OpenSAFELY Research Protocol

Disparities in COVID-19 vaccine coverage and disease burden over time

OpenSAFELY project number and summary: Project 174
https://jobs.opensafely.org/echo-evaluation-of-covid-19-vaccine-histories-using-opensafely/

GitHub repository: https://github.com/opensafely/covid-vaccine-history

Project team / authors

Edward P. K. Parker[1,2]*, Martina Pesce[3]*, Helen McDonald[4], Venexia Walker[5,6], Chris Wood[3], Helen J Curtis[3], William J. Hulme[3]

* Joint first authors

[1] Department for Infectious Disease Epidemiology and International Health, London School of Hygiene and Tropical Medicine, London, UK
[2] NIHR Health Protection Research Unit in Immunisation at the London School of Hygiene and Tropical Medicine
[3] The Bennett Institute for Applied Data Science, Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK
[4] Department of Life Sciences, University of Bath, Bath, UK
[5] Bristol Medical School: Population Health Sciences, University of Bristol, Bristol, UK
[6] MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

Version number	Version date	Summary of changes made
1.0	18 November 2025	–

Table of Contents

Abstract

Introduction
Primary vaccination with two doses of COVID-19 vaccine began in England on 08 December 2020. Since September 2021, booster campaigns have been delivered in the Spring and Autumn to reinforce protection among individuals at greatest risk of severe disease.

We present a study protocol for the standardised assessment of disparities in COVID-19 vaccine coverage and severe disease burden across vaccination campaigns in England. The protocol accommodates changes to campaign start and end dates, vaccine products available, and eligibility criteria.

Methods
For each campaign since December 2020, we will describe vaccine coverage and severe disease burden, overall and in subgroups defined by sociodemography (including age, ethnicity, deprivation index, and care home residence) and clinical characteristics (including risk groups defining eligibility for vaccination). We will use Kaplan-Meier estimators to report on cumulative vaccination coverage in each subgroup. As a sensitivity analysis, we will report coverage among those who remained alive and registered throughout the course of each campaign. We will use Poisson regression to obtain Incidence Rate Ratios and 95% confidence intervals quantifying the association between each subgroup covariate level and severe disease outcomes (COVID-19-related hospitalisation, COVID-19-related critical care admissions, and COVID-19-related mortality).

We will use the OpenSAFELY research platform, which provides secure access to routinely-collected health records for millions of people in England, to conduct analyses. All analyses are encoded in R scripts, which are fully executable against simulated dummy data before being run against real data in OpenSAFELY.

Discussion
Disparities in vaccine coverage and severe disease across sociodemographic and clinical subgroups have been evident since the early days of vaccine implementation in England. This standardised protocol aims to quantify these disparities in a harmonised way across campaigns using transparent and reproducible analytic code.

Background

The national COVID-19 vaccination programme in the UK began on 08 December 2020 in a campaign designed to deliver a two-dose primary series, first in those at high risk of hospitalisation and death from COVID-19, and later in the entire adult population (1). This was followed by the first COVID-19 vaccine “booster” campaign in September 2021, and subsequent campaigns have continued for individuals at greatest risk of severe disease in the Spring and Autumn each year.

Across each phase of vaccine implementation, the Joint Committee on Vaccination and Immunisation (JCVI) has advised the UK government on vaccine eligibility with the aim of protecting individuals at greatest risk of severe disease. Eligibility for Spring and Autumn campaigns has evolved over time, with criteria for successive campaigns defined by age, clinical risk from underlying health conditions, care home residency, and occupation.

By the end of October 2021, 85% of adults in England had received two COVID-19 vaccine doses (2), highlighting the operational effectiveness of the initial vaccine roll-out and the high acceptance of vaccines among the general population. However, certain sociodemographic and clinical subgroups experienced lower vaccination rates compared to the general population (3). For example, younger individuals, specific ethnic minority groups, pregnant women, residents of more deprived areas, individuals with chronic neurological conditions (including learning disabilities), and individuals with severe mental illness all exhibited suboptimal primary vaccine uptake (3,4), and these disparities have persisted in subsequent booster campaigns (5).

Providing robust estimates of COVID-19 vaccine coverage is crucial to: (i) monitor equity of the vaccination programme; (ii) identify under-served populations; and (iii) inform tailored delivery approaches. Efforts to monitor and enhance equitable vaccine coverage align with recommendations of the World Health Organization’s Commission on Social Determinants of Health (6), and contribute to the UK Public Health Outcomes Framework of reducing health inequalities (7) and Core20PLUS5 health inequities framework (8). To complement estimates of coverage in individual campaigns, a summary of vaccination histories at the start of each successive campaign is necessary to: (i) identify persistently un- or under-vaccinated population subgroups; and (ii) highlight common product combinations suitable for further evaluation in vaccine safety and effectiveness studies. In addition to evaluating vaccine coverage, it is important to monitor the burden of severe COVID-19 (including hospitalisation, critical care admissions, and mortality) in different population subgroups to support decision-making related to vaccine eligibility.

The OpenSAFELY research platform is a highly secure software platform that offers the capacity for near-real-time analysis of electronic health records (EHR) data covering 95% of patients registered with a GP in England. Our group has previously used OpenSAFELY to quantify disparities in COVID-19 vaccine coverage (3,4,9) and burden (10,11). However, our last comprehensive update on coverage disparities was conducted in September 2023 (5), while our most recent estimates of COVID-19-related mortality rates span up to August 2022 (11). Timely estimates are required to ensure that programmatic and policy decisions are informed by up-to-date estimates of coverage and disease burden.

Here, we present a protocol for the analysis of COVID-19 vaccine coverage and severe disease burden across the initial primary vaccine roll-out and subsequent booster vaccine campaigns in England. The analysis aims to harmonise the reporting of disparities in vaccine coverage and disease burden across sociodemographic and clinical subgroups across campaigns.

Objective

The primary objective of this study is to comprehensively describe receipt of COVID-19 vaccines, including products and timings, across the adult population in England over successive vaccination campaigns. Vaccination history and coverage will be reported overall and in sociodemographic and clinical subgroups to systematically quantify disparities in coverage across campaigns.

As a secondary objective, we will use the same analysis framework to quantify disparities in severe disease, including COVID-19-related hospitalisation, COVID-19-related critical care admissions, and COVID-19-related mortality.

The outputs will be presented via an interactive dashboard to support interpretation and decision-making.

Methods

Summary

We will conduct an observational, retrospective study to describe COVID-19 vaccine coverage and severe disease patterns in England across the initial primary vaccine roll-out (2020–2021) and each of the subsequent biannual campaigns (2021–2025).

For each campaign, the analysis will include three components:

- What is the prior COVID-19 vaccination history of individuals at the start of the campaign?
- What COVID-19 vaccine coverage is achieved over the course of the campaign?
- What is the burden of severe COVID-19 over the course of the campaign?

Each component will be reported for the total eligible population and in key sociodemographic and clinical subgroups of interest.

Campaign-specific analyses will be supplemented by a static analysis of COVID-19 vaccination frequency among adults since the roll-out began (**Appendix**).

OpenSAFELY platform

All data will be linked, stored and analysed securely using the OpenSAFELY platform, https://www.opensafely.org/, as part of the NHS England OpenSAFELY COVID-19 service. No GP data from patients who have registered a Type-1 Opt out with their GP surgery will be included in this study. Data include pseudonymised data such as coded diagnoses, medications and physiological parameters. No free text data are included. All code will be shared openly for review and re-use under MIT open license. Detailed pseudonymised patient data is potentially re-identifiable and therefore will not be shared. A detailed description of the OpenSAFELY platform is provided by Nab et al (12).

Access to the underlying identifiable and potentially re-identifiable pseudonymised EHR data is tightly governed by various legislative and regulatory frameworks, and restricted by best practice. The data in the NHS England OpenSAFELY COVID-19 service is drawn from General Practice data across England where TPP are the data processors.

TPP developers initiate an automated process to create pseudonymised records in the core OpenSAFELY database, which are copies of key structured data tables in the identifiable records. These pseudonymised records are linked onto key external data resources that have also been pseudonymised via SHA-512 one-way hashing of NHS numbers using a shared salt. University of Oxford, Bennett Institute for Applied Data Science developers and PIs, who hold contracts with NHS England, have access to the OpenSAFELY pseudonymised data tables to develop the OpenSAFELY tools.

These tools in turn enable researchers with OpenSAFELY data access agreements to write and execute code for data management and data analysis without direct access to the underlying raw pseudonymised patient data, and to review the outputs of this code. All code for the full data management pipeline — from raw data to completed results for this analysis — and for the OpenSAFELY platform as a whole is available for review at github.com/OpenSAFELY.

Study design

Study entry and follow-up

The study population for each campaign will be defined at the campaign start date (as defined in **Table 1**). Follow-up will start on the campaign start date. We will estimate coverage at 8 weeks after the campaign start date as an early reporting milestone. The primary reporting milestones for vaccine coverage will be 30 June for Spring campaigns and 28 February (of the subsequent calendar year) for Autumn campaigns, aligning with those adopted by NHS England and the UK Health Security Agency (e.g. (13,14)). The campaign end date, and final reporting milestone, will be one day before the start of the next campaign, thereby ensuring complete capture of vaccine outcome events across campaigns.

Eligibility criteria

The following eligibility criteria will be applied for each campaign to define the base population and ensure minimum data quality standards:

- Alive at the campaign start date;

- Registered at the same GP practice for at least 12 weeks up to the start date, to enable transfer of records for those who have moved between GPs;
- Aged between the campaign-specific minimum age and 104 years old, excluding missing values. Age is calculated at a date some weeks after the start of the campaign to allow those who turn 50 years old, say, during the campaign to be vaccinated (see **Table 3**);
- Sex recorded as “male” or “female”, excluding “intersex”, “unknown”, or missing values.

Individuals will be included in the campaign report if they meet all of the criteria above. In addition, vaccination eligibility criteria for each campaign will be applied for subgroup analysis (see ‘**Analysis subgroups’** below).

Outcomes

(i) COVID-19 vaccination

Receipt of vaccination in OpenSAFELY-TPP is identified by both the “target disease” or pathogen (in this case COVID-19 / SARS-CoV-2) and the specific product and dose used (e.g., for the original Pfizer/BioNTech vaccine BNT162b2, the product name is “*COVID-19 mRNA Vaccine Comirnaty 30micrograms/0.3ml dose conc for susp for inj MDV (Pfizer)”*).

Receipt of COVID-19 vaccination as a whole will be identified based on the target disease. Products used for primary vaccination or booster doses in the routine UK programme (listed in **Table 2**) will be identified specifically, with all other products classified as “other”.

When defining COVID-19 vaccination history, vaccination records occurring less than 14 days after a prior vaccination will be excluded as these are likely to represent duplicate records.

(ii) Severe COVID-19

To quantify the burden of severe COVID-19, we will report on:

- COVID-19-related hospitalisation, defined based on Secondary Use Service (SUS) in-patient hospital episodes with a relevant ICD-10 diagnosis code as the primary or non-primary diagnosis;
- COVID-19-related critical care admissions, defined based on COVID-19-related hospitalisations (as above) with at least one day spent in critical care;
- COVID-19-related mortality, defined based on death certification with a relevant ICD-10 code anywhere on the death certificate (i.e., as an underlying or contributing cause).

Table 1: Summary of dates for COVID-19 vaccination campaigns among individuals aged 16 years and over in the UK.

Campaign	Programmatic start date (source)	First record in UK Health Security Agency coverage reports (source)	Start date for common protocol	Early reporting milestone	Primary reporting milestone	Final reporting milestone
Primary series (doses 1/2)	08/12/2020 (NHS England)	Week ending 13/12/2020[a]	07/12/2020	31/01/2021	30/06/2021	05/09/2021
Autumn 2021[b]	Initial advice for extended primary series in immunosuppressed issued on 01/09/2021 (JCVI); approved by treasury on 06/092021 (UK Parliament)	6-fold rise among individuals ≥80 years in week ending 19/09/2021[a]	06/09/2021	31/10/2021	28/02/2022	20/03/2022
Spring 2022	01/04/2022 (NHS England)	Week ending 27/03/2022[a]	21/03/2022	15/05/2022	30/06/2022	28/08/2022
Autumn 2022	General delivery from 12/09/2022 (NHS England), with delivery to care homes from 05/09/2022 (UK Parliament)	Week ending 04/09/2022[a]	29/08/2022	23/10/2022	28/02/2023	02/04/2022
Spring 2023	General delivery from 17/04/2023, with delivery to care homes from 03/04/2023 (NHS England)	Week ending 09/04/2023[a]	03/04/2023	28/05/2023	30/06/2023	27/08/2023
Autumn 2023	11/09/2023 (NHS England), brought forward from initial plans for roll-out from 02/10/2023 (NHS England)	Week ending 03/09/2023[a]	28/08/2023	22/10/2023	28/02/2024	14/04/2024
Spring 2024	General delivery 22/04/2024, with delivery to care homes from 15/04/2023 (NHS England)	15/04/2024 (UKHSA; published 05/07/2024)	15/04/2024	09/06/2024	30/06/2024	29/09/2024
Autumn 2024	03/10/2024 (NHS England)	01/10/2024 (UKHSA dashboard; accessed 22/07/2025)	30/09/2024	24/11/2024	28/02/2025	30/03/2024
Spring 2025	01/04/2025 (NHS England)	01/04/2025 (UKHSA; published 03/07/2025)	31/03/2025	25/05/2025	30/06/2025	–

References within the table are included as hyperlinks to enhance usability. The start date for each campaign takes the Monday from the week containing the minimum of the programmatic start date and the first record in UK Health Security Agency coverage reports. The end date and final reporting milestone for each campaign takes the date preceding the start of the next campaign to ensure complete capture of outcomes. Primary reporting milestones for vaccine coverage are 30 June for Spring campaigns (and primary vaccination) and 28 February (of the subsequent calendar year) for Autumn campaigns.

[a] Data up to Autumn 2023 obtained from 'National flu and COVID-19 surveillance report: 4 January 2024 (week 1)' (15).

Table 2: COVID-19 vaccine products approved for use in the UK.

Product	Short name	Approval date (with source in hyperlink)	First campaign usage
Pfizer–BioNTech COVID-19 vaccine BNT162b2 (Comirnaty®)			
Pfizer–BioNTech vaccine	BNT162b2	02/12/2020	Primary series
Bivalent Pfizer–BioNTech Original/Omicron BA.1 vaccine	BNT162b2/BA.1	03/09/2022	Autumn 2022
Bivalent Pfizer–BioNTech Original/Omicron BA.4-5 vaccine	BNT162b2/BA.4-5	Extension of BNT162b2/BA.1 approval 09/11/2022	Spring 2023
Monovalent Pfizer–BioNTech XBB.1.5 vaccine	BNT162b2.XBB.1.5	Extension of BNT162b2 approval from 05/09/2023	Autumn 2023
Monovalent Pfizer–BioNTech JN.1 vaccine	BNT162b2.JN.1	24/07/2024	Autumn 2024
AstraZeneca COVID-19 vaccine ChAdOx1-S (Vaxzevria®)			
AstraZeneca COVID-19 vaccine	ChAdOx1-S	30/12/2020	Primary series
Moderna COVID-19 vaccine mRNA-1273 (Spikevax®)			
Moderna COVID-19 vaccine	mRNA-1273	08/01/2021	Primary series
Bivalent Moderna Original/Omicron vaccine	mRNA-1273/BA.1	15/08/2022	Autumn 2022
Bivalent Moderna Original/Omicron BA.4-5 vaccine	mRNA-1273/BA.4-5	Extension of mRNA-1273/BA.1 approval 21/02/2023	Spring 2023
Monovalent Moderna XBB.1.5 vaccine	mRNA-1273.XBB.1.5	15/09/2023	Autumn 2023
Monovalent Moderna JN.1 vaccine	mRNA-1273.JN.1	02/09/2024	Autumn 2024
Sanofi Pasteur COVID-19 vaccine (VidPrevtyn Beta®)			
Sanofi Pasteur COVID-19 vaccine	Vidprevtyn	21/12/2022	Spring 2023

References within the table are included as hyperlinks to enhance usability. The following vaccines have also been authorised for use in the UK but have not been widely implemented during the routine programme: Novavax COVID-19 vaccine (Nuvaxovid®); Novavax XBB vaccine; and HIPRA bivalent Beta/Alpha COVID-19 vaccine (BIMERVAX®).

Analysis subgroups

The age-eligibility threshold has differed across successive campaigns, as summarised in **Table 3**. Clinical risk groups encompass various chronic or immunosuppressive conditions and medications as set out in the Green Book chapter 14a (16). The clinical risk groups eligible for vaccination have also varied across campaigns (**Table 3**); for example, in Spring campaigns, eligibility has focused on immunosuppressed individuals, while a broader range of conditions have defined eligibility in Autumn campaigns.

We will describe outcomes in the entire eligible population and in subgroups defined by sociodemographic and clinical characteristics, as ascertained using primary care records at the start of the vaccination campaign (see **Table 4** for further details on how subgroups will be defined, including details of codelists where applicable).

Our analysis framework (summarised in **Figure 1**) comprises two levels:

- level 1 comprises *groups* that define the population of interest;
- level 2 comprises *subgroups* of interest within the selected level 1 group.

(i) Level 1: groups

Groups of interest are defined as follows:

- All adults (16–104 years)
- Age-based groups:
    - 75–104 years
    - 65–74 years
    - 50–64 years
    - 16–49 years
- Eligibility-based groups:
    - Age ≥ campaign threshold
    - ≥1 clinical risk group (any age)
    - ≥1 clinical risk group (age < threshold)
    - Care home residents
- Clinical risk-based groups:
    - Chronic respiratory disease
    - Chronic heart disease
    - Chronic kidney disease
    - Chronic liver disease
    - Chronic neurological disease
    - Diabetes mellitus
    - Immunosuppression (including asplenia)
    - Severe obesity
    - Severe mental illness

Clinical risk groups will be defined according to the national COVID-19 vaccination uptake reporting specification developed by PRIMIS – a set of codelists and business rules commissioned by the UK Health Security Agency to support identification of clinical risk groups using GP IT systems (**Table 4**).

(ii) Level 2: subgroups
Within each level 1 group, we will describe disparities in coverage and severe COVID-19 outcomes across nested subgroups defined by sociodemography and clinical characteristics (see **Table 4** for full details). Sociodemographic characteristics of interest include age, sex, region, ethnicity, deprivation of area of residence, and care home residence. Clinical characteristics of interest include the clinical risk groups used to define vaccine eligibility (e.g., chronic respiratory disease) as well as additional exploratory subgroups, including those nested within existing clinical risk groups (e.g., dialysis recipients, people with liver cirrhosis, and cancer subgroups defined by time since most recent record). All characteristics will be ascertained using primary care data at the campaign start date.

Note that some variables (e.g. age, care home residency) define both level 1 groups and level 2 subgroups. As level 1 groups, they would be used to define the overall population within which the reporting of other subgroups are nested. As level 2 subgroups, they would be used to define the covariate levels being compared within a level 1 group.

Descriptive and statistical analyses

(i) Vaccination history
In each subgroup, we will describe the distribution of the following metrics as ascertained at the start of each campaign:

- total number of prior COVID-19 vaccines recorded;
- time since most recent prior COVID-19 vaccine, if any;
- product combination of prior COVID-19 vaccines.

Distributions will be summarised using quantiles (medians, quartiles, etc) and proportions as appropriate.

(ii) Vaccination coverage
Kaplan-Meier curves will be generated for each level 1 group. In each level 2 subgroup, we will report the cumulative proportion of the population vaccinated over time. We will report this for any vaccination and for specific vaccine products, with vaccination by a different product considered as a competing event. Death is also considered as a competing event. Practice deregistrations will be considered as censoring events. We will use Kaplan-Meier estimators to report the cumulative vaccination coverage.

As a sensitivity analysis, we will report coverage overall and stratified by product among individuals who remained alive and uncensored throughout the course of the campaign.

(iii) Severe COVID-19
In each subgroup, we will report the incidence of COVID-19-related hospitalisation, COVID-19-related critical care admission, and COVID-19-related mortality per 1000 person-years during each campaign. Confidence limits for the incidences will be derived using standard errors of the log of the incidence. To quantify the association between each subgroup covariate levels and each COVID-19-related outcome, we will use Poisson regression to obtain Incidence Rate Ratios (IRRs) and 95% confidence intervals (CIs), and use Cox regression to

obtain Hazard Ratios (HRs) and 95% CIs. Estimates are calculated based on the campaign start date up to the day before the next campaign, with people removed from the risk-set (censored) at practice deregistration or death. Following our previous analyses of COVID-19-related mortality over time (11), we will adjust for age (using a restricted cubic spline with three knots) and sex in all models.

Table 3: Summary of dates and eligibility criteria for COVID-19 vaccination campaigns among individuals aged 16 years and over in the UK.

Campaign	Minimum age for non-clinical risk eligibility (years)	Reference date for age [a]	Clinical risk-based eligibility below age threshold	Eligibility based on care home residence	Vaccines used for routine vaccination	Reference for vaccines used
Primary series (doses 1/2)	16	31/03/2021	N/A (phased priority)[b]	✓	BNT162b2 mRNA-1273 ChadOx1-S	Green Book
Autumn 2021	16	31/08/2021	N/A (phased priority)[c]	✓		JCVI (01/09/2021) (14/09/2021)
Spring 2022	75	30/06/2022	Immunosuppressed	✓	BNT162b2 mRNA-1273	JCVI (21/02/2022)
Autumn 2022	50	31/03/2023	All clinical risk groups	✓	BNT162b2 mRNA-1273 BNT162b2/BA.1 mRNA-1273/BA.1	JCVI (03/09/2022)
Spring 2023	75	30/06/2023	Immunosuppressed	✓	BNT162b2/BA.4-5 mRNA-1273/BA.4-5 Vidprevtyn	JCVI (07/03/2023) UKHSA (12/10/2023)
Autumn 2023	65	31/03/2024	All clinical risk groups	✓	BNT162b2/BA.4-5 mRNA-1273/BA.4-5 BNT162b.XBB.1.5 mRNA-1273.XBB.1.5 Vidprevtyn	JCVI (30/08/2023) Kirsebom et al (2024)
Spring 2024	75	30/06/2024	Immunosuppressed	✓	BNT162b.XBB.1.5 mRNA-1273.XBB.1.5	JCVI (07/02/2024)
Autumn 2024	65	31/03/2025	All clinical risk group	✓	BNT162b.JN.1 mRNA-1273.JN.1	UKHSA (17/09/2024)
Spring 2025	75	30/06/2025	Immunosuppressed	✓	BNT162b.JN.1 mRNA-1273.JN.1	UKHSA (24/03/2025)

References within the table are included as hyperlinks to enhance usability. This table does not cover eligibility for the following groups: individuals under 16 years of age; frontline health and social care workers; carers; household contacts of people with immunosuppression; staff working in care homes for older adults. Campaign-specific eligibility is as described in the March 2025 edition of Green Book Chapter 14a (16). JCVI, Joint Committee on Vaccination and Immunisation; UKHSA, UK Health Security Agency.

[a] Age was defined as of 31 March 2021 in initial COVID-19 vaccine coverage surveillance reports, and this was updated to 31 August 2021 as of October 2021 (17). We align with these definitions for the primary and Autumn 2021 campaigns. For subsequent booster campaigns, age thresholds of 31 March of the subsequent calendar years (for Autumn campaigns) and 30 June (for Spring campaigns) are standard, with individuals eligible from the outset of the campaign if they reach the age threshold at any point during the campaign.

[b] Phased prioritisation based on age, care home residence, health and social care worker status, and clinical risk status (as described Table 2 of the Green Book

Chapter 14a (16)).

[c] The Autumn 2021 campaign had several phases reflecting the evolving epidemic situation. On 01 September 2021, JCVI advised that a third primary dose should be offered to immunosuppressed adults. These individuals would also become eligible for a booster (fourth) dose at an interval of 3 months, although for the purposes of this study we do not distinguish between third doses administered as part of an extended primary series versus a booster campaign. On 14 September 2021, JCVI advised that booster doses should be offered to priority groups including individuals living in care homes for older adults, adults aged ≥50 years, and individuals aged 16–49 years in clinical risk groups (among others). Booster roll-out was prioritised based on age and clinical risk groups, as per phase 1 of the COVID-19 programme. Following the emergence of the Omicron variant in late 2021, eligibility was subsequently expanded to adults ≥18 years (November 2021) and to those aged 16–17 years (December 2021).

Table 4: Sociodemographic and clinical characteristics used to define population subgroups.

Variable	Notes	Levels	Codelists
Sociodemographic subgroups			
Age group (11-category)	Age in years, stratified into approximately 10-year age bands among individuals under 50 years of age, and 5-year age bands among individuals over 50 years of age (representing the principal age-based target population across Spring and Autumn campaigns).	16–19; 20–29; 30–39; 40–49; 50–54; 55–59; 60–64; 65–69; 70–74; 75–79; 80+	–
Age group (4-category)	Age in years, stratified into broad age bands aligning with eligibility thresholds applied across one or more booster campaigns (50, 65, and 75 years).	16–49; 50–64; 65–74; 75+	
Sex	–	Female; Male	–
Ethnicity (16-category)	Derived from most recent record in primary care	British or Mixed British; Irish; Other White; White + Black Caribbean; White + Black African; White + Asian; Other mixed; Indian or British Indian; Pakistani or British Pakistani; Bangladeshi or British Bangladeshi; Other Asian; Caribbean; African; Other Black; Chinese; Other; Unknown	opensafely/ethnicity-snomed-0removed/22911876
Ethnicity (6-category)	Derived from most recent record in primary care	White; Mixed; Asian or Asian British; Black or Black British;	opensafely/ethnicity-snomed-0removed/22911876

Variable	Notes	Levels	Codelists
		Chinese or Other Ethnic Groups; Unknown	
Region	NHS region derived from practice address	East; East Midlands; London; North East; North West; South East; South West; West Midlands; Yorkshire and the Humber	–
Index of multiple deprivation	Social deprivation quintile derived from individual's address at Lower Super Output Area (a small geographical area defined by the ONS)	1 (most deprived); 2; 3; 4; 5 (least deprived)	–
Care home residence	Based on matching of individual's address with care homes in CQC database or presence of code in primary record before date defined	0; 1	primis-covid19-vacc-uptake/longres/v2.5
Core subgroups (Green Book criteria via PRIMIS operationalisation)[a]			
Any clinical risk group	A composite of the core subgroups below	0; 1	–
Chronic respiratory disease	Defined based on the PRIMIS variable "RESP_GROUP"	0; 1	primis-covid19-vacc-uptake/astadm/v2.5 primis-covid19-vacc-uptake/ast/v2.5 primis-covid19-vacc-uptake/astrxm2/v2.5 primis-covid19-vacc-uptake/astrxm1/v2.5 primis-covid19-vacc-uptake/resp_cov/v2.5
Chronic heart disease and vascular disease	Defined based on the PRIMIS variable "CHD_COV_DAT"	0; 1	primis-covid19-vacc-uptake/chd_cov/v2.5
Chronic kidney disease	Defined based on the PRIMIS variable "CKD_GROUP"	0; 1	primis-covid19-vacc-uptake/ckd15/v2.5 primis-covid19-vacc-uptake/ckd35/v2.5 primis-covid19-vacc-uptake/ckd_cov/v2.5
Chronic liver disease	Defined based on the PRIMIS variable "CLD_DAT"	0; 1	primis-covid19-vacc-uptake/cld/v2.5

Variable	Notes	Levels	Codelists
Chronic neurological disease (including learning disability)	Defined based on the PRIMIS variables “CNS_GROUP” and “LEARNDIS_DAT”	0; 1	primis-covid19-vacc-uptake/cns_cov/v2.5 primis-covid19-vacc-uptake/learndis/v2.5
Diabetes mellitus and other endocrine disorders	Defined based on the PRIMIS variable “DIAB_GROUP”	0; 1	primis-covid19-vacc-uptake/diab/v2.5 primis-covid19-vacc-uptake/dmres/v2.5 primis-covid19-vacc-uptake/addis_cod/v2.5 primis-covid19-vacc-uptake/gdiab_cod/v2.5
Immunosuppression (including asplenia)	Defined based on the PRIMIS variables “IMMUNO_GROUP” and "SPLN_COV_DAT"	0; 1	primis-covid19-vacc-uptake/immdx_cov/v2.5 primis-covid19-vacc-uptake/immrx/v2.5 primis-covid19-vacc-uptake/immunosuppression-admin-codes/v2.5/ primis-covid19-vacc-uptake/dxt_chemo_cod/v2.5 primis-covid19-vacc-uptake/spln_cov/v2.5
Severe obesity	Defined based on the PRIMIS variable “BMI_GROUP”	0; 1	primis-covid19-vacc-uptake/bmi/v2.5 primis-covid19-vacc-uptake/bmi_stage/v2.5 primis-covid19-vacc-uptake/sev_obesity/v2.5
Severe mental illness	Defined based on the PRIMIS variable “SEVMENT_GROUP”	0; 1	primis-covid19-vacc-uptake/smhres/v2.5 primis-covid19-vacc-uptake/sev_mental/v2.5
Extended subgroups			
CKD or RRT	Most recent serum creatinine level (in µmol/l) and age on date of creatinine measurement will be used to calculate estimated glomerular filtrate rate (eGFR), then converted to CKD status as defined below. Most recent RRT status (dialysis or transplant) will supersede CKD status where applicable.	No CKD or RRT; CKD stage 3a; CKD stage 3b; CKD stage 4; CKD stage 5; RRT (dialysis); RRT (transplant)	nhsd-primary-care-domain-refsets/renaltransp_cod/20250627/ nhsd-primary-care-domain-refsets/dialysis_cod/20250627/ nhsd-primary-care-domain-refsets/cre_cod/20250627/

Variable	Notes	Levels	Codelists
	eGFR thresholds were as follows: - Stage 3a: 45–59 mL/min/1.73 m^2 - Stage 3b: 30–44 mL/min/1.73 m^2 - Stage 4–5: 15–29 mL/min/1.73 m^2 - Stage 5: <30 mL/min/1.73 m^2		
Chronic obstructive pulmonary disease	Any prior code in primary care record	0; 1	nhsd-primary-care-domain-refsets/copd_cod/20250627
Down's syndrome	Any prior code in primary care record	0; 1	nhsd-primary-care-domain-refsets/ds_cod/20250627
Sickle cell anaemia	Any prior code in primary care record	0; 1	nhsd/sickle-spl-atriskv4-snomed-ct/7083ed37
Liver cirrhosis	Any prior code in primary care record	0; 1	nhsd-primary-care-domain-refsets/cirrhosis_cod/20250627

PRIMIS-based operationalisation of clinical risk groups is based on underlying codelists, logic, and business rules that are not in the public domain. However, the implementation of these rules in ehrQL is publicly available.

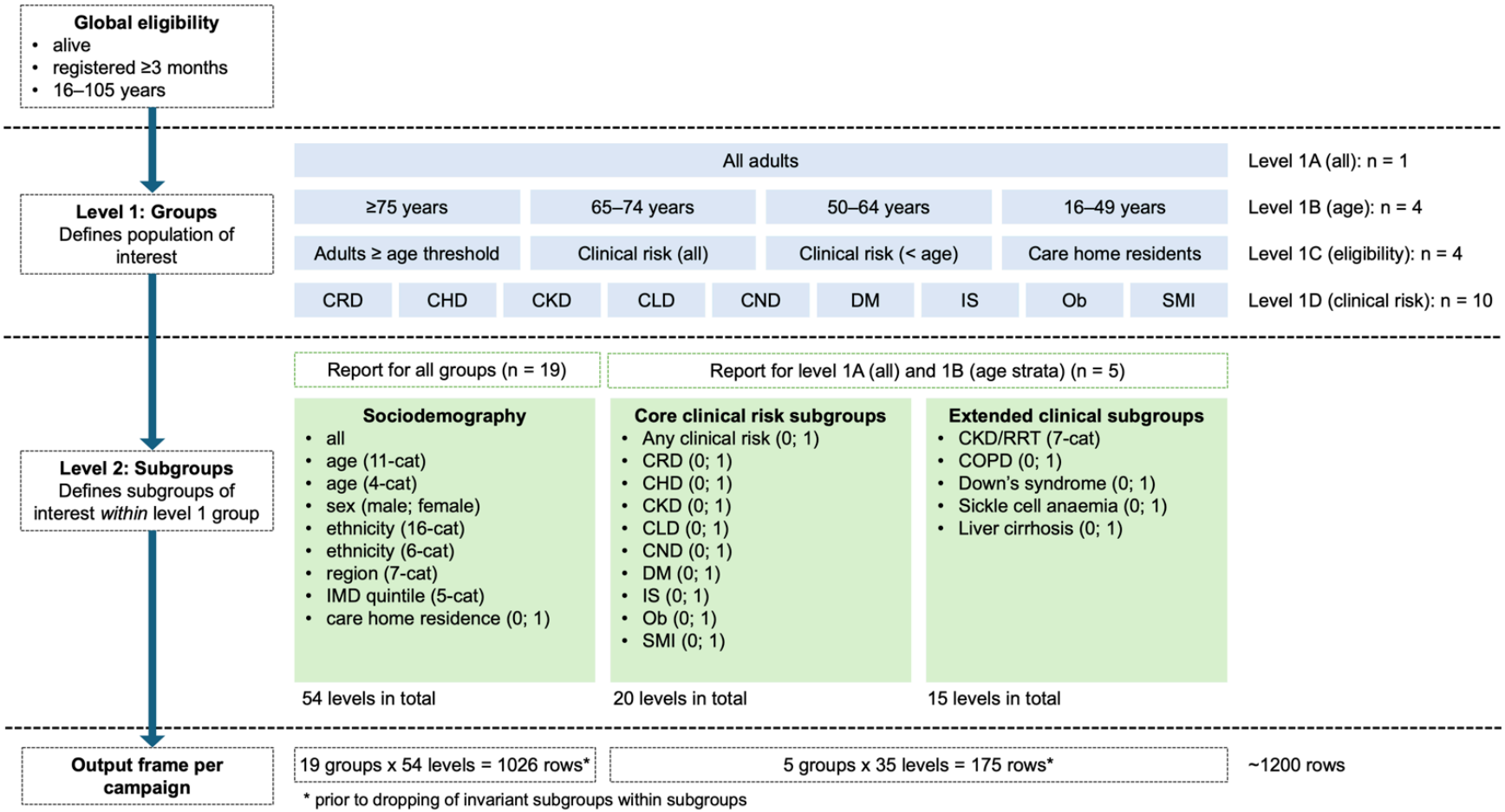

Figure 1: Analysis framework for handling of subgroups. See **Table 4** for further information on variable definitions. CHD, chronic heart disease; CKD, chronic kidney disease; CLD, chronic liver disease; CND, chronic neurological disease; COPD, chronic obstructive pulmonary disease; CRD, chronic respiratory disease; IMD, index of multiple deprivation; IS, immunosuppression including asplenia; Ob, severe obesity; RRT, renal replacement therapy; SMI, severe mental illness.

Limitations

We anticipate several limitations in our proposed analysis:

1. Vaccines administered at private clinics are likely to be under-recorded in our coverage estimates.
2. We will define clinical risk groups based on recording in primary care, supporting harmonisation with the GP IT search strategies in the national uptake reporting specification developed by PRIMIS. Notably, vaccine-eligible individuals are also identified and invited for vaccination using cohorts identified by NHS Digital's 'Cohorting as a Service' (CaaS) (18)]. Alongside primary care and demographic data, CaaS uses multiple additional datasets including Hospital Episode Statistics (HES), the Systemic Anti-Cancer Therapy (SACT) dataset, and the Radiotherapy Treatment Data Service (RTDS). In our cohort, we are likely to underascertain vaccine-eligible subgroups identified by treatments given in secondary care (e.g. those who are immunosuppressed due to active chemotherapy or radiotherapy).
3. In addition, the following subgroups have also been eligible for vaccination for one or more booster campaigns, but are considered beyond the scope of the current study given that they cannot be reliably defined based on primary care records:
    - health and social care workers;
    - care home staff;
    - close contacts of immunosuppressed individuals;
    - carers; and
    - pregnant individuals (included among eligible clinical risk groups for a subset of campaigns).

    Notably, while health and social care worker status was assigned to doses administered in the early stages of COVID-19 vaccine implementation in the UK, this was only recorded upon vaccination so cannot be used to establish a reliable denominator for vaccine coverage estimates.
4. Our analysis aligns with the implementation of COVID-19 vaccines across successive campaigns. However, these campaigns have not aligned directly with waves of COVD-19 related to the emergence of SARS-CoV-2 variants. This may pose a challenge to the interpretation of findings on severe disease burden. Some subgroups will also be small, such that rare outcomes (e.g., COVID-19-related critical care admissions and COVID-19-related mortality) will be below disclosure control thresholds in one or more campaigns (see '**Disclosure Control'** below).
5. Finally, post-COVID-19 syndrome (often termed 'long COVID') is an important (and often severe) outcome of COVID-19 but is considered beyond the scope of our study due to evolving practice in diagnosis and recording, and uncertainty in the time lags between initial infections and their long-term sequelae.

Disclosure Control

OpenSAFELY operates a strict disclosure control policy to prevent disclosure of sensitive information and preserve patient anonymity. In particular, summaries of groups of people of size 5 or fewer must not be reported, including the size of the group itself. To meet this policy, we will round counts to the nearest 10, and redact all other values that summarise groups of people of size 5 or fewer. Values at the midpoint between rounding values (15, 25, 35, …) will be rounded upwards, with the exception of 5 which will be rounded down to zero, to ensure the “5 or fewer” principle is met. Counts taking values [0-5, 6-14, 15-24, 25-34, …] will therefore map to values [0, 10, 20, 30, …]. Counts between 1 and 5 are therefore indistinguishable from zero.

Software and Reproducibility

Data management will be performed using ehrQL/Python, with analysis carried out using R. Code for data management and analysis as well as codelists archived online at https://github.com/opensafely/covid-vaccine-history.

Dissemination

We will make all underlying statistical outputs available via jobs.opensafely.org. Outputs will be presented via an interactive dashboard.

References

1. Harnden A, Earnshaw A. Lessons from the United Kingdom's COVID‐19 vaccination strategy. Med J Aust. 2021 May;214(9):417–9.

2. The rollout of the COVID-19 vaccination programme in England.

3. Curtis HJ, Inglesby P, Morton CE, MacKenna B, Green A, Hulme W, et al. Trends and clinical characteristics of COVID-19 vaccine recipients: a federated analysis of 57.9 million patients' primary care records *in situ* using OpenSAFELY. Br J Gen Pract. 2022 Jan;72(714):e51–62.

4. Hopcroft LE, Curtis HJ, Brown AD, Hulme WJ, Andrews CD, Morton CE, et al. First dose COVID-19 vaccine coverage amongst adolescents and children in England: an analysis of 3.21 million patients' primary care records in situ using OpenSAFELY. Wellcome Open Res. 2023 June 9;8:70.

5. Coverage snapshot - 01 September 2023 | OpenSAFELY: Reports [Internet]. [cited 2025 July 28]. Available from: http://reports.opensafely.org/reports/coverage-snapshot-01-september-2023/

6. WHO. Closing the gap in a generation [Internet]. [cited 2025 July 28]. Available from: https://iris.who.int/bitstream/handle/10665/69832/WHO_IER_CSDH_08.1_eng.pdf?sequence=1

7. GOV.UK [Internet]. 2019 [cited 2025 July 28]. Public Health Outcomes Framework: indicator changes 2019 to 2022. Available from: https://www.gov.uk/government/consultations/public-health-outcomes-framework-proposed-changes-2019-to-2020

8. England NHS. NHS England » Core20PLUS5 (adults) – an approach to reducing healthcare inequalities [Internet]. [cited 2025 Sept 10]. Available from: https://www.england.nhs.uk/about/equality/equality-hub/national-healthcare-inequalities-improvement-programme/core20plus5/

9. The OpenSAFELY Collaborative, Parker EP, Tazare J, Hulme WJ, Bates C, Carr EJ, et al. Factors associated with COVID-19 vaccine uptake in people with kidney disease: an OpenSAFELY cohort study. BMJ Open. 2023 Jan;13(1):e066164.

10. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. Factors associated with COVID-19-related death using OpenSAFELY. Nature. 2020 Aug 20;584(7821):430–6.

11. Nab L, Parker EPK, Andrews CD, Hulme WJ, Fisher L, Morley J, et al. Changes in COVID-19-related mortality across key demographic and clinical subgroups in England from 2020 to 2022: a retrospective cohort study using the OpenSAFELY platform. Lancet Public Health. 2023 May;8(5):e364–77.

12. Nab L, Schaffer AL, Hulme W, DeVito NJ, Dillingham I, Wiedemann M, et al.

OpenSAFELY: A platform for analysing electronic health records designed for reproducible research. Pharmacoepidemiol Drug Saf. 2024;33(6):e5815.

13. UKHSA. GOV.UK. [cited 2025 July 28]. National flu and COVID-19 surveillance report: 27 March (week 13). Available from: https://www.gov.uk/government/statistics/national-flu-and-covid-19-surveillance-reports-2024-to-2025-season/national-flu-and-covid-19-surveillance-report-27-march-week-13

14. UKHSA. National flu and COVID-19 surveillance report (18 July 2024) [Internet]. [cited 2025 July 28]. Available from: https://assets.publishing.service.gov.uk/media/6698de0f49b9c0597fdaff3e/Weekly_flu_and_COVID-19_report_w29.pdf

15. UKHSA. National Influenza and COVID-19 Surveillance Report (4 January 2024) [Internet]. [cited 2025 July 28]. Available from: https://assets.publishing.service.gov.uk/media/6596accec23a10000d8d0b7e/Weekly-flu-and-COVID-19-surveillance-report-week-1.pdf

16. UKHSA. Green Book 14a - COVID (17 Mar 2025) [Internet]. [cited 2025 June 16]. Available from: https://assets.publishing.service.gov.uk/media/67d8a1979dc953ac3bfe9382/GreenBook-chapter-14a-COVID-19-17_3_25.pdf

17. NHS. COVID-19 Vaccination Statistics (Week ending Sunday 24th October 2021) [Internet]. [cited 2025 July 28]. Available from: https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2021/10/COVID-19-weekly-announced-vaccinations-28-October-2021.pdf

18. NHS England Digital [Internet]. [cited 2025 Sept 29]. Cohorting as a Service (CaaS). Available from: https://digital.nhs.uk/services/cohorting-as-a-service-caas

Appendix: Retrospective analysis of vaccination frequency over time

Purpose

In a supplementary analysis, we provide a descriptive overview of patterns of vaccine coverage across successive campaigns.

Study entry and follow-up

The study observation period will start on Monday 07 December 2020, the Monday of the week when the first COVID-19 vaccine was administered in England outside the context of a clinical trial (the first vaccine was given on Tuesday 8 December but we start follow-up for each campaign on a Monday for consistency). It will end on Sunday 1 June 2025 in the first instance but will be extended to later dates in subsequent analyses. The index date will be defined on a rolling basis throughout the study (see 'Eligibility criteria' below).

Some participants in randomised trials may have been vaccinated before 7 December 2020, though the recording status of these vaccines in NHS health records is uncertain, particularly given that trial participants were initially blinded to their assigned treatment. We will identify vaccination events from 1 June 2020 to explore the frequency of in-trial vaccination, though reporting for our main analysis will start from 7 December 2020.

Outcome

See definition for COVID-19 vaccination in the main protocol.

Eligibility criteria

There is no fixed study population: eligibility for inclusion will be defined as at the day of vaccination.

Eligibility criteria on each day are:

- currently alive;
- currently registered at a GP, using the registration status table;
- aged between 16 and 104 years old inclusive (excluding missing values); and
- sex recorded as "male" or "female" (excluding "intersex", "unknown", or missing values).

Descriptive analysis

We will report:

- the number of vaccines received on each 4-week period for the entire observation period;
- the interval in weeks since the most recent prior dose.

Numbers will be reported overall, stratified by vaccine products (**Table 2**) and stratified by sociodemographic and clinical subgroups (**Table 4**), with the latter ascertained on the day of delivery of a given dose.

Appendix table: Identification of COVID-19 vaccine products

Product	Character string associated with product
Pfizer–BioNTech COVID-19 vaccine	
Pfizer–BioNTech vaccine (Comirnaty®)	*COVID-19 mRNA Vaccine Comirnaty 30micrograms/0.3ml dose conc for susp for inj MDV (Pfizer)* *COVID-19 mRNA Vaccine Comirnaty Children 5-11yrs 10mcg/0.2ml dose conc for disp for inj MDV (Pfizer)*
Bivalent Pfizer BioNTech (Original/Omicron BA.1 Comirnaty®)	*Comirnaty Original/Omicron BA.1 COVID-19 Vacc md vials*
Bivalent Pfizer–BioNTech Original/Omicron BA.4-5	*Comirnaty Original/Omicron BA.4-5 COVID-19 Vacc md vials*
Monovalent Pfizer–BioNTech XBB	
Monovalent Pfizer–BioNTech XBB.1.5	*Comirnaty Omicron XBB.1.5 COVID-19 Vacc md vials*
Monovalent Pfizer–BioNTech JN.1	*Comirnaty JN.1 COVID-19 mRNA Vaccine 0.3ml inj md vials*
AstraZeneca COVID-19 vaccine	
AstraZeneca COVID-19 vaccine (Vaxzevria®)	*COVID-19 Vaccine Vaxzevria 0.5ml inj multidose vials (AstraZeneca)*
Moderna COVID-19 vaccine	
Moderna COVID-19 vaccine (Spikevax®)	*COVID-19 mRNA Vaccine Spikevax (nucleoside modified) 0.1mg/0.5mL dose disp for inj MDV (Moderna)*
Moderna bivalent (Spikevax® bivalent Original/Omicron vaccine)	*COVID-19 Vac Spikevax (Zero)/(Omicron) inj md vials*
Moderna bivalent (Spikevax® bivalent Original/Omicron BA.4-5)	
Monovalent Moderna XBB	
Monovalent Moderna XBB.1.5	*COVID-19 Vacc Spikevax (XBB.1.5) 0.1mg/1ml inj md vials*
Monovalent Pfizer–BioNTech JN.1	
Novavax COVID-19 vaccine (Nuvaxovid®)	
Novavax COVID-19 vaccine (Nuvaxovid®)	*COVID-19 Vac Nuvaxovid (recombinant, adj) 5micrograms/0.5ml dose susp for inj MDV (Novavax CZ a.s.)*
Novavax XBB COVID-19 vaccine (Nuvaxovid®XBB.1.5)*	
Sanofi Pasteur COVID-19 vaccine	
Sanofi Pasteur COVID-19 vaccine	*COVID-19 Vacc VidPrevtyn (B.1.351) 0.5ml inj multidose vials*

(VidPrevtyn Beta®)	
HIPRA bivalent Beta/Alpha COVID-19 vaccine	
HIPRA bivalent Beta/Alpha COVID-19 vaccine (BIMERVAX®)	

* Vaccine has not been used in the routine UK programme, but entered the private market in Spring 2024.

Useful links:
https://reports.opensafely.org/reports/opensafely-tpp-vaccination-names/
https://github.com/opensafely/covid-vaccine-history/blob/a9830f238ee915c3398817c5f545043ed2ef94dc/analysis/0-lib/design.R#L90
https://reports.opensafely.org/reports/opensafely-tpp-database-schema/#Vaccination
https://jobs.opensafely.org/opensafely-internal/tpp-vaccination-names/outputs/106/