

OpenSAFELY Protocol: Comparison of disparities in RSV, Influenza, and COVID-19

Authors: Em Prestige, Rosalind Eggo, Edward Parker, William Hulme, Charlotte Warren-Gash, Elizabeth Williamson

Version: v1

Date: 23/09/2024

<i>Version</i>	<i>Notes</i>
<i>0.1</i>	<i>Initial draft</i>
<i>1</i>	<i>Protocol finalised to align with code being run against data</i>

Table of Contents

Table of Contents	1
Background	3
Objectives	3
Methods	3
Data Source	3
Information Governance	4
Study Design and Population	4
Study Period	4
Inclusion Criteria	5
Exclusion Criteria	5
Study Measures	5
Exposures	5
Socioeconomic status	5
Ethnicity	6
Household size and composition	6
Outcomes	6
Defining Infection 'Episodes'	7
Outcome Phenotypes	8
Covariates	12
Covariate Codelists	13
Statistical Analysis	14
Exploratory Analysis	14
Primary Analysis	14
Outcome Phenotypes	15
Secondary Analysis	15
Additional Covariate Codelists	16
Software and Reproducibility	17
Power Calculations	17
Sensitivity Analyses	17
Outcomes	17
Season Definition	17
Example Table Shells	18
Characteristics	18
Infants	18
Children and Adolescents	23
Adults	26

Older Adults	30
Results - For Example Season e.g. 2021/22	35
Infants	35
Children and Adolescents	41
Adults	44
Older Adults	47
Exploratory Tables	50
Example Figures	53
Methods	53
Results	53
Exploratory Figures	53
Primary Analysis	53
Secondary Analysis	54
Sensitivity Analysis	54
Strengths and Limitations	55
Strengths	55
Limitations	55
Administrative	55
Ethics	55
Funding	55
Conflict of Interests	55
References	56

Background

- Low socio-economic status and minority ethnicity were both identified as risk factors for Influenza, when data from the A/H1N1 epidemic was examined it shows that overall “incidence per 100k in the most deprived 20% was 2.83 times higher than the most affluent 20%” (1). These disparities were also likely to vary amongst children, with overall Influenza admissions being highest amongst healthy children under five (2).
- Similar disparities were identified during the COVID-19 pandemic (3) and were explored using OpenSAFELY (4). It is uncertain whether these disparities exist amongst children as they make up a small number of COVID-19 outcomes.
- Finally, it is thought that disparities within RSV infections exist amongst children in the UK, with studies showing the existence of disparities in both Scotland (5) and the US (6,7)
- This raises the question about how these disparities compare to one another, do characteristics influence these respiratory conditions in the same way

Objectives

1. Describe ethnic and socioeconomic health disparities in RSV, Influenza, and COVID-19
2. Compare disparities between the three respiratory conditions
3. Determine whether certain comorbidities mediate these associations

Methods

Data Source

Primary care records managed by the GP software provider, TPP were linked to ONS death data through OpenSAFELY, a data analytics platform created by our team on behalf of NHS England to address urgent COVID-19 research questions (<https://opensafely.org>).

OpenSAFELY provides a secure software interface allowing the analysis of pseudonymized primary care patient records from England in near real-time within the EHR vendor’s highly secure data centre, avoiding the need for large volumes of potentially disclosive pseudonymized patient data to be transferred off-site. This, in addition to other technical and organisational controls, minimises any risk of re-identification. Similarly pseudonymized datasets from other data providers are securely provided to the EHR vendor and linked to the primary care data. The dataset analysed within OpenSAFELY is based on 24 million people currently registered with GP surgeries using TPP SystmOne software. It includes pseudonymized data such as coded diagnoses, medications and physiological parameters. No free text data is included. Further details on our information governance can be found [here](#).

Outputs are accessed by researchers as aggregated data with suppression of small numbers and rounding of counts in compliance with re-identification minimisation requirements for OpenSAFELY.

Information Governance

NHS England is the data controller; TPP is the data processor; and the key researchers on OpenSAFELY are acting on behalf of NHS England. This implementation of OpenSAFELY is hosted within the TPP environment, which is accredited to the ISO 27001 information security standard and is NHS IG Toolkit compliant (8,9); patient data have been pseudonymized for analysis and linkage using industry standard cryptographic hashing techniques; all pseudonymized datasets transmitted for linkage onto OpenSAFELY are encrypted; access to the platform is through a virtual private network (VPN) connection; the researchers hold contracts with NHS England and only access the platform to initiate database queries and statistical models; all database activity is logged; and only aggregate statistical outputs leave the platform environment following best practice for anonymization of results such as statistical disclosure control for low cell counts (10). The OpenSAFELY research platform adheres to the data protection principles of the UK Data Protection Act 2018 and the EU General Data Protection Regulation (GDPR) 2016. In March 2020, the Secretary of State for Health and Social Care used powers under the UK Health Service (Control of Patient Information) Regulations 2002 (COPI) to require organisations to process confidential patient information for the purposes of protecting public health, providing healthcare services to the public and monitoring and managing the COVID-19 outbreak and incidents of exposure (11). Together, these provide the legal bases to link patient datasets on the OpenSAFELY platform. GP practices, from which the primary care data are obtained, are required to share relevant health information to support the public health response to the pandemic, and have been informed of the OpenSAFELY analytics platform.

Study Design and Population

Study Period

The first date of entry to the study cohort will be 1 September 2016, with the last potential date of inclusion in the study being 1 September 2022. 'Recruitment' will be performed at the beginning of each season (1 September yearly), participants will be included in multiple seasons as long as they meet the inclusion criteria for each season. Individuals will be followed up until the date of death, practice deregistration, or end of the study period (31 August 2023 or 31 August 2024 if data available). The study start date coincides with the first year of availability of HES records in OpenSAFELY and the beginning of that year's respiratory virus season. The study end date has been chosen due to this being the last date of a full season available.

Inclusion Criteria

Four cohorts will be defined in the OpenSAFELY data: infants (up to 2 years old), children and adolescents (aged 2 to 17), adults (ages 18 to 64), and older adults (aged 65 years or over). An individual in the relevant age group will be included in the relevant cohort on the first date they meet the following criteria: registered with a primary care practice in England for whom electronic health records were available through the OpenSAFELY platform, with at least 1 year of continuous registration. All infants will be included from birth, a subgroup analysis will then be conducted on infants with linkage to their birth parent. Infants will be included in this subgroup when the linked birth parent has at least 1 year of continuous registration prior to the index date in order to capture information about the pregnancy relating to the infant. The characteristics of the subgroup will be investigated to determine missingness type of linkage. The four age cohorts will be analysed separately but will include the same participants for all three viruses.

Exclusion Criteria

We will exclude people with missing data on age, sex, or index of multiple deprivation, since these are likely to indicate poor data quality. Infants eligible for treatment using Palivizumab will be excluded from the study due to their high risk of complications due to RSV. This will be infants who meet the following criteria (12):

- children under 2 years of age with severe combined immunodeficiency syndrome
- children under 1 year of age who require long-term ventilation
- children 1–2 years of age who require long-term ventilation and have an additional co-morbidity (including [cardiac disease](#) or [pulmonary hypertension](#))

Study Measures

Exposures

Socioeconomic status

Defined using quintiles of the English Index of Multiple Deprivation, and based on postcode of residence:

- 1 (least deprived)
- 2
- 3
- 4
- 5 (most deprived)

Ethnicity

Depending on outcome counts this may be separated out further into 16 categories following the census categories:

- White
- Mixed
- South Asian
- Black
- Other
- Unknown

Household size and composition

Defined in terms of numbers of household members and the age groups of household members:

- living alone: *participant* living alone (single occupancy household); this is not possible for infants, children, or adolescents
- multiple of the same generation: *participant* living with others of the same generation (reference category);
- one other generation: *participant* and other(s) of the same generation living with people from just ONE other generation;
- two other generations: *participant* and other(s) of the same generation living with people from TWO other generations;
- three other generations: *participant* and other(s) of the same generation living with people from all THREE other generations.

Outcomes

A report of any of the following in primary care or secondary care, or listed as cause of death:

- RSV
- Influenza
- COVID-19
 - COVID-19 will be included in seasons which begin after September 1st 2019, however, any outcomes identified before March 2020 will be censored to align with when testing became established
- Overall respiratory virus (RSV, Influenza, COVID-19, unspecified respiratory virus)

I will use multiple measures for these outcomes, which indicate different levels of severity as a result of each infection type. The measures will be split into mild, severe and mortality as indicated by diagnosis received in primary or secondary care or given on death certificate. For severe outcomes, diagnosis codes in secondary care will be included if in either the first or

second position on the patient record. If a participant initially falls into the mild category, but within the same episode transitions into the severe category, they will be included in the mild category until that transition.

Due to testing data being sparse for RSV we will use a diagnosis of Bronchiolitis as a proxy. Bronchiolitis is considered a good proxy for RSV in infants (13) and is defined as a condition affecting those under 2 years of age (14), hence does not indicate other viruses in adults and therefore can be included in the RSV phenotype (see 'Outcome Phenotypes' below). To capture the burden which would be placed on primary care during out of office hours I will look at emergency attendances for bronchiolitis.

Mortality will be explored as a separate outcome, provided that event numbers are high enough. If mortality rates are too low, it will be included as a composite outcome within the 'severe' outcome bracket. Mortality due to the respective viruses will be identified by the presence of a relevant ICD-10 code on the patient's death certificate as cause of death. All-cause mortality will also be explored.

Defining Infection 'Episodes'

To ensure that outcomes from the same episode are accurately captured, we needed to define episode lengths. This way when codes occur within a certain period they can be assigned to a single infection. This will be defined as 14 days as shown in literature to be the shortest time period recorded where re-infection has not occurred prior to for all viruses (15–18).

Outcome Phenotypes

RSV	Mild (Primary Care, Emergency Attendances)	Severe (Secondary Care)
Primary	<ul style="list-style-type: none"> At least one RSV diagnosis code, or bronchiolitis diagnosis code, in SNOMED CT (codelist), OR, In infant population, bronchiolitis diagnosis code (4120002), in SNOMED CT, in first or second position in ECDS (A&E) 	At least one ICD-10 diagnosis code for RSV or bronchiolitis in the first or second position in the HES APC (codelist)
Sensitivity	<p><i>Inclusion Criteria</i></p> <ul style="list-style-type: none"> At least one RSV diagnosis code, or bronchiolitis diagnosis code, in SNOMED CT, OR, At least two of the following respiratory symptoms recorded in SNOMED CT within a two-week period (two on the same day is acceptable): rhinitis, cough, fever, breathing abnormalities, viral infections, sepsis, and septicaemia OR, (codelist for above) In infant population, at least one bronchiolitis (4120002) or viral wheeze (276191000000107) diagnosis code, in SNOMED CT, in any position in ECDS (A&E), OR, At least one prescription for antibiotics (amoxicillin, doxycycline, trimethoprim, clarithromycin, phenoxymethylpenicillin, azithromycin, erythromycin, cefalexin, ciprofloxacin, amoxicillin / clavulanic acid (co-amoxiclav), levofloxacin, co-trimoxazole, ceftriaxone or moxifloxacin) or antivirals (oseltamivir or zanamivir) in Pseudo BNF relevant for the treatment of infectious respiratory diseases (codelist), with the presence of at least one diagnosis code from maximally sensitive codelist 	<p><i>Inclusion Criteria</i></p> <ul style="list-style-type: none"> At least one ICD-10 diagnosis code for RSV or bronchiolitis in any position in the HES APC (codelist), OR, ICD-10 diagnosis code for unspecified acute lower respiratory infection (J22) in any position in the HES APC <p><i>Exclusion Criteria</i></p> <p>ICD-10 diagnosis code in any position in HES APC for influenza, pneumonia, human metapneumovirus, rhinovirus, adenovirus, parainfluenza, enterovirus, mycoplasma pneumonia, bronchitis, sinusitis, tonsillitis, laryngitis, pertussis, or coronavirus, including COVID-19, within the one-month period (\pm) of respiratory symptoms (codelist)</p>

	<p><i>Exclusion Criteria</i></p> <p>Diagnosis code in SNOMED CT for Influenza, pneumonia, human metapneumovirus, rhinovirus, adenovirus, parainfluenza, enterovirus, mycoplasma pneumonia, bronchitis, sinusitis, tonsillitis, laryngitis, pertussis, or coronavirus, including COVID-19, within the one-month period (\pm) of respiratory symptoms or receipt of a relevant prescription as noted in the inclusion criteria above (codelist)</p>	
<i>Influenza</i>	<i>Mild (Primary Care)</i>	<i>Severe (Secondary Care)</i>
Primary	At least one SNOMED CT code for Influenza (codelist)	At least one ICD-10 diagnosis code for influenza in the first or second position in the HES APC (codelist)
Sensitivity	<p><i>Inclusion Criteria</i></p> <ul style="list-style-type: none"> At least one SNOMED CT code for influenza, suspected influenza, influenza-like-illness, test run for influenza (codelist), OR, Record of an acute respiratory illness (codelist) with a measured temperature of $\geq 38^{\circ}\text{C}$ (codelist) within the same episode (adapted from definition of influenza-like-illness according to WHO (19)) At least one prescription for antivirals (oseltamivir or zanamivir) in Pseudo BNF relevant for the treatment of infectious respiratory diseases (codelist) <p><i>Exclusion criteria</i></p> <p>Diagnosis code in SNOMED CT for RSV, pneumonia, human metapneumovirus, rhinovirus, adenovirus, parainfluenza, enterovirus, mycoplasma pneumonia, tonsillitis, pertussis, or coronavirus, including COVID-19, within the one-month period (\pm) of respiratory symptoms or receipt of a relevant prescription as noted in the inclusion criteria above (codelist)</p>	<p><i>Inclusion Criteria</i></p> <ul style="list-style-type: none"> At least one ICD-10 diagnosis code for influenza in any position in the HES APC (codelist), OR, At least one ICD-10 diagnosis code for acute respiratory illness (codelist) <p><i>Exclusion criteria</i></p> <p>ICD-10 diagnosis code in any position in HES APC for RSV, pneumonia, human metapneumovirus, rhinovirus, adenovirus, parainfluenza, enterovirus, mycoplasma pneumonia, tonsillitis, pertussis, or coronavirus, including COVID-19, within the one-month period (\pm) of respiratory symptoms (codelist)</p>

COVID	Mild (Primary Care)	Severe (Secondary Care)
Primary	At least one SNOMED CT code for COVID-19 (codelist)	At least one ICD-10 diagnosis code for COVID-19 in the first or second position in the HES APC (codelist)
Sensitivity	<p><i>Inclusion criteria</i></p> <ul style="list-style-type: none"> At least one SNOMED CT code for COVID-19, coronavirus (without specific strain), suspected COVID-19, tests performed and referrals to support, OR, At least two of the following symptoms of COVID-19 recorded in SNOMED CT within a two-week period (two on the same day is acceptable): fever, cough, change to sense of smell or taste, shortness of breath, fatigue, body aches/muscle aches, headache, sore throat, rhinitis, loss of appetite, diarrhoea, nausea and vomiting , OR, (codelist for above) At least one prescription for antivirals (nirmatrelvir plus ritonavir (Paxlovid) and molnupiravir (Lagevrio)) in Pseudo BNF (codelist) <p><i>Exclusion criteria</i></p> <p>Diagnosis code in SNOMED CT for RSV, Influenza, pneumonia, human metapneumovirus, rhinovirus, adenovirus, parainfluenza, enterovirus, mycoplasma pneumonia, bronchitis, sinusitis, tonsillitis, pertussis, laryngitis, or specific coronavirus strains which are not COVID-19 within the one-month period (\pm) of respiratory symptoms or receipt of a relevant prescription as noted in the inclusion criteria above (codelist)</p>	<p><i>Inclusion criteria</i></p> <ul style="list-style-type: none"> At least one ICD-10 diagnosis code for COVID-19 in any position in the HES APC (codelist), OR, ICD-10 diagnosis code for Coronavirus as the cause of diseases classified to other chapters (B972) in any position in the HES APC, OR, ICD-10 diagnosis code for Coronavirus infection, unspecified site (B342), in any position in the HES APC <p><i>Exclusion Criteria</i></p> <p>ICD-10 diagnosis code in any position in HES APC for RSV, influenza, pneumonia, human metapneumovirus, rhinovirus, adenovirus, parainfluenza, enterovirus, mycoplasma pneumonia, bronchitis, sinusitis, tonsillitis, pertussis, laryngitis, or specific coronavirus strains which are not COVID-19 within the one-month period (\pm) of respiratory symptoms (codelist)</p>

Overall	Mild (Primary Care, Emergency Attendances)	Severe (Secondary Care)
Sensitive	<p><i>Inclusion criteria</i></p> <ul style="list-style-type: none"> Identified from RSV maximally sensitive inclusion criteria, OR, Identified from influenza maximally sensitive inclusion criteria, OR, Identified from COVID-19 maximally sensitive inclusion criteria, OR, Identified from codelist with non-specific conditions associated with respiratory viruses (codelist), OR, At least one SNOMED CT code, in any position in ECDS (A&E), for LRTI (50417007) or UTRI (54150009), OR, In older adult population, identified as having an exacerbation for COPD (codelist) or Asthma (codelist), in SNOMED CT, OR, In older adult population, acute exacerbation of COPD diagnosis code (195951007), in SNOMED CT, in any position in ECDS (A&E) <p><i>Exclusion criteria</i></p> <p>Diagnosis code in SNOMED CT for metapneumovirus, rhinovirus, adenovirus, parainfluenza, enterovirus, mycoplasma pneumonia or pertussis within the one-month period (\pm) of respiratory symptoms (codelist)</p>	<p><i>Inclusion criteria</i></p> <ul style="list-style-type: none"> Identified from RSV maximally sensitive inclusion criteria, OR, Identified from influenza maximally sensitive inclusion criteria, OR, Identified from COVID-19 maximally sensitive inclusion criteria, OR, Identified from codelist with non-specific conditions associated with respiratory viruses (codelist), in ICD-10, in any position in HES APC, OR, In older adult population, identified as having an exacerbation for COPD (codelist) or Asthma (codelist), in ICD-10, in any position in HES APC <p><i>Exclusion criteria</i></p> <p>ICD-10 diagnosis code in any position in HES APC for metapneumovirus, rhinovirus, adenovirus, parainfluenza, enterovirus, mycoplasma pneumonia or pertussis within the one-month period (\pm) of respiratory symptoms (codelist)</p>

Covariates

For all individuals:

- Age - will be taken at the start of the season in which the participant is included. For infants, the age will be updated monthly (see 'Statistical Analysis' section for more details). We will define age bands as follows:
 - Infants (20):
 - 0-2 months (separated further into 0-1m and 1-2m if accurate ages can be attained)
 - 3-5 months
 - 6-11 months
 - 12-23 months
 - Children and adolescents:
 - 2-5 years
 - 6-9 years
 - 10-13 years
 - 14-17 years
 - Adults:
 - 18-39 years
 - 40-64 years
 - Older Adults:
 - 65-74 years
 - 75-89 years
 - 90+ years
- Sex
- Rurality (in five categories of urban to rural density)
 - Urban major conurbation
 - Urban minor conurbation
 - Urban city and town
 - Rural town
 - Rural village

For infants a subgroup analysis will be conducted for those which can be linked to their birth parent, for these infants the following characteristics will be included:

- Gestational age of the infant where available, grouped as follows (21)
 - <35 weeks
 - 35-36 weeks
 - 37-40 weeks
 - ≥41 weeks

If gestational age is not available on infant records, we will aim to use birth weight instead.

The following maternal characteristics will also be included (where available):

- Age
- Smoking (never, former, current, no information)
- Drinking and drug usage (flag for each)

Following base models, additional models will include:

- Vaccination status against pertussis, during pregnancy of corresponding infant, as a proxy for engagement with health services. This is a per-pregnancy vaccination so for multiple children vaccinations should be linked to the correct infant
- Vaccination status against maternal Influenza vaccination, during pregnancy of corresponding infant, this will be both an indicator for health-seeking behaviour as well as a protective factor relevant for neonatal Influenza risk

And for children and adolescents, adults and older adults additional models will include:

- Prior vaccinations (e.g. COVID-19 and Influenza)
 - Influenza:
 - Vaccinated within one year prior to season start date
 - Vaccinated within current season
 - COVID-19:
 - Time since last vaccination (0-6 months, 6-12 months, 12 months plus or unvaccinated)
 - Vaccinated within current season

Covariate Codelists

<i>Covariate</i>	<i>Codelist</i>
Smoking	Smoking (clear) codelist , Smoking (unclear) codelist
Drinking and drug usage	Hazardous drinking codelist , Illicit substance abuse , Illicit substance abuse intervention , Illicit substance abuse - declined assessment
<i>Exclusion Criteria</i>	<i>Codelist</i>
Severe combined immunodeficiency syndrome	SNOMED-CT code - 31323000, in primary care record
Requiring long-term ventilation	ICD-10 code - Z99.1 for dependence on a ventilator OR ICD-10 code - P28.8 for ventilation, newborn, in secondary care record

Cardiac disease	Cardiac disease codes
Pulmonary hypertension	Pulmonary hypertension codes
Care home resident	Care home residency codes , Temporary care home residency codes , there is also a care home flag in TPP

Statistical Analysis

Exploratory Analysis

A number of exploratory feasibility counts will be conducted to explore:

- Number of mortality outcomes
 - If too few mortality events occur, a composite outcome of severe and mortality will be used
- The number of participants who are infected by the same virus (RSV, Influenza, COVID-19, respectively) more than once within the same season
 - Theoretically, the number of these occurrences should be low which will mean that it is safe to explore one episode per disease per patient per season
- The differences in individuals with multiple outcomes (mild and severe respectively) within one season
 - E.g. are there differences in the individuals with more than one episode across mild and severe outcomes respectively (mild/severe: RSV, influenza, COVID-19, unspecified respiratory infection) when compared to those with only one episode
- Length of stay amongst different individuals

Primary Analysis

Rates per 1000 person-years (5,22,23) of each outcome (see 'Outcomes' section) will be estimated per season in each cohort overall and within categories of ethnicity, socioeconomic status and household size/composition. An additional analysis will be conducted using data combining all seasons to examine how disparities have evolved over the study period. I will then perform a Poisson regression to model the rate of infection over time in different groups. Particularly for the infant cohort, two timescales are of interest: calendar time (to capture seasonality) and infant age. Initial models will use calendar time as the timescale and include infant age as a time-varying covariate, updating monthly. Models will be adjusted for age and sex, as well as the covariates given in the 'Covariates' section. Models including each of the exposures above will be explored.

After investigating risk factors individually, risk factors will be explored in pairs. Finally a full model containing all risk factors will be defined.

There will be seven seasons analysed as follows:

1. September 1st 2016 - August 31st 2017
2. September 1st 2017 - August 31st 2018
3. September 1st 2018 - August 31st 2019
4. September 1st 2019 - August 31st 2020
5. September 1st 2020 - August 31st 2021
6. September 1st 2021 - August 31st 2022
7. September 1st 2022 - August 31st 2023
8. *September 1st 2023 - August 31st 2024 (if available)*

Outcome Phenotypes

For each outcome, phenotypes which are of maximal specificity will be used in an attempt to correctly identify each disease accurately - differentiating between the conditions. These are listed in the 'Outcomes' section for each virus under 'specific'. A second set of phenotypes will be used in a sensitivity analysis to determine how often misclassification occurs when phenotypes are broader, listed in the 'Outcomes' section for each virus under 'sensitive' - see 'Sensitivity Analyses' sections for further detail.

Secondary Analysis

To determine whether there is mediation of risk occurring due to conditions which increase risk of poor outcomes I will conduct a secondary analysis to determine whether some of the disparities identified are mediated by the presence of risk factors.

The secondary analyses will be limited to older adults, this is because they are one of the primary targets for vaccination campaigns and are likely to have the comorbidities of interest. These analyses will be performed on one season of interest per pathogen, looking at a season where the pathogen of interest has had particularly high burden.

From COVID-19 and/or Influenza vaccine eligibility criteria:

- Asthma
- COPD
- Cystic fibrosis
- Other chronic respiratory diseases
- Diabetes
- Addison's disease
- Body Mass Index >40
- Chronic Heart Diseases
- Chronic Kidney Disease
- Chronic Liver Disease
- Chronic Neurological Disease

- Cancer within three years
- Immunosuppression
- Sickle cell disease

Lifestyle factors for adults and older adults:

- Smoking (never, former, current, no information)
 - These data will be missing not at random (MNAR). Current smoking is likely to be well recorded whereas former never is unlikely to be as well recorded.
- Drinking and drug usage

Additional Covariate Codelists

<i>Covariate</i>	<i>Codelist</i>
Asthma diagnosis	Asthma diagnosis codes , Asthma medications (inhalers) , Asthma medications (oral steroids)
Reactive airway disease	SNOMED CT code - 991000119106
COPD	COPD codelist , COPD medications , COPD resolved codes
Cystic fibrosis	Cystic fibrosis codelist
Other chronic respiratory disease	Pulmonary fibrosis codelist , Chronic respiratory disease codes
Diabetes	Diabetes codes , Diabetes resolution codes
Addison's Disease	Addison's disease and hypoadrenalism codelist
BMI	BMI event , BMI stage
Chronic Heart Diseases	Chronic heart disease codelist , Heart failure , Coronary heart disease codelist
Chronic Kidney Disease	Chronic kidney disease - all stages , Chronic kidney disease - stages 3-5 , Chronic kidney disease diagnosis codes (no stage)
Chronic Liver Disease	Chronic liver disease codes
Chronic Neurological Disease	Chronic neurological disease - including significant learning disability
Cancer	Cancer codes (excluding lung and haematological) , Haematological cancer codes , Lung cancer codes
Immunosuppression	Immunosuppression diagnosis codelists , Immunosuppression medications codelist , Receiving chemo or radiotherapy
Sickle Cell Disease	Sickle cell disease codelists

Smoking	Smoking (clear) codelist , Smoking (unclear) codelist
Drinking and drug usage	Hazardous drinking codelist , Illicit substance abuse , Illicit substance abuse intervention , Illicit substance abuse - declined assessment

Software and Reproducibility

Data management will be performed using Python, with analysis carried out using Python/R. Code for data management and analysis as well as codelists archived online.

Power Calculations

OpenSAFELY contains records for approximately 45% of England and is broadly representative, hence power calculations are extraneous detail.

Sensitivity Analyses

To determine whether results are due to decisions made during the analysis process (as detailed above) a number of sensitivity analyses will be conducted.

Outcomes

For the primary analysis, a phenotype which is highly specific will be used to detect each disease (RSV in particular due to the lack of testing available). A more sensitive phenotype will be used as a sensitivity analysis, this phenotype will be broader and have more room for error, so there will be specific exclusion criteria to attempt to prevent incorrect identification.

Subsequently I will conduct an analysis on an overall respiratory outcome using a phenotype combining all maximally sensitive phenotypes with a reduced exclusion criteria codelist.

Season Definition

For the season by season analyses, data will be extracted for one year periods between September-September. Sensitivity analyses will be conducted to determine the impact of 'out-of-season' infections, this will be done by extracting data for October-March for each year.

Example Table Shells

Characteristics

Infants

Characteristic	Category	Number of individuals (col %) 2016/17	Number of individuals (col %) 2017/18	Number of individuals (col %) 2018/19	Number of individuals (col %) 2019/20	Number of individuals (col %) 2020/21	Number of individuals (col %) 2021/22	Number of individuals (col %) 2022/23	Number of individuals (col %) 2023/24
Total									
Age	0-2m								
	3-5m								
	6-11m								
	12-23m								
Sex	Female								
	Male								
Ethnicity	White								
	Mixed								
	Asian or Asian British								
	Black or								

	Black British								
	Other Ethnic Groups								
	Unknown								
IMD Quintile	1 (least deprived)								
	2								
	3								
	4								
	5 (most deprived)								
Rurality	Urban major conurbation								
	Urban minor conurbation								
	Urban city and town								
	Rural town								
	Rural village								

	Unknown								
Subgroup for Which Maternal Linkage is Available									
Total									
Age	0-2m								
	3-5m								
	6-11m								
	12-23m								
Sex	Female								
	Male								
Ethnicity	White								
	Mixed								
	Asian or Asian British								
	Black or Black British								
	Other Ethnic Groups								
	Unknown								
IMD Quintile	1 (least deprived)								

	2								
	3								
	4								
	5 (most deprived)								
Rurality	Urban major conurbation								
	Urban minor conurbation								
	Urban city and town								
	Rural town								
	Rural village								
	Unknown								
Gestational Age	<35w								
	35-36w								
	37-40w								
	≥41w								

	Unknown								
Maternal Age									
Maternal Smoking Status	Never								
	Former								
	Current								
	Unknown								
Maternal Drinking	Yes								
	No								
Maternal Drug Usage	Yes								
	No								
Maternal Pertussis Vaccination Status	Vaccinated during pregnancy								
	No								
Maternal Influenza Vaccination Status	Vaccinated during pregnancy								
	No								

Children and Adolescents

Characteristic	Category	Number of individuals (col %) 2016/17	Number of individuals (col %) 2017/18	Number of individuals (col %) 2018/19	Number of individuals (col %) 2019/20	Number of individuals (col %) 2020/21	Number of individuals (col %) 2021/22	Number of individuals (col %) 2022/23	Number of individuals (col %) 2023/24
Total									
Age	2-5y								
	6-9y								
	10-13y								
	14-17y								
Sex	Female								
	Male								
Ethnicity	White								
	Mixed								
	Asian or Asian British								
	Black or Black British								
	Other Ethnic Groups								

	Unknown								
IMD Quintile	1 (least deprived)								
	2								
	3								
	4								
	5 (most deprived)								
Rurality	Urban major conurbation								
	Urban minor conurbation								
	Urban city and town								
	Rural town								
	Rural village								
	Unknown								
Asthma or reactive airway disease	Yes								

	No								
Time since last COVID-19 vaccination	0-6m								
	6-12m								
	12m+								
Vaccination against COVID-19 in current season	Yes								
	No								
Vaccinated against flu in previous season	Yes								
	No								
Vaccinated against flu in current season	Yes								
	No								

Adults

Characteristic	Category	Number of individuals (col %) 2016/17	Number of individuals (col %) 2017/18	Number of individuals (col %) 2018/19	Number of individuals (col %) 2019/20	Number of individuals (col %) 2020/21	Number of individuals (col %) 2021/22	Number of individuals (col %) 2022/23	Number of individuals (col %) 2023/24
Total									
Age	18-39y								
	40-64y								
Sex	Female								
	Male								
Ethnicity	White								
	Mixed								
	Asian or Asian British								
	Black or Black British								
	Other Ethnic Groups								
	Unknown								
IMD Quintile	1 (least deprived)								

	2								
	3								
	4								
	5 (most deprived)								
Rurality	Urban major conurbation								
	Urban minor conurbation								
	Urban city and town								
	Rural town								
	Rural village								
	Unknown								
Time since last COVID-19 vaccination	0-6m								
	6-12m								
	12m+								
Vaccination	Yes								

against COVID-19 in current season									
	No								
Vaccinated against flu in previous season	Yes								
	No								
Vaccinated against flu in current season	Yes								
	No								
Smoking Status	Never								
	Former								
	Current								
	Unknown								
Drinking	Yes								
	No								
Drug Usage	Yes								
	No								
COPD	Yes								

	No								
Cystic Fibrosis	Yes								
	No								
Other Chronic Respiratory Diseases	Yes								
	No								
Diabetes	Yes								
	No								
Addison's disease	Yes								
	No								
BMI > 40	Yes								
	No								
Chronic Heart Diseases	Yes								
	No								
Chronic Kidney Disease	Yes								
	No								
Chronic Liver Disease	Yes								

	No								
Chronic Respiratory Disease	Yes								
	No								
Chronic Neurological Disease	Yes								
	No								
Cancer within 3 years	Yes								
	No								
Immunosuppression	Yes								
	No								
Sickle cell disease	Yes								
	No								

Older Adults

Characteristic	Category	Number of individuals (col %) 2016/17	Number of individuals (col %) 2017/18	Number of individuals (col %) 2018/19	Number of individuals (col %) 2019/20	Number of individuals (col %) 2020/21	Number of individuals (col %) 2021/22	Number of individuals (col %) 2022/23	Number of individuals (col %) 2023/24
Total									

Age	65-74y								
	75-89y								
	90y+								
Sex	Female								
	Male								
Ethnicity	White								
	Mixed								
	Asian or Asian British								
	Black or Black British								
	Other Ethnic Groups								
	Unknown								
IMD Quintile	1 (least deprived)								
	2								
	3								
	4								
	5 (most								

	deprived)								
Rurality	Urban major conurbation								
	Urban minor conurbation								
	Urban city and town								
	Rural town								
	Rural village								
	Unknown								
Time since last COVID-19 vaccination	0-6m								
	6-12m								
	12m+								
Vaccination against COVID-19 in current season	Yes								
	No								
Vaccinated	Yes								

against flu in previous season									
	No								
Vaccinated against flu in current season	Yes								
	No								
Smoking Status	Never								
	Former								
	Current								
	Unknown								
Drinking	Yes								
	No								
Drug Usage	Yes								
	No								
Asthma	Yes								
	No								
COPD	Yes								
	No								
Cystic Fibrosis	Yes								

	No								
Other Chronic Respiratory Diseases	Yes								
	No								
Diabetes	Yes								
	No								
Addison's disease	Yes								
	No								
BMI > 40	Yes								
	No								
Chronic Heart Diseases	Yes								
	No								
Chronic Kidney Disease	Yes								
	No								
Chronic Liver Disease	Yes								
	No								
Chronic	Yes								

Respiratory Disease									
	No								
Chronic Neurological Disease	Yes								
	No								
Cancer within 3 years	Yes								
	No								
Immunosuppression	Yes								
	No								
Sickle cell disease	Yes								
	No								

Results - For Example Season e.g. 2021/22

Infants

Characteristic	Category	Mild RSV Outcomes (Rate per 1000 person-years)	Severe RSV Outcomes (Rate per 1000 person-y	Mortality Due to RSV (Rate per 1000 person-y	Mild Influenza Outcomes (Rate per 1000	Severe Influenza Outcomes (Rate per 1000	Mortality Due to Influenza (Rate per 1000 person-y	Mild COVID-19 Outcomes (Rate per 1000	Severe COVID-19 Outcomes (Rate per 1000	Mortality Due to COVID-19 (Rate per 1000 person-y	All-cause mortality (Rate per 1000 person-y ears)
----------------	----------	--	---	--	--	--	--	---------------------------------------	---	---	---

			ears)	ears)	person-y ears)	person-y ears)	ears)	person-y ears)	person-y ears)	ears)	
Total											
Age	0-2m										
	3-5m										
	6-11m										
	12-23m										
Sex	Female										
	Male										
Ethnicity	White										
	Mixed										
	Asian or Asian British										
	Black or Black British										
	Other Ethnic Groups										
	Unknow n										
IMD Quintile	1 (least deprive										

	d)										
	2										
	3										
	4										
	5 (most deprive d)										
Rurality	Urban major conurba tion										
	Urban minor conurba tion										
	Urban city and town										
	Rural town										
	Rural village										
	Unknow n										
Subgroup for Which Maternal Linkage is Available											
Total											

Age	0-2m										
	3-5m										
	6-11m										
	12-23m										
Sex	Female										
	Male										
Ethnicity	White										
	Mixed										
	Asian or Asian British										
	Black or Black British										
	Other Ethnic Groups										
	Unknow n										
IMD Quintile	1 (least deprive d)										
	2										
	3										

	4										
	5 (most deprived)										
Rurality	Urban major conurbation										
	Urban minor conurbation										
	Urban city and town										
	Rural town										
	Rural village										
	Unknown										
Gestational Age	<35w										
	35-36w										
	37-40w										
	≥41w										

	Unknown										
Maternal Age											
Maternal Smoking Status	Never										
	Former										
	Current										
	Unknown										
Maternal Drinking	Yes										
	No										
Maternal Drug Usage	Yes										
	No										
Maternal Pertussis Vaccination Status	Vaccinated during pregnancy										
	No										
Maternal Influenza	Vaccinated										

Vaccination Status	during pregnancy										
	No										

Children and Adolescents

Characteristic	Category	Mild RSV Outcomes (Rate per 1000 person-years)	Severe RSV Outcomes (Rate per 1000 person-years)	Mortality Due to RSV (Rate per 1000 person-years)	Mild Influenza Outcomes (Rate per 1000 person-years)	Severe Influenza Outcomes (Rate per 1000 person-years)	Mortality Due to Influenza (Rate per 1000 person-years)	Mild COVID-19 Outcomes (Rate per 1000 person-years)	Severe COVID-19 Outcomes (Rate per 1000 person-years)	Mortality Due to COVID-19 (Rate per 1000 person-years)	All-cause mortality (Rate per 1000 person-years)
Total											
Age	2-5y										
	6-9y										
	10-13y										
	14-17y										
Sex	Female										
	Male										
Ethnicity	White										
	Mixed										
	Asian or Asian										

	British										
	Black or Black British										
	Other Ethnic Groups										
	Unknow n										
IMD Quintile	1 (least deprive d)										
	2										
	3										
	4										
	5 (most deprive d)										
Rurality	Urban major conurba tion										
	Urban minor conurba tion										
	Urban										

	city and town										
	Rural town										
	Rural village										
	Unknown										
Time since last COVID-19 vaccination	0-6m										
	6-12m										
	12m+										
Vaccination against COVID-19 in current season	Yes										
	No										
Vaccinated against flu in previous season	Yes										
	No										

Vaccinated against flu in current season	Yes										
	No										

Adults

Characteristic	Category	Mild RSV Outcomes (Rate per 1000 person-years)	Severe RSV Outcomes (Rate per 1000 person-years)	Mortality Due to RSV (Rate per 1000 person-years)	Mild Influenza Outcomes (Rate per 1000 person-years)	Severe Influenza Outcomes (Rate per 1000 person-years)	Mortality Due to Influenza (Rate per 1000 person-years)	Mild COVID-19 Outcomes (Rate per 1000 person-years)	Severe COVID-19 Outcomes (Rate per 1000 person-years)	Mortality Due to COVID-19 (Rate per 1000 person-years)	All-cause mortality (Rate per 1000 person-years)
Total											
Age	18-39y										
	40-64y										
Sex	Female										
	Male										
Ethnicity	White										
	Mixed										
	Asian or Asian British										

	Black or Black British										
	Other Ethnic Groups										
	Unknown										
IMD Quintile	1 (least deprived)										
	2										
	3										
	4										
	5 (most deprived)										
Rurality	Urban major conurbation										
	Urban minor conurbation										
	Urban city and town										

	Rural town										
	Rural village										
	Unknown										
Time since last COVID-19 vaccination	0-6m										
	6-12m										
	12m+										
Vaccination against COVID-19 in current season	Yes										
	No										
Vaccinated against flu in previous season	Yes										
	No										
Vaccinated against flu in	Yes										

current season											
	No										

Older Adults

Characteristic	Category	Mild RSV Outcomes (Rate per 1000 person-years)	Severe RSV Outcomes (Rate per 1000 person-years)	Mortality Due to RSV (Rate per 1000 person-years)	Mild Influenza Outcomes (Rate per 1000 person-years)	Severe Influenza Outcomes (Rate per 1000 person-years)	Mortality Due to Influenza (Rate per 1000 person-years)	Mild COVID-19 Outcomes (Rate per 1000 person-years)	Severe COVID-19 Outcomes (Rate per 1000 person-years)	Mortality Due to COVID-19 (Rate per 1000 person-years)	All-cause mortality (Rate per 1000 person-years)
Total											
Age	65-74y										
	75-89y										
	90y+										
Sex	Female										
	Male										
Ethnicity	White										
	Mixed										
	Asian or Asian British										
	Black or										

	Black British										
	Other Ethnic Groups										
	Unknown										
IMD Quintile	1 (least deprived)										
	2										
	3										
	4										
	5 (most deprived)										
Rurality	Urban major conurbation										
	Urban minor conurbation										
	Urban city and town										

	Rural town										
	Rural village										
	Unknown										
Time since last COVID-19 vaccination	0-6m										
	6-12m										
	12m+										
Vaccination against COVID-19 in current season	Yes										
	No										
Vaccinated against flu in previous season	Yes										
	No										
Vaccinated against flu in	Yes										

current season											
	No										

Exploratory Tables

- Table to show number of multiple episodes of different viruses within a season within one individual broken down by characteristics (specific phenotypes to begin)

For an example season e.g. 2022-23 for an example cohort e.g. infants

Characteristic	Category	RSV + influenza	RSV + COVID-19	Influenza + COVID-19	RSV + influenza + COVID-19
Age	0-2m				
	3-5m				
	6-11m				
	12-23m				
Sex	Female				
	Male				
Ethnicity	White				
	Mixed				
	Asian or Asian British				
	Black or Black British				
	Other Ethnic Groups				

	Unknown				
IMD Quintile	1 (least deprived)				
	2				
	3				
	4				
	5 (most deprived)				
Rurality	Urban major conurbation				
	Urban minor conurbation				
	Urban city and town				
	Rural town				
	Rural village				
	Unknown				

- Table to show multiple infections of same virus within a season within one individual (specific phenotypes to begin)

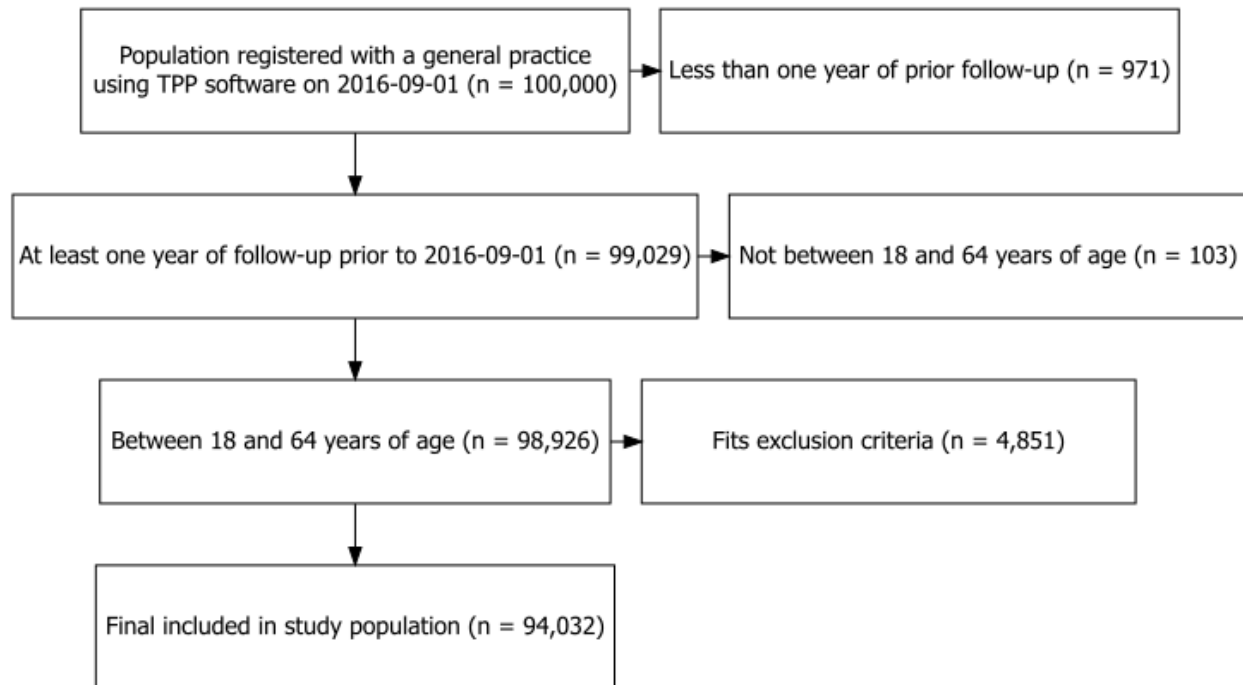
Season	Prop. RSV 2nd	Med. time to 2nd infection	Prop. with 2nd infection under 15-28 days	Prop. COVID-19 2nd	Med. time to 2nd infection	Prop. with 2nd infection under 15-28 days	Prop. influenza 2nd	Med. time to 2nd infection	Prop. with 2nd infection under 15-28 days
2016-17									
2017-18									
2018-19									
2019-20									
2020-21									
2021-22									
2022-23									

- Proportion of those with one episode that go on to have a second

Example Figures

Methods

- Flow diagram of the cohort
<https://www.nature.com/articles/s41586-020-2521-4/figures/1>



(Adult cohort: created using dummy data)

Results

Exploratory Figures

- Figure to show number of multiple viruses within a season within one individual
- Venn diagram showing overlap in outcomes identified dependent on phenotype
- Venn diagram showing how the overall respiratory outcome overlaps with the three viruses

Primary Analysis

- Graphs indicating rate ratio in subgroups ([e.g. figures 3 and 5](#))
 - Outcome by IMD
 - Outcome by ethnicity

- Outcome by household size/composition
- Outcome by pairs
- Outcome by all three exposures
- Rate ratios for each patient characteristic
(<https://www.nature.com/articles/s41586-020-2521-4/figures/3>)

Secondary Analysis

- Graphs indicating rate ratio in subgroups ([e.g. figures 3 and 5](#))
 - Outcome by IMD
 - Outcome by ethnicity
 - Outcome by household size/composition
 - Outcome by pairs
 - Outcome by all three exposures

Sensitivity Analysis

For one season per pathogen:

- Graphs indicating rate ratio in subgroups ([e.g. figures 3 and 5](#))
 - Outcome by IMD
 - Outcome by ethnicity
 - Outcome by household size/composition
 - Outcome by pairs
 - Outcome by all three exposures
- Rate ratios for each patient characteristic
(<https://www.nature.com/articles/s41586-020-2521-4/figures/3>)

Strengths and Limitations

Strengths

This is the first analysis of its kind, looking at differences in potential - and already identified - disparities in the three most resource intensive respiratory viruses. In addition to comparing disparities in these three viruses, this analysis will also explore how these disparities are characterised in different age cohorts.

Limitations

Differentiating between these respiratory viruses, especially in primary care, is particularly difficult and relies on the definition of phenotypes (given in the 'Outcome Phenotypes' section). Determining how much the definition of these phenotypes impacts the results I will conduct a sensitivity analyses using less specific phenotypes which contain more overlap for the three viruses.

Another limitation of this analysis is that it relies on testing in secondary care settings, access to which differs depending on the disease in question, the time period, and location. This will influence how many cases are identified for each virus, especially when using the most specific outcome phenotype.

Administrative

Ethics

Local ethical approval application will be submitted once protocol accepted, as according to LSHTM procedure

London School of Hygiene & Tropical Medicine will act as the main sponsor for this study. Delegated responsibilities will be assigned locally.

Funding

Project undertaken as part of an NIHR funded doctoral fellowship

Conflict of Interests

None

References

1. Munday JD, Pebody R, Atkins KE, van Hoek AJ. Changing socio-economic and ethnic disparities in influenza/A/H1N1 infection early in the 2009 UK epidemic: a descriptive analysis. *BMC Infect Dis.* 2021 Dec 11;21(1):1243.
2. Cromer D, van Hoek AJ, Jit M, Edmunds WJ, Fleming D, Miller E. The burden of influenza in England by age and clinical risk group: A statistical analysis to inform vaccine policy. *J Infect.* 2014 Apr 1;68(4):363–71.
3. GOV.UK [Internet]. [cited 2023 Oct 10]. Chapter 2: disparities. Available from: <https://www.gov.uk/government/publications/technical-report-on-the-covid-19-pandemic-in-the-uk/chapter-2-disparities>
4. 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;584(7821):430–6.
5. Hardelid P, Verfuenden M, McMenamin J, Smyth RL, Gilbert R. The contribution of child, family and health service factors to respiratory syncytial virus (RSV) hospital admissions in the first 3 years of life: birth cohort study in Scotland, 2009 to 2015. *Eurosurveillance.* 2019 Jan 3;24(1):1800046.
6. Wang L, Davis PB, Berger NA, Kaelber DC, Volkow ND, Xu R. Disruption in seasonality, patient characteristics and disparities of respiratory syncytial virus infection among young children in the US during and before the COVID-19 pandemic: 2010-2022 [Internet]. *medRxiv*; 2022 [cited 2022 Dec 30]. p. 2022.11.29.22282887. Available from: <https://www.medrxiv.org/content/10.1101/2022.11.29.22282887v1>
7. Holmen JE, Kim L, Cikesh B, Kirley PD, Chai SJ, Bennett NM, et al. Relationship between neighborhood census-tract level socioeconomic status and respiratory syncytial virus-associated hospitalizations in U.S. adults, 2015–2017. *BMC Infect Dis.* 2021 Mar 23;21(1):293.
8. NHS Digital [Internet]. [cited 2023 Oct 16]. NHS digital, data and technology standards. Available from: <https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/nhs-digital-data-and-technology-standards>
9. NHS Digital [Internet]. [cited 2023 Oct 16]. Data Security and Protection Toolkit assessment guides. Available from: <https://digital.nhs.uk/cyber-and-data-security/guidance-and-assurance/data-security-and-protection-toolkit-assessment-guides>
10. NHS Digital [Internet]. [cited 2023 Oct 16]. ISB1523: Anonymisation Standard for Publishing Health and Social Care Data. Available from: <https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/isb1523-anonymisation-standard-for-publishing-health-and-social-care-data>
11. Coronavirus (COVID-19): notification to organisations to share information - GOV.UK [Internet]. 2020 [cited 2023 Oct 16]. Available from: <https://web.archive.org/web/20200421171727/https://www.gov.uk/government/publication/s/coronavirus-covid-19-notification-of-data-controllers-to-share-information>
12. Respiratory syncytial virus | Treatment summaries | BNF content published by NICE [Internet]. [cited 2023 Nov 20]. Available from:

- <https://bnf.nice.org.uk/treatment-summaries/respiratory-syncytial-virus/>
13. Rybak A, Cohen R, Kramer R, Béchet S, Delobbe JF, Dagrenat V, et al. Respiratory Syncytial Virus in Outpatient Children with Bronchiolitis: Continuous Virus Circulation During the Nonepidemic Period. *Pediatr Infect Dis J*. 2023 Dec;42(12):e488–90.
 14. nhs.uk [Internet]. 2017 [cited 2024 Feb 5]. Bronchiolitis. Available from: <https://www.nhs.uk/conditions/bronchiolitis/>
 15. Wishaupt JO, van der Ploeg T, de Groot R, Versteegh FGA, Hartwig NG. Single- and multiple viral respiratory infections in children: disease and management cannot be related to a specific pathogen. *BMC Infect Dis*. 2017 Jan 11;17(1):62.
 16. Price O, Birrell FA, Mifsud EJ, Sullivan SG. Epidemiology of repeat influenza infection in Queensland, Australia, 2005–2017. *Epidemiol Infect*. 2022 Jan;150:e144.
 17. Molecular Analysis of Respiratory Syncytial Virus Reinfections in Infants from Coastal Kenya | The Journal of Infectious Diseases | Oxford Academic [Internet]. [cited 2024 Jan 4]. Available from: <https://academic.oup.com/jid/article/193/1/59/870060>
 18. Tang CY, Wang Y, McElroy JA, Li T, Hammer R, Ritter D, et al. Reinfection with two genetically distinct SARS-CoV-2 viruses within 19 days. *J Med Virol*. 2021;93(10):5700–3.
 19. Fitzner J, Qasmieh S, Mounts AW, Alexander B, Besselaar T, Briand S, et al. Revision of clinical case definitions: influenza-like illness and severe acute respiratory infection. *Bull World Health Organ*. 2018 Feb 1;96(2):122–8.
 20. Noble M, Khan RA, Walker B, Bennett E, Gent N. Respiratory syncytial virus-associated hospitalisation in children aged ≤5 years: a scoping review of literature from 2009 to 2021. *ERJ Open Res*. 2022 May 30;8(2):00593–2021.
 21. Zylbersztejn A, Pembrey L, Goldstein H, Berbers G, Schepp R, Klis F van der, et al. Respiratory syncytial virus in young children: community cohort study integrating serological surveys, questionnaire and electronic health records, Born in Bradford cohort, England, 2008 to 2013. *Eurosurveillance*. 2021 Feb 11;26(6):2000023.
 22. Wellcome Open Research | Open Access Publishing Platform [Internet]. 2021 [cited 2024 Jan 26]. Ethnic differences in the incidence of clinically diagnosed ... Available from: <https://wellcomeopenresearch.org/articles/6-49/v3>
 23. 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 1;8(5):e364–77.