Title	Sensitivity analysis of design choices for monitoring adverse events of special interest in vaccine safety for BEST
Protocol version identifier	2.0
Date of last version of protocol **Authors**	March 4, 2021 Anna Ostropolets George Hripcsak Patrick Ryan Marc Suchard David Madigan

RATIONALE AND BACKGROUND

There are a number of assumptions, definitions, and parameter choices in monitoring the safety of vaccines. This document is written in preparation for monitoring the safety of COVID-19 vaccines, using administrative open- and closed-claims databases and a variety of electronic health records (EHR) databases.

Safety monitoring requires defining the adverse events of special interest (AESI) in the data sources (phenotyping) and applying methods to detect a possible rise in incidence of AESI following vaccination. One such method is to compare post-vaccination incidence with baseline incidence rates. The baseline rates are sensitive to the exact definition of the cohort in which they are estimated, as well as other assumptions, definitions, and parameter choices. We assess the sensitivity to these factors in this protocol.

SENSITIVITY TO DESIGN CHOICES IN AESI BASELINE INCIDENCE RATE ESTIMATION

We identify the assumptions, definitions, and parameter choices related to estimating baseline incidence rates, and then we specify a set of experiments to estimate the sensitivity to those factors. Our overall research question, is, **how do factors like data sources, outcomes, population subgroups, cohort definitions, anchoring event, and time-at-risk influence baseline incidence rate estimation**?

A. Baseline incidence rate estimation design assumptions

A.i. Data sources

We intend to study data sources with different data source provenance as well data representing different populations (privately insured employed patients in IBM MarketScan Commercial Claims and Encounters [CCAE] or patients with limited income in IBM MarketScan Multi-state Medicaid [MDCD]) and data with different origins (US and non-US).

- Columbia University Irving Medical Center (CUIMC) EHR
- IBM MarketScan Commercial Claims and Encounters (CCAE)
- IBM MarketScan Medicare Supplemental Beneficiaries (MDCR)
- IBM MarketScan Multi-state Medicaid (MDCD)
- IQVIA Disease Analyzer (DA) Germany
- IQVIA Disease Analyzer (DA) France
- IQVIA Australia EMR
- Japan Medical Data Center (JMDC)
- Optum® de-identified Electronic Health Record Dataset (PanTher)
- Optum® de-identified Clinformatics Data Mart Database (SES)
- Clinical Practice Research Datalink (CPRD)

A.ii. Outcomes

For the purposes of this research, we use the outcomes outlined in the U.S. Food and Drug Administration (FDA) "Background Rates of Adverse Events of Special Interest for COVID-19 Vaccine Safety Monitoring" protocol[1]. They include Guillain-Barré syndrome, facial nerve (Bell's) palsy, anaphylaxis, encephalomyelitis, narcolepsy, appendicitis, non-hemorrhagic stroke, hemorrhagic stroke, acute myocardial infarction, myocarditis and pericarditis, deep vein thrombosis, pulmonary embolism, disseminated intravascular coagulation, immune thrombocytopenia, and transverse myelitis. We dropped the two age-limited pediatric outcomes because as currently set up, the experiment will produce invalid incidence rates (as a COVID-19 vaccine is not approved for use in the pediatric population). All phenotype definitions are publicly available in the OHDSI phenotype library and were based on the FDA protocol with minor clinical context broadening (and minor code broadening as needed for example to accommodate SNOMED) to allow them to run on our wide range of data sources; note that our goal is not to estimate actual rates, but to estimate sensitivity of rates.

We consider the effect of clean windows to define the incidence of outcomes. A clean window is the prespecified time prior to the vaccination (reference master protocol), but there are several ways that this can be defined. One choice is an outcome-specific clean window, within which an outcome is not considered incident. The clean window, as defined in the FDA protocol, is 365 days for all outcomes except 30 days for anaphylaxis, 183 days for facial nerve palsy, and encephalomyelitis. Alternatively, the clean window for this analysis may be the first recorded occurrence of the outcome in the data source.

- Clean window
  - Outcome-specific clean window
  - First recorded outcome event

A.iii. Population subgroups

In order to appropriately use a background incidence characterization as context for interpreting the observed rate of adverse events in a defined exposed population, it is important to consider the extent to which the target population that used to estimate the background incidence is generalizable to the exposed individuals. In any observed vs. expected comparison, the comparator serves as a proxy for a counterfactual of the exposed population—what would have happened to those same individuals had they not been exposed—and any deviation between the comparator and that counterfactual represents a potential bias.

One aspect of the target population is the population subgroup. We will consider the following subgroups:

- Age groups
  - 0−5
  - 6−17
  - 18−35
  - 36−55
  - 56−64
  - 65−74
  - 74−85
  - > 85
- Sex
  - Male

- Female
- Race
  - Black
  - White
- Population
  - No requirement
  - 'At-risk' condition in the 365 before entry into the cohort
  - Vaccination in the 365 days before entry into the cohort
  - Pregnancy

'At-risk' condition is defined as at least one condition associated with a higher risk of COVID-19 illness as defined by CDC (removing pregnancy): cancer, chronic kidney disease, chronic obstructive pulmonary disease, heart failure, coronary artery disease, cardiomyopathies, immunocompromised state, obesity, sickle cell disease, smoking, and type 2 diabetes mellitus. We recognize that obesity and smoking are not reliably ascertained in many of these data sources. Pregnancy is assessed separately.

A.iii. Cohort definitions

Another aspect of the target population relates to the cohort definition because it affects the population that is included. The timing of the cohort can have an effect on the population and the incidence rates. For example, medicine and the health of the population change over time, so rates can drift over the years. A pandemic like COVID-19 can affect both health and use of health services. The season affects prevalence of disease and use of health services. We will consider the following time periods:

- Year
  - 2017–2019
  - 2017
  - 2018
  - 2019
  - 3/1/2019–12/31/2019
  - 3/1/2020–12/31/2020 — peri-COVID-19 period
- Season
  - 1/1–3/31 dates
  - 4/1–6/30 dates
  - 7/1–9/30 dates
  - 10/1–12/31 dates

In the above, a COVID-19 interval, 3/1/2020–12/31/2020, is compared to the same period the year before.

The population can be altered by requiring visits during the time period. This is sometimes used in EHRs to ensure that the patient is actually receiving care during the time period. The choice of visit type can affect the population. We will study the following:

- Encounter requirement
  - No required visit
  - Any health care visit

- Well-patient visit

Requiring a pre-entry observation period in the database ensures that a previous outcome event can be observed for an individual. On the other hand, requiring the pre-entry observation period reduces the sample size.

- Pre-entry observation requirement
  - At least 365 days
  - 0-364 days

A.iv. Anchoring events

The entry time into a cohort can have a very large effect on the estimated incidence rate, especially for short times at risk. A patient who just had a medical encounter is more likely to develop other conditions than someone at an arbitrary time (e.g., because of medications ordered). In addition, detection bias may cause previous conditions to be newly recognized and reported. The question for COVID-19 vaccination is whether vaccination will behave more like a medical encounter or more like a random date. This is likely to be affected by the triggering event for the vaccination. Influenza vaccines are often given during other medical encounters, but COVID-19 vaccination may be driven by vaccine availability. Therefore, we study several entry date anchors: an arbitrary date like January 1, a medical visit, or a well-patient visit, which may behave somewhere between the arbitrary date and the medical visit. We also consider vaccination other than COVID-19 vaccines as an anchor to better study whether vaccines behave more like arbitrary dates or medical encounters (using outcomes not expected for those vaccines).

- Anchoring
  - Jan 1 of each year
  - Start of period (for seasons)
  - Random visit
  - Random well visit
  - Vaccination

A.v. Time-at-risk (TAR)

For some outcomes, there is a known pathophysiological relationship between vaccine and outcome that can justify a choice of time at risk [6]. For other outcomes, less is known. In estimating baseline incidence rates, there is a tension between using long time-at-risk windows to maximize the sample size and using windows that match the planned windows in the vaccinated patients. It is therefore essential to know the effect of time at risk on baseline incidence estimation. We will compare the following time-at-risk windows (start -> end offset, relative to target cohort entry date), where the end offset is a maximum and patients with shorter available observation periods are included for the time they are available:

- 0 – 1 days
- 1 – 28 days
- 1 – 42 days
- 1 – 90 days
- 1 – 365 days

I.b. Baseline incidence rate estimation sensitivity experiments

We here list the specific research questions, and then we define the cohorts used to test them.

> *1. Measure the effect of population characteristics on baseline incidence estimates*
> *2. Measure the effect of a history of at-risk conditions*
> *3. Measure the effect of anchoring on baseline incidence estimates*
> *4. Measure the effect of time at risk (TAR) on baseline incidence estimates*
> *5. Measure the effect of season on baseline incidence estimates*
> *6. Measure the effect of peri-COVID-19 time period on baseline incidence estimates*
> *7. Measure the effect of year on baseline incidence estimates*
> *8. Measure the effect of requiring observation before entry on baseline incidence estimates*
> *9. Measure the effect of the clean window on baseline incidence estimates*
> *10. Measure the effect of race on baseline incidence estimates*

We have not included the effect of excluding the outcome from an encounter inclusion requirement because that is not part of our initial testing infrastructure. The purpose of such an exclusion would be to avoid creating bias where patients without the outcome are on average forced to be less healthy (one more visit) than those with the outcome. The encounter inclusion requirement may raise other biases independent of this effect.

The same cohort may address several questions, so they are grouped by cohort below. For each cohort, we compute the following:

- Absolute number of outcomes
- Absolute number of outcomes prior to TAR start date
- Absolute number of patients at risk
- Absolute number of patients with outcomes prior to TAR start date
- Absolute number of patients with outcomes within TAR interval
- Incidence proportion of outcomes
- Incidence rate of outcomes

In the above outcome counts, multiple data elements (e.g., multiple diagnosis codes) within a clean window count as one outcome.

Then for each research question, we compare incidence rates using incidence rate ratios for binary comparisons, and we compare incidence rates for larger comparisons. Exact Poisson confidence intervals[16] are calculated for incidence rates (ignoring the small variance due to the very large aggregate time at risk). For incidence rate ratios (IRRs), the variance of the log(IRR) is given by $(1/n_1 + 1/n_2)$, where n_1 and n_2 are the observed number of outcome events in the two groups, under the null hypothesis that the incidence rates are the same and assuming that the aggregate time at risk is the same in the two groups.

1. Measure the effect of population characteristics on baseline incidence estimates
2. Measure the effect of a history of at-risk conditions

Assess the effect of age group, sex, and data source on the incidence of each outcome. Apply this to the pre-COVID years 2017-2019, with TAR 365 days, entering on each January 1, and ensure that each patient is observed at least 365 days before the entry date.

Assess the effect of a history of at-risk conditions with pregnancy separated out, and look at a history of vaccination as a marker of health-seeking behavior.

```
age = "All ages",
      "Age 0-5",
      "Age 6-17",
      "Age 18-35",
      "Age 36-55",
      "Age 56-64",
      "Age 65-74",
      "Age 75-84",
      "Age 85+"
sex = "All sexes",
      "Male",
      "Female"
race = "All persons"
encounterreq = "No visit req",
               "At-risk condition in 365 days before entry",
               "Vaccination in 365 days before entry ",
               "Pregnancy at time of entry"
anchoring = "Jan 1 of each year"
TAR = "1-365 days"
season = "All seasons"
year = "2017-2019"
preobservation = "At least 365 days"
cleanwindow = "Outcome-specific clean window"
outcome = "Acute myocardial infarction",
          "Anaphylaxis",
          "Appendicitis",
          "Bells palsy",
          "Deep vein thrombosis",
          "Disseminated intravascular coagulation",
          "Encephalomyelitis",
          "Guillian-Barre syndrome",
          "Hemorrhagic stroke",
          "Immune thrombocytopenia",
          "Myocarditis pericarditis",
          "Narcolepsy",
          "Non-hemorrhagic stroke",
          "Pulmonary embolism",
          "Transverse myelitis"
database = "CUIMC",
           "CCAE",
           "MDCR",
           "MDCD",
           "DA Germany",
           "DA France",
           "IQVIA Australia",
```

```
        "JMDC",
        "PanTher",
        "SES",
        "CPRD"
```

3. Measure the effect of anchoring on baseline incidence estimates
4. Measure the effect of time at risk (TAR) on baseline incidence estimates

The cohort entry time may be assigned to an arbitrary date like January 1, or it may be anchored to another event, like a medical encounter. For short times at risk, this can produce a very large effect on outcome baseline incidence rates. We are unsure how vaccination encounters will behave, so we provide three choices, anchored at the beginning of the study period, anchored on a random visit, and anchored on a well visit, where the latter may be a compromise between the first two. (We will compare these three anchorings in the context of a well-visit encounter requirement to keep the populations the same in the comparison. Similarly for vaccination.) These are each tested on a set of TAR, both to confirm the invariance of estimated incidence rates when the anchor is an arbitrary time and to confirm the variance of estimated incidence rates when the anchor is a medical encounter. Apply this to the pre-COVID years 2017-2019, ensure that each patient is observed at least 365 days before the entry date, and stratify by age group because of its large effect.

```
age = "All ages",
      "Age 0-5",
      "Age 6-17",
      "Age 18-35",
      "Age 36-55",
      "Age 56-64",
      "Age 65-74",
      "Age 75-84",
      "Age 85+"
sex = "All sexes"
race = "All persons"
encounterreq = "Visit during the cohort study period",
            "Well visit during the cohort study period",
            "Vaccination the cohort study period"
anchoring = "Jan 1 of each year"
            "Random visit",
            "Random well visit",
            "Vaccination"
TAR = "0-1 days",
      "1-28 days",
      "1-42 days",
      "1-90 days",
      "1-365 days"
season = "All seasons"
year = "2017-2019"
preobservation = "At least 365 days"
cleanwindow = "Outcome-specific clean window"
outcome = "Acute myocardial infarction",
        "Anaphylaxis",
```

                "Appendicitis",
                "Bells palsy",
                "Deep vein thrombosis",
                "Disseminated intravascular coagulation",
                "Encephalomyelitis",
                "Guillian-Barre syndrome",
                "Hemorrhagic stroke",
                "Immune thrombocytopenia",
                "Myocarditis pericarditis",
                "Narcolepsy",
                "Non-hemorrhagic stroke",
                "Pulmonary embolism",
                "Transverse myelitis"
database = "CUIMC",
           "CCAE",
           "MDCR",
           "MDCD",
           "DA Germany",
           "DA France",
           "IQVIA Australia",
           "JMDC",
           "PanTher",
           "SES",
           "CPRD"

5. Measure the effect of season on baseline incidence estimates
6. Measure the effect of peri-COVID-19 time period on baseline incidence estimates

Vaccination will at first be restricted to a single season, so the goal is to measure the effect of season on baseline incidence rates. We will use four calendar points in the year with TAR 90 days. Apply this to the pre-COVID years 2017-2019, ensure that each patient is observed at least 365 days before the entry date, and stratify by age group because of its large effect.
Compare also a period with COVID-19, seasons within the period 3/1/2020-12/31/2020, to a similar span the year before, using at TAR of 90 days.

age = "All ages",
      "Age 0-5",
      "Age 6-17",
      "Age 18-35",
      "Age 36-55",
      "Age 56-64",
      "Age 65-74",
      "Age 75-84",
      "Age 85+"
sex = "All sexes"
race = "All persons"
encounterreq = "No visit req"
anchoring = "Start of period"
TAR = "1-90 days"

season = "1/1-3/31 dates",
        "4/1-6/30 dates",
        "7/1-9/30 dates",
        "10/1-12/31 dates"
year = "2017-2019",
      "2020"
preobservation = "At least 365 days"
cleanwindow = "Outcome-specific clean window"
outcome = "Acute myocardial infarction",
        "Anaphylaxis",
        "Appendicitis",
        "Bells palsy",
        "Deep vein thrombosis",
        "Disseminated intravascular coagulation",
        "Encephalomyelitis",
        "Guillian-Barre syndrome",
        "Hemorrhagic stroke",
        "Immune thrombocytopenia",
        "Myocarditis pericarditis",
        "Narcolepsy",
        "Non-hemorrhagic stroke",
        "Pulmonary embolism",
        "Transverse myelitis"
database = "CUIMC",
        "CCAE",
        "MDCR",
        "MDCD",
        "DA Germany",
        "DA France",
        "IQVIA Australia",
        "JMDC",
        "PanTher",
        "SES",
        "CPRD"

7. Measure the effect of year on baseline incidence estimates

Compare baseline incidence rates among three years, 2017-2019, to look for temporal trends. Ensure that each patient is observed at least 365 days before the entry date and stratify by age group because of its large effect.

age = "All ages",
      "Age 0-5",
      "Age 6-17",
      "Age 18-35",
      "Age 36-55",
      "Age 56-64",
      "Age 65-74",
      "Age 75-84",
      "Age 85+"
sex = "All sexes"

race = "All persons"
encounterreq = "No visit req"
anchoring = "Start of period"
TAR = "1-365 days"
season = "All seasons"
year = "2017",
     "2018",
     "2019"
preobservation = "At least 365 days"
cleanwindow = "Outcome-specific clean window"
outcome = "Acute myocardial infarction",
     "Anaphylaxis",
     "Appendicitis",
     "Bells palsy",
     "Deep vein thrombosis",
     "Disseminated intravascular coagulation",
     "Encephalomyelitis",
     "Guillian-Barre syndrome",
     "Hemorrhagic stroke",
     "Immune thrombocytopenia",
     "Myocarditis pericarditis",
     "Narcolepsy",
     "Non-hemorrhagic stroke",
     "Pulmonary embolism",
     "Transverse myelitis"
database = "CUIMC",
     "CCAE",
     "MDCR",
     "MDCD",
     "DA Germany",
     "DA France",
     "IQVIA Australia",
     "JMDC",
     "PanTher",
     "SES",
     "CPRD"

8. Measure the effect of requiring observation before entry on baseline incidence estimates

Compare baseline incidence estimates among patients with at least 365 days of observation before entry versus patients who do not. Apply this to the pre-COVID years 2017-2019, with TAR 365 days, entering on each January 1, and stratify by age group because of its large effect.

age = "All ages",
    "Age 0-5",
    "Age 6-17",
    "Age 18-35",
    "Age 36-55",
    "Age 56-64",

```
        "Age 65-74",
        "Age 75-84",
        "Age 85+"
sex = "All sexes"
race = "All persons"
encounterreq = "No visit req"
anchoring = "Jan 1 of each year"
TAR = "1-365 days"
season = "All seasons"
year = "2017-2019"
preobservation = "At least 365 days",
                "0-364 days"
cleanwindow = "Outcome-specific clean window"
outcome = "Acute myocardial infarction",
        "Anaphylaxis",
        "Appendicitis",
        "Bells palsy",
        "Deep vein thrombosis",
        "Disseminated intravascular coagulation",
        "Encephalomyelitis",
        "Guillian-Barre syndrome",
        "Hemorrhagic stroke",
        "Immune thrombocytopenia",
        "Myocarditis pericarditis",
        "Narcolepsy",
        "Non-hemorrhagic stroke",
        "Pulmonary embolism",
        "Transverse myelitis"
database = "CUIMC",
        "CCAE",
        "MDCR",
        "MDCD",
        "DA Germany",
        "DA France",
        "IQVIA Australia",
        "JMDC",
        "PanTher",
        "SES",
        "CPRD"
```

9. Measure the effect of the clean window on baseline incidence estimates

Assess the effect of the clean window on the incidence of each outcome. Apply this to the pre-COVID years 2017-2019, with TAR 365 days, entering on each January 1, ensure that each patient is observed at least 365 days before the entry date, and stratify by age group because of its large effect.

```
age = "All ages",
      "Age 0-5",
      "Age 6-17",
```

"Age 18-35",
    "Age 36-55",
    "Age 56-64",
    "Age 65-74",
    "Age 75-84",
    "Age 85+"
sex = "All sexes"
race = "All persons"
encounterreq = "No visit req"
anchoring = "Jan 1 of each year"
TAR = "1-365 days"
season = "All seasons"
year = "2017-2019"
preobservation = "At least 365 days"
cleanwindow = "Outcome-specific clean window",
        "First-ever outcome"
outcome = "Acute myocardial infarction",
        "Anaphylaxis",
        "Appendicitis",
        "Bells palsy",
        "Deep vein thrombosis",
        "Disseminated intravascular coagulation",
        "Encephalomyelitis",
        "Guillian-Barre syndrome",
        "Hemorrhagic stroke",
        "Immune thrombocytopenia",
        "Myocarditis pericarditis",
        "Narcolepsy",
        "Non-hemorrhagic stroke",
        "Pulmonary embolism",
        "Transverse myelitis"
database = "CUIMC",
        "CCAE",
        "MDCR",
        "MDCD",
        "PanTher",
        "SES"

10. Measure the effect of race on baseline incidence estimates

Assess the effect of race on the incidence of each outcome. Apply this to the pre-COVID years 2017-2019, with TAR 365 days, entering on each January 1, ensure that each patient is observed at least 365 days before the entry date, and stratify by age group because of its large effect. This characteristic is separated out because it is absent from many databases and because it is often missing and inaccurate. We need to verify the list of databases with sufficiently accurate race. We focus on black and white here but can expand if sufficient data are available.

age = "All ages",
    "Age 0-5",
    "Age 6-17",

```
    "Age 18-35",
    "Age 36-55",
    "Age 56-64",
    "Age 65-74",
    "Age 75-84",
    "Age 85+"
sex = "All sexes"
race = "All persons",
    "Black",
    "White"
encounterreq = "No visit req"
anchoring = "Jan 1 of each year"
TAR = "1-365 days"
season = "All seasons"
year = "2017-2019"
preobservation = "At least 365 days"
cleanwindow = "Outcome-specific clean window"
outcome = "Acute myocardial infarction",
    "Anaphylaxis",
    "Appendicitis",
    "Bells palsy",
    "Deep vein thrombosis",
    "Disseminated intravascular coagulation",
    "Encephalomyelitis",
    "Guillian-Barre syndrome",
    "Hemorrhagic stroke",
    "Immune thrombocytopenia",
    "Myocarditis pericarditis",
    "Narcolepsy",
    "Non-hemorrhagic stroke",
    "Pulmonary embolism",
    "Transverse myelitis"
database = "CUIMC",
    "CCAE",
    "MDCR",
    "MDCD",
    "PanTher",
    "SES"
```

BIBLIOGRAPHY

1. U.S. Food and Drug Administration Center for Biologics Evaluation and Research. Background Rates of Adverse Events of Special Interest for COVID-19 Vaccine Safety Monitoring, 2021.J. Schuemie M, Soledad Cepede M, A. Suchard M, *et al*. How Confident Are We About Observational Findings in Health Care: A Benchmark Study. *Harvard Data Science Review* Published Online First: 31 January 2020. doi:10.1162/99608f92.147cc28e

2. Schuemie MJ, Ryan PB, Pratt N, *et al*. Principles of Large-scale Evidence Generation and Evaluation across a Network of Databases (LEGEND). *Journal of the American Medical Informatics Association* 2020;**27**:1331–7. doi:10.1093/jamia/ocaa103

3. Maass C, Kuske S, Lessing C, *et al.* Are administrative data valid when measuring patient safety in hospitals? A comparison of data collection methods using a chart review and administrative data. *International Journal for Quality in Health Care* 2015;**27**:305–13. doi:10.1093/intqhc/mzv045

4. Hua-Gen Li M, Hutchinson A, Tacey M, *et al.* Reliability of comorbidity scores derived from administrative data in the tertiary hospital intensive care setting: a cross-sectional study. *BMJ Health & Care Informatics* 2019;**26**:e000016. doi:10.1136/bmjhci-2019-000016

5. Hazell L, Shakir SAW. Under-Reporting of Adverse Drug Reactions: A Systematic Review. *Drug Safety* 2006;**29**:385–96. doi:10.2165/00002018-200629050-00003

6. Farrington P, Rush M, Miller E, *et al.* A new method for active surveillance of adverse events from diphtheria/tetanus/pertussis and measles/mumps/rubella vaccines. *The Lancet* 1995;**345**:567–9. doi:10.1016/S0140-6736(95)90471-9

7. Hak E, Wei F, Grobbee DE, *et al.* A nested case-control study of influenza vaccination was a cost-effective alternative to a full cohort analysis. *Journal of Clinical Epidemiology* 2004;**57**:875–80. doi:10.1016/j.jclinepi.2004.01.019

8. Weldeselassie YG, Whitaker HJ, Farrington CP. Use of the self-controlled case-series method in vaccine safety studies: review and recommendations for best practice. *Epidemiology and Infection* 2011;**139**:1805–17. doi:10.1017/S0950268811001531

9. De Smedt T, Merrall E, Macina D, *et al.* Bias due to differential and non-differential disease- and exposure misclassification in studies of vaccine effectiveness. *PLOS ONE* 2018;**13**:e0199180. doi:10.1371/journal.pone.0199180

10. Shi M, An Q, Ainslie KEC, *et al.* A comparison of the test-negative and the traditional case-control study designs for estimation of influenza vaccine effectiveness under nonrandom vaccination. *BMC Infectious Diseases* 2017;**17**. doi:10.1186/s12879-017-2838-2

11. Guide to Contraindications and Precautions to Commonly Used Vaccines. ;:2.

12. Birth-18 Years Immunization Schedule | CDC. 2020.https://www.cdc.gov/vaccines/schedules/hcp/imz/child-adolescent.html (accessed 17 Nov 2020).

13. Squeri R. HPV vaccine and autoimmune diseases: systematic review and meta-analysis of the literature. *Journal of Preventive Medicine and Hygiene* 2018;**Vol 59**:E194 Pages. doi:10.15167/2421-4248/JPMH2018.59.3.998

14. Tricco AC, Zarin W, Cardoso R, *et al.* Efficacy, effectiveness, and safety of herpes zoster vaccines in adults aged 50 and older: systematic review and network meta-analysis. *BMJ*

15. SAS macro: https://pubmed.ncbi.nlm.nih.gov/1424580/