for MCS whether it is the CM or parent respondent. It is set as follows:

1 = NCDS

2 = BCS70

3 = Next Steps

4 = MCS Cohort Member (CM)

5 = MCS Parent

4.3 Use of individual identifiers to merge with cohort study data

For NCDS, BCS70, Next Steps and MCS, the data are identified with the same research IDs used for the rest of cohort data and COVID-19 survey data available at the UKDS. This enables the data to be easily merged with all current waves of COVID-19 data.

While merging COVID-19 serology data with cohort study data should be straightforward using their respective identifiers, users should consult individual user guides for specific information beforehand.

For MCS, researchers need to use both the MCS family identifier (MCSID) and the two individual person identifiers (CNUM00/PNUM00) to merge on with other cohort data. As CNUM00 and PNUM00 include the “CS_” prefix they may need consistent naming across datasets beforehand depending on the method of merging used.

There are different ways the data of MCS can be merged depending on the focus of the research project (Parent/Carers, Cohort Members or family). Details, syntax and examples on merging is provided by the MCS Data Handling Guide.

4.4 Variable order

The order in which variables appear in the dataset is:

- IDs for each cohort
- Cohort study
- Date kits were processed in the lab
- Antibody test results
- Sample collection questions and related data

The included data variables are as follows:

Variable Name	**Variable Label**
CS_PROCESSDATEM	Month in 2021 of antibody test
CS_COVN_RESULT	N assay: result classification

Variable Name	Variable Label
CS_COVS_VALUE	S assay: numeric result of the test (U/mL)
CS_COVS_RESULT	S assay: result classification
CS_COV_RESULT	(Derived) Combined result of two antibody tests
CS_SURDATM	Month in 2021 consent web survey completed
CS_BEENVAC	Whether vaccinated for COVID-19 at time of consent web survey
CS_NUMJAB	Number of jabs received at time of consent web survey
CS_VACTYP	Type of vaccine received
CS_VACDAT1	Date of first vaccine (Month Grouped)
CS_VACDAT2	Date of second vaccine (Month Grouped)
CS_VACFLAG	Respondent entered vaccine date later than survey date
CS_VVTIMED	(Derived) Number of days between first vaccine and second vaccine
CS_V1PTIMED	(Derived) Number of days between first vaccine and date of antibody tests
CS_V2PTIMED	(Derived) Number of days between second vaccine and date of antibody tests

4.5 Test results

CS_COVN_RESULT

This variable gives the result of the N: assay – positive, negative or void. This is a qualitative assay and as such no continuous result is available for this test.

CS_COVS_VALUE

This variable gives a numeric result of the S: assay in (U/mL). The test has an upper threshold of 250 U/ml.

Note: A small number of participants have a result of over 250 U/ml. This is because the laboratory conducted a pilot study where some of the samples that hit a value of 250 u/ml were run diluted 1:100 to assess whether this would allow higher values to be accurately identified.

CS_COVS_RESULT

This variable classifies the result of the S: assay – positive, negative or void. Values of higher than 0.8 U/ml are classified as positive.

CS_COV_RESULT

This variable combines the result of the N: assay and S: assay. Participants classified as positive on either the N: assay or S: assay are classified as positive. Participants classified as negative on both the N: assay and the S: assay are classified as negative. Participants who had one negative result and one void result are classified as void.

4.6 Variable names

Non-identifier variables have been named with a preceding “CS_” in the same manner as COVID-19 survey variables are named to include the wave number in order to distinguish them as Serology variables.

4.7 Variable description

Variable labels

For result details regarding specific antibody tests, variable labels are preceded by either “N assay” or “S assay” to aid identification. Variable labels based on questions are based on the wording from the sample collection survey with additional details or modifications where necessary for context.

Value labels

The N: assay and S: assay result classification variables are given are given ‘Positive’, ‘Negative’ or ‘Void’ values dependent on the level of antibodies detected.

Values for sample collection data typically match the question responses. CS_VACDAT1 and CS_VACDAT2 are the exceptions as dates have been banded (see 4.9 for more information on this).

4.8 Missing values

Antibody result data

Missing data on the antibody test result variables (CS_COVN_RESULT, CS_COVS_RESULT, and CS_COV_RESULT) are identified as follows:

-1 = Not applicable

-8 = Void

Not applicable (-1) indicates that test results were never received for the selected case.

Void (-8) results occur when the returned blood samples could not be analysed and no valid results for that measure could be obtained.

Vaccination questions

Missing data on the vaccination questions are consistently labelled as follows:

-1 = Not applicable

-8 = Don’t know

-9 = Don’t want to answer

Not applicable (-1) occurs when the specific question does not apply.

4.9 Coding of disclosive information

Vaccine dates

Vaccine dates have been grouped due to the potential extra level of disclosivity, as there were very few cases in December of 2020 and from April 2021 onwards:

CS_VACDAT1:

1. December 2020 – January 2021
2. February 2021
3. March 2021 or later

CS_VACDAT2:

1. December 2020 – February 2021
2. March 2021 or later

As the efficacy of the vaccine changes over time, relevant information in this regard is available to users in the bands provided. In general, cases who had vaccines before March 2021 will all show results when the vaccines effects are at their highest. Those after this time may or may not be, depending on the type of vaccine and when the sample was taken.

A number of cases input a first or second vaccine date later than the survey date. It is likely these cases had their vaccines booked for these times, though the survey design only intended for completed vaccinations. As a result, a flag (CS_VACFLAG) is included with the data to allow users to decide whether to use these data. The vaccine dates of cases identified with this flag do all still precede the dates their respective kits were returned.

As the time between vaccines and samples is very relevant to result data, the number of days between the following events has been derived:

CS_VVTIMED = (Derived) Number of days between the first and second vaccine

CS_V1PTIMED = (Derived) Number of days between first vaccine and date of antibody tests

CS_V2PTIMED = (Derived) Number of days between second vaccine and antibody tests

Vaccine types

Very few cases in the sample reported having the Moderna vaccine (CS_VACTYP). As such the ‘Something else’ category has been expanded to cover the Moderna vaccine.

MCS triplets

The MCS families containing triplets have been removed from the data. This affects 2 cohort members and 2 parents.

4.10 Inconsistencies / errors

Gap between vaccine date and laboratory analysis

The dataset contains the vaccination date self-reported by the participants on the online consent questionnaire, as well as the date when the antibody tests were conducted at the lab.

Participants completed the online questionnaire over a two week period from 22nd March and returned their blood samples from 21st April until 2nd July. Therefore, there is a gap of at least several weeks between the vaccine questions being answered, and the blood samples being returned to the lab. The UK vaccine rollout means that many respondents could have received first and/or second jabs within this time frame. Data users should account for this when using vaccine information in analysis.

Incorrect year of vaccination

Several vaccine dates have been amended in the data in cases where it was apparent participants entered the year incorrectly. For cases where real dates could not be established from dates before December 2020 or after the survey date values have been set to -8 (Don’t know). No other amendments have been made to vaccine dates including in cases where respondents entered vaccine dates very close together or very far apart. Users should look at CS_VACDAYS, which is the number of days between first and second vaccine, if they wish to make their own assessment on the validity of such responses.

Date of vaccination later than kit returns

While CS_VACFLAG has been used to identify cases whose vaccination date was later than the survey date, a very small number of cases input dates which followed the date their kit was returned. The related vaccine information for these cases has been set to -8 (Don’t know).

5. Weighting

Non-response is common in longitudinal surveys. Missing values mean less efficient estimates because of the reduced size of the analysis sample, but also introduce the potential for bias since respondents are often systematically different from non-respondents. To support researchers in producing robust analysis, we have developed comprehensive advice on how to deal with missing data (1). The approaches we recommend to researchers capitalise on the rich data cohort members provided over the years before their nonresponse. These include well known methods such as Multiple Imputation (MI), Inverse Probability Weighting (IPW), and Full Information Maximum Likelihood (FIML).

To correct for non-response in the COVID-19 Wave 1, 2 and 3 surveys and facilitate analysis in all cohorts, non-response weights have been provided, so that IPW analysis can be undertaken, either in isolation or in combination with MI (2).

Because participation in the antibody testing project does not correspond exactly to participation in the COVID-19 surveys (see section 3), antibody testing project-specific non-response weights could be derived. Such non-response weights for each cohort will follow at later date. In the meantime, users may wish to follow the procedure outlined for the COVID-19 surveys (2) to derive their own weights or, alternatively, to use the existing COVID-19 survey non-response weights as an approximation. As antibody testing project participants may have participated at any one or more of the COVID-19 survey waves (and not participated at the remaining waves; see section 2), one approach that would ensure a non-missing weight for each antibody testing project participant would be to take the mean of the non-missing COVID-19 survey non-response weights across Waves 1, 2 and 3.

1. Silverwood R, Narayanan M, Dodgeon B, Ploubidis G. 2021. *Handling missing data in the National Child Development Study: User Guide (Version 2)*. London: UCL Centre for Longitudinal Studies.

2. Brown, M., Goodman, A., Peters, A., Ploubidis, G.B., Sanchez, A., Silverwood, R., Smith, K. (2021) *COVID-19 Survey in Five National Longitudinal Studies: Waves 1, 2 and 3 User Guide (Version 3)*. London: UCL Centre for Longitudinal Studies and MRC Unit for Lifelong Health and Ageing.

Appendix 1: Vaccination questions

BEENVAC {ASK IF CONSENT 1 = 1}

It would help be helpful to know whether you have been vaccinated for COVID-19.

Have you been vaccinated for COVID-19?

Yes

No

-8 Don’t Know

-9. Don’t want to answer

NUMJAB {ASK IF BEENVAC = 1}

How many jabs have you received?

One

Two

-8 Don’t Know

-9 Don’t want to answer

VACDAT1 {ASK IF BEENVAC = 1}

Please enter the date you had your {IF NUMJAB = 2: ‘first’} jab.

CALENDAR ENTRY

VACDAT2 {ASK IF NUMJAB = 2}

Please enter the date you had your second jab.

CALENDAR ENTRY

VACTYP {ASK IF BEENVAC = 1}

Which vaccine did you receive?

1. Oxford/AstraZeneca

2. Pfizer/BioNTech

3. Moderna

4. Janssen (Johnson & Johnson)

5. Something else

-8 Don’t Know

-9 Don't want to answer