Statistical analysis plan (SAP) to investigate effectiveness of COVID-19 vaccines in children and young people (CYP) aged 12 to 17 years in Scotland

Contents

Introduction

Although many have reported that severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is less likely to cause serious illness in children and young people (CYP) in comparison to older age groups, there is a need for further studies in CYP to understand who among them is at greatest risk of any serious COVID-19 outcomes and why. In a recent study of hospitalised CYP, Harwood and colleagues[1] reported increased risk of severe disease requiring hospitalization or death among infants, teenagers, and those with cardiovascular or neurological co-morbid conditions, or having two or more co-morbid conditions, suggesting a need to prioritise these groups of CYP for vaccination.

Taking into consideration the latest available data among CYP, the Joint Committee on Vaccination and Immunisation (JCVI), on 19 July 2021, advised the UK government on COVID-19 vaccination among CYP.[2] They advised that 12-15 year-olds who are at increased risk of serious illness and hospitalisation from COVID-19, including those with severe neuro-disabilities, immunosuppression, Down's syndrome, and severe learning disabilities, should be offered two doses of BNT162b2 mRNA (Pfizer-BioNTech) vaccine eight weeks apart. Aside from those CYP who are considered high risk, all young people aged 16 and 17 years are being offered only a first dose of vaccine. The timing of a second dose will be confirmed pending further evidence on safety and effectiveness in the group.

This analysis is part of the ongoing Early Pandemic Evaluation and Enhanced Surveillance of COVID-19 (EAVE II) study.[3] Given the available dataset on vaccination among CYP in Scotland, this study aims to investigate effectiveness of Pfizer-BioNTech vaccine, and any other vaccines licensed for this age group in the future, against COVID-19 infection, hospitalisation and death in CYP aged 12 to 17 years in Scotland. We theorise the vaccination programme as a natural experiment with impacts that result from behavioural and other responses that occur as a consequence of broader programmatic effects. We will test the hypothesis that vaccine is safe and effective in this age group and determine its effectiveness (VE) against COVID-19 infection, hospitalisation and death in this age group. To do this, we will use pseudonymised individual level data that are obtained through linking routinely collected primary, laboratory and vaccination healthcare data on CYP across Scotland. All data and analyses will be hosted within Public Health Scotland (PHS) or the Scottish National Safe Haven.

Aims and objectives

Aims

To investigate the association between receiving one and two doses of the Pfizer-BioNTech and COVID-19 infection, hospitalisation and death in children and young people aged 12 to 17 years in Scotland.

Objectives

We seek to:

i. Estimate VE against the composite outcome of COVID-19 infection, hospitalisation or death as a function of time since first SARS-CoV-2 vaccine dose;

ii. Estimate VE against the separate outcomes of COVID-19 infection, hospitalisation and death as a function of time since first COVID-19 vaccine dose;

iii. Assess if VE differs as a function of time since first vaccine dose by vaccine type, age group (12-13; 14-15; 16-17 years) and sex;

iv. Estimate VE for first/second dose timing amongst different age groups (12-13; 14-15; 16-17 years) and risk groups (to be completed when adequate numbers for statistical analysis);

Methods

Study design

First, we will conduct an open prospective cohort study with risk-set matching comparing outcomes (COVID-19 infection, hospitalization and death) amongst first dose vaccinated (partially vaccinated) and second dose vaccinated (fully vaccinated) to people who have not yet been vaccinated. To maximise statistical power, we will conduct secondary analyses using the entire study population for 12-17 year olds. Data from multiple sources will be linked deterministically using the Community Health Index (CHI), a unique identifier used for all health contacts in Scotland.

Setting

Scotland.

Population

Children and young people aged 12-17 years who are resident in Scotland.

Data sources

i. Primary care data: General Practices (n=940) for information on demographics, other population characteristics and vaccination data.
ii. Vaccination centre data: Vaccines administered in national vaccination centres and data available via the Turas Vaccination Management Tool (TVMT).
iii. Secondary care data: Hospital admissions through the Scottish Morbidity Record (SMR) and Rapid Preliminary Inpatient Data (RAPID).
iv. Laboratory test data: RT-PCR laboratory confirmed SARS-CoV-2 infection and data available via the Electronic Communication of Surveillance in Scotland (ECOSS) database.

Selection criteria

Exclusion criteria:

- Age 18 or more at the study start date;
- Deceased prior to the study start date;

Sample size calculations

Although there is not yet any study on vaccine effectiveness among CYP in Scotland, sample size estimations will be based on the Scottish testing and vaccination data. From the first EAVE-II paper on VE against COVID-19 hospitalisation,[4] VE for combined vaccine status at 28-34 days post vaccination was estimated as 0.84, with a standard deviation of 0.06. If we assume VE estimates are asymptotically normally distributed, this gives almost a 100% power to detect a VE of ≥ 0.5. We note a possibility of the study being underpowered for estimating the association between receiving the vaccine and the secondary outcome of death.

Data and data validation

Data variables available

Table 1 lists the groupings of variables available for this study by data source.

Table 1. Groupings of variables by sources

Category	Item	Source
Demographics	Sex	GP
	Age	GP
Others	BMI-for-age	GP
Socio-economic	SMID	GP
Residential settlement	Urban Rural Index (UR6), Health Board, council area	GP
Housing	Private housing, care home or social housing	GP
Clinical diagnoses/Co-morbidities	Underlying conditions (eg. severe neuro-disabilities, immunosuppression, Down syndrome, and severe learning disabilities)	GP
Vaccinations	Vaccine	GP, TMVT
	Vaccine dose	GP, TMVT
	Vaccine date	GP, TMVT
Laboratory tests	RT-PCR SARS-CoV-2 test result	ECOSS
	Date of RT-PCR SARS-CoV-2 test	ECOSS
Sequencing of SARS-CoV-2	Variant of the virus	COG UK
Secondary care	Hospital admission	SMR, RAPID
Mortality	Admission ICD code	SMR
	Death with COVID-19 as the main cause according to death certificate, or death within 28 days of a positive RT-PCR test for COVID-19	NRS

BMI=body mass index, ECOSS= Electronic Communication of Surveillance in Scotland, GP=general practice, NRS=National Records of Scotland; RAPID=Rapid Preliminary Inpatient Data, SMID=Scottish Index of Multiple Deprivation, COG=Centre of Genomics; SMR=Scottish Morbidity Record, TMVT=Turas Vaccination Management Tool.

Exposure data are described in the vaccinations category. Outcome data are described in the secondary care, mortality and laboratory tests categories. The rest of categories contain data on potential confounding factors and effect modifiers.

Consistency and error checking

Data will be checked for patterns of missing and implausible values (e.g. date of second vaccine dose being earlier than the first) for all analytical variables being used, with reasons for exclusion of any records from analysis noted. Alternative variables that are closely related will be considered for a variable of interest with high levels of missingness.

Statistical Analyses

Exposures of interest

Derived from date of receiving PfizerBioNTech vaccine. Exposure categories are as follows:

- 0-13 days after dose 1 or no vaccine record;

- ≥14 days after dose 1 and before dose 2;
- 0-6 days after dose 2
- ≥7 days after dose 2
- ≥14 days after dose 2

To increase statistical power in secondary analyses, we will conduct analyses of whole population data. We expect part of the effects of the vaccination programme to arise from behavioural responses (e.g., behavioural advice received with the invitation letter).

Outcomes of interest

The primary outcome will be a composite outcome of time to COVID-19 infection, hospitalisation or death. COVID-19 infection will be defined as any of COVID-19 symptoms in children with the virus confirmed by RT-PCT; hospitalisation will be defined as a RT-PCR confirmed positive test for SARS-CoV-2 in the 28 days prior to admission, or with ICD-10 code for COVID-19 (in any diagnostic position). COVID-19 deaths will be defined as COVID-19 as the underlying ICD-10 cause of death recorded on the death certificate, or death from any cause within 28 days of a positive RT-PCR test for SARS-CoV-2 infection.
Secondary outcomes will be the single outcomes of: a) COVID-19 infection (as defined above); b) COVID-19 hospitalisation (as defined above), c) COVID-19 deaths, and d) RT-PCR confirmed positive test. We anticipate the RT-PCR confirmed SARS-CoV-2 infection results to be more susceptible to bias arising from differential ascertainment and therefore anticipate treating these results as exploratory.

Potential confounders

Age, sex, socio-economic status (SES) measured by quintiles of the Scottish Index of Multiple Deprivation (SIMD) (1 refers to most deprived and 5 refers to least deprived), residential settlement measured by the urban/rural 6 fold classification (1 refers to large urban areas and 6 refers to small remote rural areas), household size, number and types of comorbidities commonly associated with COVID-19 illness (QCOVID conditions; severe neuro-disabilities, immunosuppression, Down syndrome, severe learning disabilities and other severe conditions, asthma, type 1 diabetes, inborn errors of metabolism), risk factors (BMI for age, relevant QCOVID conditions) and Health Board. We will also include care home status as a potential confounder when data are available.

Potential effect modifiers

Stratification into different population groups by age group (12-13, 14-15, 16-17), sex, time intervals and vaccine exposures and a study of their possible interactions will be performed.

Analytical techniques

We will commence analysis by conducting descriptive analyses to visually inspect trends in vaccination uptake, age-specific COVID-19 hospitalisations and COVID-19 deaths, including by age group and sex. This will include inspecting the number of CYP who have received no doses, one dose and two doses and the length of time between the receipt of one dose and two doses.

To create control groups, we will emulate a 'target trial'. We will do this by conducting 1:1 risk-set matching to identify individuals who had not yet become exposed (i.e., remained unvaccinated) on the date an exposed person received their first vaccine dose. We will do this using a propensity score matching algorithm and will consider incorporating the following characteristics: sex, age, geography, comorbidities, risk factors, number of previous SARS-CoV-2 tests, SES, presence in hospital pre-vaccination and urban-rural settlement. The adequacy of the matching will be assessed by checking for imbalance of the individual covariates across exposure groups.

We expect that any effects before 14 days for the first dose vaccination will largely reflect programmatic effects (e.g., being instructed not to attend vaccination if symptomatic and reinforcement of behavioural advice to reduce COVID-19 related risks). Follow-up will therefore start from day 14 after the date of the first dose vaccination (and day 14 after second dose) for both the exposed and control CYP. It will end on the first of: experiencing the outcome of interest, death (from any cause) or end of follow up period. Unvaccinated controls who become vaccinated will be eligible to become considered as exposed, after follow up is censored for the exposed-unexposed pair at the date of vaccination of the control.

For the matched cohort analysis, we will draw Kaplan-Meier curves to inspect cumulative incidence. We will then fit Cox proportional hazards models and conduct Poisson regression for our outcomes of interest. Given we are particularly interested in the potential for vaccine waning, we will fit a flexible model over time (using a fractional polynomial or smooth spline with person-time defined in days) to explore the timing of any waning of vaccine effectiveness.

Given the relatively small number of anticipated COVID-19 deaths during the follow up period, the secondary analysis focusing on mortality will be based on data from the whole cohort. In addition, secondary analyses will be conducted for hospitalisation and having confirmed SARS-CoV-2 infection.

Vaccination status (unvaccinated, one dose and two doses) will be defined as a time-varying exposure. Poisson regression adjusted for an offset representing the time at risk, with rate ratios (RRs) and 95% confidence intervals (CIs) will be calculated. Vaccine effectiveness and 95% CIs will be calculated as (1-Rate Ratio)*100. Models will be adjusted for relevant confounders, including age, sex, SEP, geography, time period and comorbidities. Stratification variables for each week post-vaccination will be included in the Poisson regression. A polynomial/spline will be fitted to the resulting discrete set of weekly VE estimates. We will conduct hypothesis testing on the resulting vaccine effectiveness fit as a function of time.

Test-Negative Design

A Test Negative Design (TND) case control study nested within a national cohort that compares outcomes of interest between COVID-19 vaccine doses in CYP and unvaccinated controls will be carried out. The time period of this study will begin on the first date of vaccine administration in CYP in Scotland and until the end of follow-up, in line with the TND protocol that has been defined and applied within the EAVE-II collaboration.

Sub-group analysis

Subgroup analyses by vaccine type, age group and sex will be performed.

Sensitivity analysis

We will consider exploring the use of different time intervals following administration of the vaccine to define exposure.

Other analysis

We will consider conducting falsification analyses (negative controls) for alternative time periods (e.g. repeating analyses using time periods two months prior to first vaccination dose) to check the comparability of our exposure groups.

Missing data

Missing data will be reported as percentages of total or raw numbers where possible. Previous analyses have demonstrated that little missing data exist for our key variables of interest. For

covariates which may have a higher proportion of missing data (such as body mass index), we will either use records with no item missingness or use a missing category.

Statistical software

All analyses will be carried out using R/RStudio, version 3.6.1.

Reporting results

Reporting guidelines and conventions

Results will be reported according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) and REporting of studies Conducted using Observational Routinely-collected Data (RECORD) (via the COVID-19 extension) guidelines. P-values will be quoted to two decimal places, unless they are less than 0.001 (whereby the p-value will be given as <0.001) or between <0.005 and >0.001, in which case they will be stated to three decimal places. Measures of association will be reported with 95% CIs.

Dissemination

The analysis will be written in a manuscript and submitted to a peer-reviewed journal. We will also seek to provide near real-time reports on vaccine safety, effectiveness and uptake for the various vaccines to the funders and the government's COVID-19 advisory bodies as appropriate. All code will be made publicly available via the EAVE II GitHub repository. Meta-data will be made available via the HDR Gateway.

References

1. Harwood R, Yan H, Da Camara NT, et al. Which children and young people are at higher risk of severe disease and death after SARS-CoV-2 infection: a systematic review and individual patient meta-analysis. *medRxiv* 2021: 2021.06.30.21259763.
2. NHS Inform. Vaccinating 12 to 15 year olds. 2021. https://www.nhsinform.scot/covid-19-vaccine/invitations-and-appointments/vaccinating-12-to-15-year-olds (accessed 01 September 2021.
3. Simpson CR, Robertson C, Vasileiou E, et al. Early pandemic evaluation and enhanced surveillance of COVID-19 (EAVE II): protocol for an observational study using linked Scottish national data. *BMJ open* 2020; **10**(6): e039097.
4. Vasileiou E, Simpson CR, Shi T, et al. Interim findings from first-dose mass COVID-19 vaccination roll-out and COVID-19 hospital admissions in Scotland: a national prospective cohort study. *The Lancet* 2021; **397**(10285): 1646-57.
5. Hodgson SH, Mansatta K, Mallett G, Harris V, Emary KR, Pollard AJ. What defines an efficacious COVID-19 vaccine? A review of the challenges assessing the clinical efficacy of vaccines against SARS-CoV-2. *The lancet infectious diseases* 2020.
6. US Food and Drug Administration. Development and Licensure of Vaccines to Prevent COVID-19: Guidance for Industry. Washington DC: U.S. Department of Health and Human Services; Food and Drug Administration; Center for Biologics Evaluation and Research; 2020.