**Statistical analysis plan**

Determinants and impacts of under-vaccination against SARS-CoV-2 in the United Kingdom – a pooled meta-analysis of data from 4 UK nations.

# Analysts

**England:**  Genevieve Cezard, Angela Wood

**Northern Ireland:** Declan Bradley, Lynsey Patterson, Siobhán Murphy

**Scotland:**  Chris Robertson, Steven Kerr, Kirsty Morrison

**Wales:** Ashley Akbari, Stuart Bedston

# Version History

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Last edited** | **Initials** | **Comment** |
| 0.1 | 20-06-2022 | CR | Initial draft. |
| 0.2 | 15-07-2022 | CR | Draft 2 |
| 0.3 | 10-10-2022 | CR | Draft 3 |
| 0.4 | 24-10-2022 | CR | Draft 4 |
| 0.5 | 25-10-2022 | CR | To Steering Committee |
| 0.6 | 10-11-2022 | CR | Incorporation of comments by Steering Committee |

# Lay Summary

The benefits of Covid-19 vaccinations are well-known; they reduce the risk of infection and lower the risk of serious illness or death associated with Covid-19. Evidence in England and Scotland suggests rates of COVID-19 hospitalisation and COVID-19 death are around five times higher in unvaccinated individuals in comparison to fully vaccinated individuals. Therefore, Covid-19 vaccination is key to society’s recovery from the Covid-19 pandemic.

However, across the four nations of the UK and despite vaccines being freely available, millions of individuals remain completely unvaccinated or are under-vaccinated (defined as not having had all available doses of the vaccine). The rates of unvaccinated or under-vaccinated individuals vary by nation, age, ethnicity and other demographic characteristics.

We will extend current knowledge by investigating unvaccination and under-vaccination across all four nations of the UK and will harmonise the results to give an overall UK representation.

This UK wide national study will analyse how many individuals are completely unvaccinated, and which populations are under-vaccinated or sub-optimally vaccinated as of 1st June 2022. Each of the four nations are responsible for collating the patient data but this has not previously been undertaken using a consistent methodology in each nation. We will investigate children and young people, aged between 5-15 years old, adults whose ages lie in the age range 16-74 and finally adults over 75+ years. These three age ranges will be examined separately as the standard recommended vaccine schedule is different for each age groups.

The standard vaccine schedule for individuals without a clinical condition (putting them at increased risk) is currently 3 doses and sub-optimal vaccination for adults is taken to be 0, 1 or 2 doses only. For people aged 75 or over the standard schedule was 4 doses and sub-optimal vaccination means no vaccination or less than 4 doses.

We will also describe the demographic and clinical characteristics of those who are unvaccinated and sub-optimally vaccinated. Further, we will investigate and describe the relationship between demographics, socioeconomics and co-morbidities and sub-optimal vaccination. Finally, we will also study the relationship between sub-optimal vaccination and severe COVID-19 outcomes over the period from 1st June 2022 to 30th September 2022. We will combine the results across the four nations thereby creating a national record.

The research will be undertaken using electronic healthcare record (EHR) data sources already available and approved to researchers across the four UK nations.

We will provide insights for governments and national public health agencies to help improve vaccine uptake and coverage to as many people in the UK as possible. Our proposed work will feed into public messaging that will highly likely save lives, reduce hospitalisation and morbidity, particularly for the most vulnerable members of society.

# Aim and research questions

The study aims to identify and analyse suboptimal COVID-19 vaccination across the UK in adults, aged 16 and over, and children and young people (CYP), aged 5-15. Sub-optimal vaccination is the state where an individual has not received the standard target scheduled number of doses for their age group, and includes unvaccinated and under-vaccinated. Individuals are under-vaccinated if they have had at least one vaccine dose but not the full target number. The study period is the summer of 2022 during the omicron BA4/BA5 wave.

**RQ1:** What proportion of the eligible UK population has not received any COVID-19 vaccination dose?

**RQ2:** What proportion of the UK population is currently under-vaccinated with respect to COVID-19?

**RQ3:** How are demographic, socio-economic, geographic and co-morbidity characteristics of individuals associated with non-vaccination and under-vaccination?

**RQ4:** How are severe outcomes (COVID-19 related hospitalisations and death) related to sub-optimal vaccination across the UK?

# Study design

We will carry out a prospective observational cohort study of individuals aged 5 years and over who are eligible for vaccination in all four UK nations.

Sub-optimal vaccination within each nation will be defined using the same date in each nation. This is the **study\_start\_date (1st June 2022).** The **study\_end\_date (30th September 2022)** is the date to be used for end of follow-up for severe COVID-19 outcomes.

Two separate analyses are planned. The first is a descriptive analysis of the sub-optimal vaccination groups at the study\_start\_date. The second is a follow up study of severe covid outcomes among the different vaccination groups from the study\_start\_date to the study\_end\_date.

## Sub-optimally vaccinated

The population cohort in each nation will be stratified into different eligibility groups at the study\_start\_date. These groups are defined by age 5-15, 16-74, and 75+. These are age groups where the standard vaccine schedule is 2, 3 or 4 doses respectively by study\_start\_date. Individuals are defined as sub-optimally vaccinated if they have received fewer doses than their eligibility group has been offered by study\_start\_date. Some individuals may have been eligible for more doses than has been offered to their eligibility group by study\_start\_date (e.g. if they are immunosuppressed), but will not be considered under-vaccinated. The reason for this simplification is the difficulty in identifying exactly who has been offered more doses than their eligibility group in the four nations. The table below gives the definition of sub-optimal vaccination in each age group.

|  |  |
| --- | --- |
| Age Group | Sub optimal Vaccination |
| 5-11 | Unvaccinated |
| 12-15 | Unvaccinated, 1 dose |
| 16-74 | Unvaccinated, 1 or 2 doses |
| 75+ | Unvaccinated, 1, 2 or 3 doses |

Descriptive analyses will be conducted on the observational cohorts within each nation to characterise those who are sub-optimally vaccinated at the study\_start\_date with the tabulations pooled over the four nations to give a UK-wide description. Logistic or multinomial regression analyses will be carried out in each nation and the estimated coefficients and variance-covariance matrix shared for combining via an inverse variance weighted meta-analysis. For a restricted set of important demographics, socio-economic factors and summary co-morbidities we will form multiway tables of counts and pool these for a ‘pooled individual level’ meta-analysis.

For the logistic/multinomial regression analysis the dependent variable will be the vaccination status at the study\_start\_date. In some analyses, the vaccination status variable will be sub-optimally vaccinated versus fully vaccinated. In others this it will be number of vaccine doses received. Suboptimal vaccination status at the study\_end\_date will also be tabulated.

## Severe outcomes from study\_start\_date to study\_end\_date

A severe COVID-19 outcome is defined as either:

1. COVID-19 Death, where COVID–19 is mentioned as one of the causes of death on the death certificate, or
2. COVID-19 Hospitalisation, where COVID–19 is recorded as one of the reasons for the admission to hospital on the discharge record.

The primary outcome of the study will be a composite outcome of COVID-19 Hospitalisation or COVID-19 death. The date of outcome will be taken as the date of the earliest of these outcomes. Severe COVID-19 outcomes among the sub-optimally vaccinated groups will be enumerated from the study\_start\_date to the study\_end\_date. We will fit a Cox model with vaccine status as a time-dependent exposure.

## Meta Analysis and pooling of results

To give a UK perspective, we will provide cohort summary tables stratified by number of vaccine doses in each nation which will then be aggregated to give UK totals.

For the logistic/multinomial regression and Cox model coefficients the estimated coefficients and full covariance matrix will be shared for combining in an inverse variance weighted meta-analysis.

# Data sources

* CVD-COVID-UK/COVID-IMPACT in England (57M population)
* Honest Broker Service in Northern Ireland (1.9M population)
* EAVE II in Scotland (5.4M population)
* SAIL Databank in Wales (3.2M population)

Principal data sources required in each nation:

* Vaccination data: Type, date of administration of COVID-19 vaccines.
* PCR testing data: Date and outcome of SARS-CoV-2 polymerase chain reaction (PCR) tests.
* LFT data: Date and outcome of lateral flow tests.
* Primary care data: Routinely collected records from GPs across the UK containing information on QCovid Risk Groups. (This is not available in NI where Prescribing data are used to identify BNF prescriptions)
* Hospital admission data: Date of admission to hospital, date of discharge, COVID-19 as primary or secondary reason for admission.
* Mortality data: Date of death and whether COVID-19 was a primary, secondary or underlying cause.
* Demographic data: Age, sex, ethnic group, area of residence, household size, deprivation, urban/rural status.
* Other healthcare data which can be used to establish if the individual is likely to be in the nation after 8th December 2020 and so eligible for vaccination. This could include data on childhood /seasonal vaccinations; telehealth calls (NHS24/NHS111), out of hours GP services, Accident and Emergency attendances, Prescriptions, Maternity Services and Outpatients attendance. This will be different in each participating nation depending on data availability.

# Sample selection

* Has sex and age, or year of birth, recorded.
* Recorded as alive and living in the nation on study\_start\_date.
* Has local authority area recorded so that we have information on urban/rural status.
* Aged 5 years old and over at study\_start\_date
* England: Registered in GDPPR at study\_start\_date (about 96% of England population and 98% of English general practices).
* Wales: Registered with a SAIL providing GP at study\_start\_date (not all GPs have agreed to share their data with SAIL, 86% have agreed).

# Dates of Vaccine roll out in each nation

The following tables give important vaccine roll out dates for dose 1, 2, 3, 4 by age group in each UK nation

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **Date of eligibility for last dose offered (prior to June 1st 2022)** | | | |
| **Age group** | **Minimum number of doses eligible for** | **Scotland** | **Wales** | **England** | **Northern Ireland** |
| 5-11 | 1 vaccine dose | 19/03/2022 | 08/03/2022 | 01/04/2022 | 01/02/2022 |
| 12-15 | 2 vaccine doses | 03/01/2022 | 14/12/2021 | 20/12/2021 | 01/12/2021 |
| 16-74 | 2 vaccine + 1 booster doses | 28/02/2022 |  | 17/01/2022 | 01/12/2021 |
| 75+ | 2 vaccine + 2 booster doses | 25/03/2022 |  | 23/03/2022 | 01/03/2022 |

# Data cleaning for vaccination data

In all 4 nations the vaccination records are not perfect with individuals recorded as having 2 vaccinations a short time apart and individuals with missing vaccinations such as a record for dose 1 and dose 3 but no dose 2 or no records for dose 1 or 2 but records for dose 3 and 4. The following data cleaning principles are suggested.

Any individual appearing in the vaccination data base with a valid ID number will be taken as vaccinated.

There will be no exclusions of individuals with an irregular vaccination record as that might inadvertently inflate the unvaccinated populations.

For individuals with multiple dose 1 records pick the first one, after Dec 08 2020, omit the duplicate records, unless there is no dose 2 and the gap is at least 19 days. In which case rename the duplicated dose 1 as dose 2 and retain this record.

For individuals with a short time between dose 1 and dose 2, less than 19 days, will have the second record omitted and subsequent vaccine dose numbers renumbered. 21 days was the stipulated gap between dose 1 and dose 2 of Pfizer and some of the early vaccinees in December 2020 had the company schedule. 19 is used for individuals validly getting dose 1 on a Monday and dose 2 on a Friday.

For individuals with multiple dose 2 records pick the first one, after Dec 29 2020, omit the duplicate records, unless there is no dose 3 and the gap is at least 9 weeks. In which case rename the duplicated dose 2 as dose 3 and retain this record.

The gap between dose 2 and dose 3 for most people was 24+ weeks but immune-compromised individuals had a 12-week gap and younger individuals who got the second dose in autumn 2021 had and even shorter gap of 9 weeks. Individuals with a gap between dose 2 and dose 3 of less than 63 days will have the dose 3 record omitted and subsequent records renumbered.

For individuals with multiple dose 3 records pick the first one, after Sept 13 2021, omit the duplicate records, unless there is no dose 4 and the gap is at least 9 weeks. In which case rename the duplicated dose 3 as dose 4 and retain this record.

The gap between dose 3 and dose 4 is only relevant for individuals aged 75 or more or 16-64 and immune-compromised. The gap for the 75+ was 24+ weeks but immune-compromised individuals had a 12-week gap. Individuals with a gap between dose 3 and dose 4 of less than 84 days will have the dose 4 record omitted and subsequent records renumbered.

Vaccine status at June 01 2022 will be recorded as the number of doses received and this will have values 1,2,3,4+

# Study Cohorts

Vaccine rollout and eligibility for varying numbers of doses is largely driven by age and specific risk groups. With the data availability in each of the nations, it is difficult to identify those in the targeted risk groups to be able to report on sub-optimal vaccination taking into account risk group. It is easy to identify cohorts based upon age. The three cohorts we will investigate for the study of severe COVID-19 events are:

* Children and Young People (CYP) aged 5-15
* Adults aged 16-74
* Adults aged 75+

At the study\_start\_date, CYP aged 5-11 at least one vaccination represents fully vaccination; for CYP aged 12-15, 2 doses represent full vaccination; for adults aged 16-74, full vaccination is 3+ doses and for adults aged 75+ 4+doses represent full vaccination.

# Measures

## Outcomes

The main outcome in the first analysis is number of doses received at study\_start\_date

Individuals with inconsistent vaccine records will not be excluded from the analysis as the aim is to enumerate vaccine uptake at the study\_start\_date. For example, if the vaccine record indicates that an individual only has one record but this is for a third dose before the study\_start\_date, then this individual will be counted in the dose 3 group and fully vaccinated if they are an adult aged 16-74.

For the second analysis of outcomes between study\_start\_date and study\_end\_date:

* COVID-19 Death (where COVID–19 is mentioned as one of the causes of death on the death certificate), and date of death or
* COVID 19 Hospitalisation (where COVID–19 is mentioned as one of the reasons for the admission to hospital on the discharge record) and date of first hospital admission.
* The primary outcome of the study will be a composite outcome of COVID-19 Hospitalisation or COVID-19 Death. The date of the outcome will be the date of hospitalisation for those who died after a COVID-19 hospital admission, and the date of death for those who died without being admitted to hospital for COVID-19.

## Covariates

### Minimal Analysis

* Age Groups: 5-11; 12-15; 16-17; 18-24; 25-29, and then 5-year age groups up until 80-84; 85+
* Sex: Male, Female.
* Index of Multiple Deprivation Quintile: 1 = Most deprived, 5 = Least deprived.
* Urban/rural classification of residence (2 levels): urban, rural.
* Ethnic group (mapped to ONS 5 classification (White, Black, Mixed, Asian, Other, Null)
  + Not in Northern Ireland.
* Number of co-morbid conditions: 0, 1, 2, 3, 4, 5+
  + England, Scotland, Wales use QCovid
  + Northern Ireland use BNF chapters.

### Extended Analysis

* Time since last positive PCR test: no positive test, 0-13 weeks, 14-26, 27 or more.
* Number of PCR tests (all and positive) in the 6-month period prior to the study\_start\_date.
* Variant of last previous infection, if any, based upon time period.
* Previous hospital admission for COVID-19 (Yes/No) since March 01 2020.
* Appropriate geography (e.g., NHS or region).
* Ever in Shielding Group (Yes/No). Not available in Northern Ireland
* Was another household member asked to shield (as an indicator of if they may have been offered an early vaccination). Not available in Northern Ireland.
* Individual QCovid Risk Groups.
* BMI <20, 20-24, 25-29, 30-34, 35-39, 40+, NA.
* COVID-19 vaccination types.
* Identification of individuals with prescriptions for various BNF chapters – e.g. anti-depressants, anxiolytics, antipsychotics.

### Second Aim – Follow up Study

* Vaccination dates so individuals changing vaccination group can be tracked for the follow up analysis.
* Date of first admission to hospital for any reason between the study\_start\_date and the study\_end\_date so that a non covid admission to hospital can be included as a time dependent covariate.

# Exploratory analysis

## Ghost Individuals

It is likely that, except for the data in Wales, there will be individuals in the cohorts who represent records of individuals who were in the nation at some time in the past but have left, or also individuals who are registered at two separate general practices with different identifiers, such as students or individuals who work away from home on a regular basis. We propose two solutions to this issue which will have a big impact on the enumeration of those who are unvaccinated.

One solution is to keep the whole denominators but to calculate sample weights to weight the e-cohort population back to the official population sizes, although these may be an underestimate of the current resident population. These weights could be derived using age, sex and appropriate geography. These weights would normally be below 1 for everyone in the e-cohort. Reweighting by assigning a weight of 1 to individuals who had some contact with the health service over the last 2 years could be used and the weights for those with no contact down scaled so that the sum of the weights is equal to the official population size.

For cohorts based upon GP registered populations, which are often 5-15% higher than the official population estimates, the initial study weights are calculated as , where is the number of individuals in the official population in age group *i* with sex *j*, and is the corresponding figure for the GP population; *k* is the index for the individual records. Typically, in the age ranges up to 18 and from 50 onwards up to 90, , so that the weights will be close to 1; in the age range from 18 to 40 where there is much more movement of individuals for work and study , and the weights will be much less than 1. With this initial choice of sample weights, , where the sum, over *k*, is over all individuals in age group *i* and sex *j* in the GP population.

This initial weighting scheme treats all individuals in the GP population records the same within a particular age group and sex combination. However, many individuals have had an interaction with the health service in the two or five years prior to the study\_start\_date and so are known to exist in the populations. In particular the covid testing and covid vaccination data records have a huge population coverage. By trawling through all the appropriate health data sets, including ones for tele health and accident and emergency we can reweight the GP population such that =1, if individual *k* in in age group *i* and sex *j* has had an interaction with the health service. Individuals who have not had any interaction are the ones who are most likely to be ghost records and their reweight is given by

,

where , is the sum of the original weights and so is the total official population size, , where the sum is only over those individuals who are in at least one of the population databases in the previous period and represents the number of records corresponding to individuals with confirmed health service contact, and where the sum is over those individuals in the GP denominator who have not had any health service contact in the previous period.

With this modification to the sampling weights individuals who have had a health service contact in the recent period are upweighted to 1 as they are not likely to be ‘ghost’ patients; individuals with no contact are down-weighted. The sum of both sets of weights are identical so the population size is preserved but there is no longer the restriction that . This is reasonable if a particular age group has more individuals accessing the health service than in the official population estimates; this might arise with immigration.

An alternative approach is to base the analysis of the unvaccinated only among those who have had some contact with the health service over the last two years and accept that there will be some individuals who are in each nation, who have not had contact with the health service and who are unvaccinated. It is likely that this approach will underestimate the total number unvaccinated, but at least we will know that they are active users of the health service, including for mass vaccination and mass covid testing.

## Plots and Tables

### Not Eligible for Vaccination

Tabulate the number of individuals (aged 5+) who are recorded as being ineligible for vaccination at study\_start\_date

* 1. Overall Total
  2. By age group
  3. By sex
  4. By deprivation
  5. By ethnicity

### Trends in Vaccine Uptake

* For full vaccinated plot time trends of week of first/second/third (booster)/fourth dose.
* For one dose only plot time trends of week of first dose.
* For two doses only plot time trends of week of first/second dose.
* Plot histogram of the weeks between dose 1 and 2 vaccinations for fully vaccinated and for two doses only.

### Tabulations of Vaccine Uptake and serious covid outcomes

* Counts and percentages of each event type at study\_start\_date:
  + First dose, Second dose, Third, Fourth dose.
* Counts and rates of each event type at study\_start\_date by number of doses:
  + COVID-19 Death.
  + COVID-19 ICU admission, if available.
  + COVID-19 Hospitalisation.

# Statistical-analysis

## Sub-optimal vaccination at study\_start\_date

A separate analysis will be carried out within each of the four cohorts, Children and young people age 5-11, 12-15, Adults 16-74 and Adults 75+. As the definition of full vaccination differs between CYP aged 5-11 and those aged 12-15 this cohort will be analysed in two parts.

Logistic regression with response variable vaccination status - Under Vaccinated or Fully Vaccinated, modelling the odds of being under vaccinated.

Multinomial regression with response variable vaccination status - Under Vaccinated or Fully Vaccinated, modelling the odds of being under vaccinated.

Models to fit:

Unadjusted and adjusted models for all 6 variables in the minimal analysis group. Age will be included as a categorical variable rather than a spline. This makes the interpretation and meta-analysis easier.

Unadjusted and adjusted models for all variables in the minimal and extended analysis groups. When including the individual QCovid groups/BNF categories the composite variable – number of risk groups will not be included.

## Severe Covid-19 outcomes from study\_start\_date to study\_end\_date

A separate analysis will be carried out within each of the three cohorts, Children and young people age 5-15, Adults 16-74 and Adults 75+. These endpoints are rare in CYP and so the cohort will be analysed as a whole using the appropriate definitions of full vaccination for the age groups.

Cox proportional hazards regression with response variables confirmed (1) COVID-19 Hospitalisation, (2) COVID-19 Death, and, (3) COVID-19 Hospitalisation or Death, modelling the hazard of the serious COVID-19 outcome. The principal exposure variable is vaccination status. Time will start at the study\_start\_date and end at the earliest of the date of COVID-19 hospitalisation or death, death from other causes, study\_end\_date or date of leaving the study (if known). Individuals who are fully vaccinated for their age group will remain in that group for the remainder of the study period, until censored or end of study, even if they receive additional vaccine doses.

For this short follow up, age group and all covariates apart from vaccine status, will be measured at the study\_start\_date. The exposure variables will be time dependent during the follow up period. We do not propose to study waning of vaccine protection here and for simplicity will use the date of vaccination during the study period to denote when an individual changes from one vaccine group to another, as opposed to using a 7- or 14-day period to allow the immune response to develop.

# Meta-analysis

Results to be shared:

* Sample descriptive table of characteristics (counts), stratified by explanatory variables.
* Raw model results: log odds ratios/hazard ratios and variance covariance matrices.

Pooling of the results will take place in one Trusted Research Environment (TRE) which is still to be decided. We plan to use a fixed effects meta-analysis as we are using the same analytical technique in each nation with similar data sets and definitions. This will be based upon the rma.mv function within the metafor package for R.

# Sensitivity analysis

The primary analysis will use sample weights to correct for the likely over count of patients no longer resident in the country among the unvaccinated.

A sensitivity analysis will establish the denominator for the whole cohort as those who have been in contact with any part of the health service electronic records available in the nations in the 3-year period prior to the study\_start\_date.

In the follow up analyses of severe COVID-19 events in children and young people (CYP) we will undertake a stratified analysis of those aged 5-11 and 12-15, in view of the differing dose strength of the vaccines in these two age groups.

# Tables to prepare for sharing

These should be ‘long thin’ tables with one row per combination and the number in each combination recorded. There will be no zeroes in these tables.

Week of vaccination dose 1 by age group by vaccine type, tabulating number vaccinated each week.

Similar tables for dose 2, 3, 4.

At the study\_start\_date (we may also want to do this at the study\_end\_date for the most up to date classification).

Sub-optimal vaccination group by age group by sex, tabulating number in each combination.

Sub-optimal vaccination group by age group by deprivation, tabulating number in each combination.

Sub-optimal vaccination group by age group by number of co-morbid groups, tabulating number in each combination.

Sub-optimal vaccination group by age group by ethnicity, tabulating number in each combination.

Details of the tables are in the excel file Coalesce\_Draft\_Tables\_v0.1.xlsx

# Definition of Variables

|  |  |  |
| --- | --- | --- |
| Variable name to be used in the tables |  | Groups / Levels |
| age\_group |  | 5-11; 12-15;16-17; 18-24; 25-29, up until; 80-84; 85+ |
| sex |  | Male, Female |
| deprivation |  | 1 – High, 2,3,4 5-Low, |
| urban\_rural | Urban Rural classification | Rural/Urban |
| ethnic\_group |  | White, Black, Mixed, Asian, Other, Missing |
| n\_risk\_gps | Number of Q Covid Risk groups/ BNF categories in NI | 0,1,2,3,4,5+ |
| time\_last\_pos\_pcr | Time since last positive PCR test | no positive test, 0-13 weeks, 14-26, 27 or more |
| n\_pcr\_tests | Number of PCR tests in the 6-month period prior to the study\_start\_date | 0,1,2,3,4,5+ |
| n\_pos\_pcr\_tests | Number of positive PCR tests in the 6-month period prior to the study\_start\_date | 0,1,2+ |
| variant\_last\_inf | Variant of last previous infection, if any, based upon time period | never\_positive, wild\_type, alpha, delta, omicron, |
| prior\_hosp\_covid | Previous hospital admission for COVID-19 (Yes/No) since March 01 2020 – confirmed admission using discharge records | No, Yes |
| ever\_shielding | Ever in Shielding Group - Not in Northern Ireland | No, Yes |
| household\_ever\_sheild | Was another household member asked to shield (as an indicator of if they may have been offered an early vaccination) | No, Yes |
| bmi | Body Mass Index | <20, 20-24, 25-29,30-34, 35-39, 40+, Missing |

## Definition of QCOVID variables

**QCOVID-** Definition of QCOVID items used for co-morbidity score.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Comorbidity (value range) | Description | Details |
| 1 | GP prescribed immunosuppressant medication (Yes/No) | Have you been prescribed immunosuppressant’s prescribed by your GP | Prescribed four or more times in the previous 6 months |
| 2 | Prescribed leukotriene or LABA (Yes/No) | Are you taking anti-leukotriene or long acting beta2-agonists (LABA)? | Prescribed four or more times in the previous 6 months |
| 3 | GP prescribed Oral steroids (Yes/No) | Have you been prescribed oral steroids by your GP in the last 6 months? | oral prednisolone containing preparations prescribed four or more times in the previous 6 months |
| 4 | Atrial Fibrillation (Yes/No) | Do you have atrial fibrillation? |  |
| 5 | Heart failure (Yes/No) | Do you have heart failure? |  |
| 6 | Asthma (Yes/No) | Do you have asthma? |  |
| 7 | Cancer of blood or bone marrow (Yes/No) | Have you a cancer of the blood or bone marrow such as leukaemia, myelodysplastic syndromes, lymphoma or myeloma and are at any stage of treatment? |  |
| 8 | Coronary heart disease (Yes/No) | Do you have coronary heart disease? |  |
| 9 | Cirrhosis of liver (Yes/No) | Do you have cirrhosis of the liver? |  |
| 10 | Congenital heart disease (Yes/No) | Do you have congenital heart disease or have you had surgery for it in the past? | The categorization should be based either a Read code for congenital heart disease OR an HES OPCS code for surgery for congenital heart disease ever |
| 11 | Chronic obstructive pulmonary disease (Yes/No) | Do you have chronic obstructive pulmonary disease (COPD)? |  |
| 12 | Dementia (Yes/No) | Do you have dementia? |  |
| 13 | Epilepsy (Yes/No) | Do you have epilepsy? |  |
| 14 | Osteoporotic fracture (Yes/No) | Have you had a prior fracture of hip, wrist, spine or humerus? |  |
| 15 | Motor neurone disease (Yes/No) | Do you have motor neurone disease, multiple sclerosis, myaesthenia, or Huntingtons's Chorea? |  |
| 16 | Parkinson’s disease (Yes/No) | Do you have Parkinson’s disease? |  |
| 17 | Pulmonary hypertension or pulmonary fibrosis (Yes/No) | Do you have pulmonary hypertension or pulmonary fibrosis? |  |
| 18 | Cystic fibrosis, bronchiectasis or alveolitis (Yes/No) | Do you have cystic fibrosis or bronchiectasis or alveolitis? |  |
| 19 | Peripheral vascular disease (Yes/No) | Do you have peripheral vascular disease? |  |
| 20 | Rheumatoid arthritis or SLE (Yes/No) | Do you have rheumatoid arthritis or SLE? |  |
| 21 | Respiratory cancer (Yes/No) | Do you have lung or oral cancer? |  |
| 22 | Severe mental illness (Yes/No) | Do you have severe mental illness? |  |
| 23 | Stroke (Yes/No) | Have you had a stroke or TIA? |  |
| 24 | Diabetes type (0:none;  1:Type 1;  2:Type 2) | Do you have diabetes? |  |
| 25 | Thrombosis or pulmonary embolus (Yes/No) | Have you had a thrombosis or pulmonary embolus? |  |
| 26 | Body Mass Index (15.0 to 47.0) | Body Mass Index | The most recently recorded patient BMI within the last 5 years. |
| 27 | Chemotherapy  0:none;  1: Group A;  2: Group B;  3: Group C | Have you had chemotherapy in the last 12 months? | Chemotherapy prescribed in preceding 12 months as recorded on the Systemic Anti Cancer Treatment (SACT) data. Chemotherapy classified into 3 categories (sheet 3) |
| 28 | Housing status  0: neither;  1: care home;  2: homeless | what is your housing category - care home or homeless or neither? | The most recently recorded accommodation status recorded on GP record |
| 29 | Learning disability/Down’s syndrome  0: neither;  1: learning disability;  2: Down's | Do you have a learning disability or Down's Syndrome? | The most recently recorded value. If some has a code both for learning disability and Downs, they should be coded as Downs |
| 30 | Radiotherapy in previous 6 months (Yes/No) | Have you had radiotherapy in the last 6 months? | coded as having radiotherapy in the preceding 6 months on either HES or RTDS |
| 31 | CKD stage 3 or 4 (1 to 6) | Do you have kidney disease? | if ckd3==1 then code as 2; if ckd4==1 then replace as 3; if ckd5=1 then replace as 4; if ckd5=1 & dialysis in last 12 months replace as 5; if ckd5=1 & transplant ever code as 6 |