D7.1.2 Core protocol for type/brand - specific influenza vaccine effectiveness studies (test-negative design studies)

DRIVE 116134-2

DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

[WP7 – Influenza Vaccine Effectiveness Pilot Studies]

Lead contributor	Stefania Bellino (ISS), Ornella Punzo (ISS)
	stefania.bellino@iss.it, ornella.punzo@iss.it
Other contributors	Caterina Rizzo (OPBG), Kaatje Bollaerts (P95), Margarita Riera (P95), Anke Stuurman (P95), Jorne Biccler (P95), Simon de Lusignan (University of Surrey), Uy Hoang (University of Surrey), Harshana Liyanage (University of Surrey), Alexandre Descamps (INSERM), Christine Pereira (INSERM)

Due date	30 June 2020
Delivery date	15 July 2020
Deliverable type	R[1]
Dissemination level	PU[2]

[1] **Use one of the following codes:**
R: Document, report (excluding the periodic and final reports)
DEM: Demonstrator, pilot, prototype, plan designs
DEC: Websites, patents filing, press & media actions, videos, etc.
OTHER: Software, technical diagram, etc.

[2] **Please choose the appropriate reference and delete the rest:**
PU = Public, fully open, e.g. web;
CO = Confidential, restricted under conditions set out in Model Grant Agreement;
CI = Classified, information as referred to in Commission Decision 2001/844/EC.

Document History

Version	Date	Description
V1.1	21/05/2020	First draft
V1.2	02/06/2020	Comments WP7 members
V1.3	12/06/2020	Comments EFPIA members
V1.4	16/06/2020	Comments ISC members
V1.5	08/07/2020	Comments EFPIA/ISC members
V1.6	15/07/2020	Final version

List of abbreviations

DRIVE	Development of Robust and Innovative Vaccine Effectiveness
ECDC	European Centre for Disease Prevention and Control
EEA	European Economic Area
ENCePP	European Network of Centres for Pharmacoepidemiology & Pharmacovigilance
EU	European Union
GEP	Good Epidemiological Practice
GP	General Practitioner
ICD	International Classification of Diseases
IMI	Innovative Medicines Initiative
ILI	Influenza-like illness
IVE	Influenza vaccine effectiveness
OR	Odds ratio
RT-PCR	Real-Time Polymerase Chain Reaction
SARI	Severe Acute Respiratory Infection
TND	Test-negative design
VC	Vaccination coverage
VE	Vaccine effectiveness
WHO	World Health Organization

Preface

The Innovative Medicines Initiative (IMI) project Development of Robust and Innovative Vaccine Effectiveness (DRIVE) aims to create a European platform for studying brand-specific influenza vaccine effectiveness (IVE) and to develop a governance model for scientifically robust, independent and transparent implementation of IVE studies in a public-private partnership.

In DRIVE, data from several independently operating national or regional study sites will be analysed jointly to obtain sufficient geographical coverage and sample size for brand-specific IVE estimates. DRIVE recognizes the value of current study networks and strives to include secondary data from existing studies and initiatives. This is expected to foster European cooperation and maximize the sustainability of the pooled IVE studies.

In 2017/18, a pilot study was performed to test the different operational aspects of the DRIVE project, including the IT infrastructure, the DRIVE governance for conducting IVE studies and to streamline key processes such as data collection, statistical analyses and dissemination of study results. In the pilot study, there were four test-negative design studies (TND) and one register-based cohort study. The DRIVE network is continuously expanding. The 2018/19 season was based on a multi-centre study with data available from five primary care based TND studies, six hospital based TND studies, one register-based cohort and two clinical cohorts (in pregnant women and their young infants and in healthcare workers). For the 2019/20 season, the DRIVE network included 13 TND studies and one register-based cohort study.

This amended generic protocol is intended to be adapted to the local procedures at each individual study site from season 2020/21 onwards. Its aim is to achieve maximum harmonization between the different sites while respecting their different backgrounds. Experience from the pilot studies, together with the completion of other, interconnected DRIVE tasks, will inform the subsequent versions of the protocol.

Specifically, it was found that, despite the continuous efforts of DRIVE to expand the network, challenges are still experienced to achieve a sufficient sample size to perform all stratified analyses in any age groups or population and setting for all vaccine brands. In addition, to increase sample size and capture as many vaccine brands as possible, DRIVE has needed to expand the network with an increasing mix of independent sites and surveillance systems, with a different set of data collections increasing the (clinical) heterogeneity and posing challenges for the interpretation of the pooled analysis. To overcome these challenges and to improve the feasibility of the data collection, the annual study tender will be focused on populations with highest disease burden of influenza and relatively high vaccine coverage in EU and therefore only adults and older adults will be enrolled in hospital setting for the upcoming seasons from sites included through the tender. DRIVE will however continue to incorporate data available from its associate partners, which may still include the pediatric population and primary care.
In addition, the required number of confounders to be collected has been reduced to permit the inclusion of sites, which are unable to expand on the confounder data collection.
In addition, the amendment reflects the anticipated impact of COVID-19 on IVE estimation; for the 2020/21 and potentially beyond. Influenza and Coronavirus disease-2019 (COVID-19) are both respiratory infections which might share similar clinical presentation (e.g. acute respiratory syndrome and influenza-like illness). During the season 2019/20, circulation of SARS-CoV-2 in the EU may have impacted on influenza VE studies in terms of case detection, health care usage, and testing practices though the period of co-circulation of influenza and SARS-CoV-2 was limited.

For the coming season 2020/21 the generic TND protocol has been updated to reflect the experience of the DRIVE pilot studies and to anticipate that SARS-CoV-2 and influenza viruses will both circulate in the next season. In addition, the required number of confounders to be collected has been reduced to accommodate the sites, which are unable to expand on the confounder data collection. Furthermore,

study sites investigating influenza VE will need to collect epidemiological data on the COVID-19 impact in participating countries to address the potential impact of COVID-19 on IVE estimation.

Background

Influenza is a major public health burden. It is responsible for an estimated 50 million disease episodes and 15,000 to 70,000 deaths in the European Union (EU) and European Economic Area (EEA) Member States each year, although with considerable variation from season to season [2] and by methodology used [3]. Complications including deaths are more common in the elderly and in children younger than one year of age [4]. Vaccination is considered as the most effective means for preventing influenza and its complications [5] and the World Health Organization (WHO) has set a vaccination coverage (VC) target of at least 75% in the elderly population and among risk groups [6].

Due to frequent genetic and antigenic changes in influenza viruses, the seasonal vaccine is regularly reformulated (almost annually) to match with the characteristics of the viruses circulating and annual vaccination is recommended.

Observed IVE varies year-to-year due to a variety of reasons including mismatch between the vaccine virus strains and the circulating strains, waning immunity and possible interference from previous vaccinations [6, 7]. In the last two decades, controversies have sprung around the effectiveness of influenza vaccines [8]. While past IVE estimation efforts have led to significant achievements using generic protocols, standard methodologies and laboratory confirmation, several questions about IVE remain open.

In its new guideline on influenza vaccines, the European Medicines Agency (EMA) [9] requires that observational IVE studies be conducted in the EU/EEA as part of the post-licensure requirements of the vaccine manufacturers. Specifically, manufacturers are requested to replace the annual clinical immunogenicity trials (with no clear correlates of protection) with vaccine effectiveness (VE) to provide product (brand) specific data. To reach this goal, manufacturers are encouraged to liaise with organisations/institutions/public health authorities. The studies are expected to be conducted in line with Good Epidemiological Practice (GEP) guidelines and with European Network of Centres for Pharmacoepidemiology & Pharmacovigilance (ENCePP) guidelines.

This document presents the generic DRIVE protocol for the field-based test-negative design (TND) study with patients seeking care for influenza-like illness (ILI) or severe acute respiratory infections (SARI). This TND protocol is intended to be adapted to each individual study site for the next season 2020/21. While each of the study sites can be analyzed separately, pooling them into one analysis is expected to provide a sample size large enough to answer more specific study questions (such as type and age specific VE estimates) with a reasonable/greater precision.

The protocol builds upon the European Centre for Disease Prevention and Control (ECDC) Protocol for case-control studies to measure pandemic and seasonal influenza vaccine effectiveness in the European Union and European Economic Area Member States [10] and the WHO guide to the design and interpretation of observational studies [11]. It will be updated according to the pilot conducted in the participating EU member states, starting from the 2018/2019 season.

For the coming season 2020/21, the generic TND protocol has been updated to account for probable co-circulation of SARS-Cov-2 and influenza viruses during the next season. The study sites already investigating influenza VE will also be able to collect epidemiological data on the COVID-19 impact in participating countries.

The details of each site-specific study will be provided in the study annexes (e.g. ethical committee clearance, data collection strategy, etc.).

Objectives

Primary objective

To estimate seasonal overall, age-specific (6m-17 yr, 18-64 yr, ≥65 yr) and brand-specific IVE against medically attended (primary care/hospital) laboratory-confirmed influenza, by virus type/subtype [A, (A/H1N1, AH3/N2), B (B/Victoria, B/Yamagata)].

Secondary objectives

To estimate seasonal vaccine-type IVE against laboratory-confirmed influenza by virus type/subtype [A, (A/H1N1, AH3/N2), B (B/Victoria, B/Yamagata)]. The following vaccine types will be considered:

- trivalent inactivated vaccine (TIV)
- high-dose TIV
- high-dose QIV
- adjuvanted TIV (aTIV)
- live attenuated quadrivalent egg-based vaccine (LAIV)
- quadrivalent inactivated egg-based vaccine (QIVe)
- quadrivalent inactivated cell-based vaccine (QIVc)

To estimate COVID-19 impact on IVE (SARS-CoV-2 positive vs SARS-CoV-2 negative) within the adults/older adults in hospital setting, given the COVID-19 epidemiology.

Exploratory objectives

To describe clinical signs and symptoms as well as laboratory features, around the point of admission, among hospitalised COVID-19 cases as compared to influenza cases.

The details of the analyses are described in the statistical analysis plan (SAP), updated annually according to the characteristics and structure of data available at the participating study sites.

Methods

Study design

➢ In each participating study site, an observational case-control study using the test-negative design will be conducted

Study setting

The studies will take place in primary care or hospital setting. The study setting is defined by each study site depending on the available data. Each patient with ILI or hospitalized with Severe Acute Respiratory Infection (SARI) will be tested for the detection of both influenza and SARS-CoV-2 viruses.

➢ *Each study site to specify if the study is nested into the national influenza surveillance scheme (the ILI/SARI surveillance system) or is organized differently*

- ➢ *Each study site to specify national policy for influenza surveillance and vaccination and available vaccine brands on the market*
- ➢ *Each study site to specify the target groups for which influenza vaccination is recommended and provide information on vaccination uptake for previous season by age groups*
- ➢ *Each site to specify how they manage influenza and COVID-19 in terms of screening/triage strategy and testing (i.e. with the same swab or not, simultaneous tests, first SARS-CoV-2 test and then influenza or the opposite)*

Study period

The seasonal assessment will start when the influenza virus circulation begins (first week of two consecutive weeks when influenza viruses are detected at the study site level) in the country/region and will finish at the end of the influenza season (the end of the week prior to the first of two consecutive weeks when no influenza viruses are detected at the study site level, or 30th of April, whichever is first).

Study population

The study population consists of patients seeking care (i.e. subjects consulting an emergency department/hospital) for symptoms compatible with ILI/SARI aged 6 months and above, with no contraindication for influenza vaccination.

For the inclusion in the study, the subject will be tested for influenza, and preferably also for SARS-CoV-2. If only influenza test will be done, the record of SARS-CoV-2 testing will be recorded as unknown.

A subgroup analysis will be considered based on the status of SARS-CoV-2 test availability.

Influenza-like illness

A case of influenza like illness (ILI) will be defined by the EU case definition (European Commission Decision of 30 April 2009 amending Decision 2002/253/EC laying down case definitions for reporting communicable diseases to the Community network under Decision N. 2119/98/EC of the European Parliament and of the Council. Luxembourg: Publications Office of the European Union. 1.5.2009. L 110/58) as an individual who presents with a sudden onset of symptoms including at least one of the following four systemic symptoms:

- fever or feverishness;
- malaise;
- headache;
- myalgia;

AND

at least one of the following three respiratory symptoms:

- cough;
- sore throat; and
- shortness of breath.

Severe acute respiratory infection

A case of severe acute respiratory infection (SARI) will be defined by the IMOVE+ 2017/2018 case definition as a hospitalised person with a suspicion of a respiratory infection, with at least one of the following systemic symptoms or signs;

- fever or feverishness;
- malaise;
- headache;
- myalgia;
- deterioration of general condition (asthenia or loss of weight or anorexia or confusion or dizziness)

AND at least one respiratory symptom or sign e.g.

- cough;
- sore throat;
- shortness of breath;

at admission or within 48 hours after admission.

The symptoms should not have started (or, if chronic, clearly worsened) more than 7 days before swabbing.

- *Study sites that do not follow the previous ILI or SARI clinical case definitions should specify the modified definitions (i.e. WHO, modified EU, etc.)*

Clinical information (according to ILI/SARI definition), selected co-morbidities, selected medications, and influenza vaccination status (date of vaccination and brand) will be collected.

COVID-19 case definition

Laboratory criteria

Detection of SARS-CoV-2 nucleic acid in a clinical specimen

Outcomes

The outcome of interest is virologically-confirmed influenza in the study population. More specifically:

- subtype-specific virologically-confirmed influenza A (A/H1N1, A/H3N2),
- virologically-confirmed influenza B overall and if available by lineage (B Victoria/B Yamagata),

Case definition

Primary care studies

- Case: ILI laboratory-confirmed influenza. An ILI patient will be defined as a person in the study population, meeting the ILI - EU case definition with a respiratory sample positive for influenza (see Laboratory testing section).
- Control: ILI negative for Influenza. A control will be defined as an ILI patient in the study population, meeting the ILI - EU case definition for clinical criteria, with a respiratory sample negative for influenza.

The case/control definition will be based only on the influenza test, irrespective from the SARS-CoV-2 result. Patients who will be tested only for COVID-19 and do not have a test result for influenza (within a certain timeframe) will be excluded.

Hospital studies

Case: SARI confirmed as Influenza. A SARI patient will be defined as a person in the study population, meeting the clinical case definition with a respiratory sample positive for influenza.

Control: SARI negative for Influenza. A control will be defined as a SARI patient in the study population, meeting the clinical case definition with a respiratory sample negative for influenza.

Cases and controls will include positive or negative influenza patients irrespective to SARS-CoV-2 testing results. Exploratory analysis will be eventually performed on COVID-19 patients.

Case finding

ILI and SARI patient identification

Patients will be identified among people who present at a healthcare provider (general practitioners (GPs) or hospitals) with ILI or SARI.

- ➢ *Each study site to provide exclusion criteria applied, if different from the list described below*
- ➢ *Each study site to describe procedures to identify study participants*
- ➢ *Any aspect of physician's discretion to decide on inclusion should be described*
- ➢ *Each site should describe how the concepts of random swabbing or systematic swabbing are defined*
- ➢ *Each site should describe the study procedure (i.e. the procedure to screen and test patients)*

Inclusion criteria

ILI/SARI patients are eligible if they accept to participate and do not fulfill any of the exclusion criteria.

Exclusion criteria

The following exclusion criteria will be applied to subjects presenting with ILI:

1. is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)

2. is less than 6 months of age at the time of the onset of the symptoms
3. has a contraindication for influenza vaccine
4. lives in a communal establishment (e.g. in a long-term care facility) at the time of symptoms onset
5. will have the respiratory specimen taken ≥ 8 days after ILI onset
6. tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalization
7. influenza vaccine administration <14 days before ILI symptoms onset

The following exclusion criteria will be applied to subjects presenting with SARI:

1. is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)
2. is less than 6 months of age at the time of the onset of the symptoms
3. has a contraindication for influenza vaccine
4. lives in a communal establishment (e.g. in a long-term care facility) at the time of symptoms onset
5. will have the respiratory specimen taken ≥ 8 days after SARI onset
6. tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalisation
7. was previously hospitalised < 48 hours prior to SARI onset
8. had his/her ILI/SARI onset ≥ 48 hours after hospital admission
9. influenza vaccine administration <14 days before SARI symptoms onset

Exposure

Exposure of interest

The exposure of interest is vaccination with any influenza vaccine in the season under investigation. It is crucial to know precisely the date of vaccine administration, the type/brand of the vaccine, and the date of symptoms' onset as well as the date of specimen collection.

Vaccination status ascertainment

The sources of information for the vaccination status may include:

- vaccination registry
- medical records
- vaccination card

When vaccination status is positive according to any of the above sources but not recalled by patient, they will be coded as vaccinated.

Additional information will be requested if the patient has been vaccinated while undergoing active infection with COVID-19.

➢ *Each study site to describe the precise way of vaccination status ascertainment (i.e. vaccinated yes/no) and vaccine brand and vaccination date ascertainment.*

Definition of vaccination status

An individual aged ≥9 years, or a child aged <9 who has been fully vaccinated (at least two injectable doses or one LAIV dose) during the current influenza season will be considered as

- **vaccinated** with the influenza vaccine of interest if he/she has a record of influenza vaccine administration >14 days before ILI/SARI symptom onset
- **unvaccinated** if he/she has no influenza vaccine record for the current season

A child aged < 9 years who has not been fully vaccinated (see above) during the current influenza season will be considered as

- **vaccinated** with the influenza vaccine of interest if >14 days have elapsed since the second record of injectable vaccination or the first record of LAIV vaccination during the current season
- **unvaccinated** until the first vaccination record during the season

Baseline descriptive, potential confounders and effect modifiers

The following list, based on available literature, presents known and potential confounders and effect modifiers in population-based influenza vaccine effectiveness studies (please also refer to DRIVE D4.1: Framework for analysis of influenza vaccine effectiveness studies).

A parsimonious approach to adjust for confounding will be used, defining a minimum confounder adjustment with the aim to avoid discarding data due to missing values and permit participation of sites who have limited data on confounders.

In the context of the planned inclusion of COVID-19 testing and results to DRIVE objectives for the season 2020/21, this item will be included as a covariate in the sensitivity/exploratory data analysis. To this aim, the following variables will be collected:
- COVID-19 testing result
- Clinical symptoms to distinguish between the 2 diseases (influenza and COVID-19)
- Use of antiviral treatment (treatment for COVID-19 or not), and if yes which
- Co-morbidities to identify risk specific groups for COVID-19

The mandatory set of covariates to be collected for the pooled analysis is marked with an asterisk (*); however, data collection of additional covariates other than the mandatory variables is preferred.

If available and relevant, the other variables may be used in individual study site analyses, and if possible, they will may be harmonised between the study sites for pooled analysis according to the SAP.

- Age*
- Sex*
- Date of symptoms onset*

- COVID-19 positivity in the current season (RT-PCR test)
- Time of COVID-19 test
- COVID-19 positivity in the previous season
- Any chronic underlying conditions*
- If possible, to define the type of chronic condition (like chronic pulmonary disease, cardiovascular disease, diabetes, liver disease, renal disease, neurologic/neuromuscular conditions, treatment-induced immunosuppression and disease-induced immunosuppression (see Annex 1)
- Number of primary care visits 12 months prior to the study period describing a study subject's healthcare seeking behaviour
- Number of hospitalisations 12 months prior to the study period to be used as proxy for the severity of the chronic conditions
- Influenza vaccination in previous influenza seasons (at least one)
- Contraindication to influenza vaccination
- Pregnancy
- Use of influenza antivirals
- Use of COVID-19 treatments
- Use of statins and other selected co-medications
- Pneumococcal vaccination
- Socio-economic status or applicable proxy
- Smoking behaviour or parental smoking behaviour (for subjects ≤18 years)
- Symptoms related to COVID-19 (anosmia, ageusia)
- Morbidity related to COVID-19 (pneumonia)
- Death

➢ *Each study site to describe the covariates included in the study & how these are identified.*

Data collection

Data collection and entry will be conducted at the site level. Data will be collected using a standardised questionnaire/data collection form, administered by clinicians at the moment of swabbing. The questionnaire will be developed before the beginning of the study period according with the list of variables adopted at the study site level.

- *Each study site to describe the data collection tools used*
- *Each study site to describe if and how informed consent is obtained*
- *Each study site to document any protocol violations*

Virological testing

Respiratory specimens will be collected from all eligible patients (ILI and/or SARI). We strongly encourage the use of systematic sampling for primary care studies recruiting ILI (e.g. swabbing the first 3 ILI cases presenting to a GP on the second day of the week of practice) and all hospitalized SARI cases (e.g. all SARI cases presenting at the Emergency department of a Hospital and being admitted to the hospital).

Laboratory confirmation should be done through one of the following laboratory tests: reverse transcription-polymerase chain reaction (RT-PCR, recommended option), viral culture, and immunofluorescence or rapid influenza diagnostic tests. Each positive test result is to be classified by influenza type (A and B) and preferably also subtype/lineage (A/H1N1, A/H3N2, B/Victoria, and B/Yamagata).

In case point of care tests (PoCT) are used,the method should be clearly described. However, the primary analysis should be performed using RT-PCR only, whereas PoCT could be considered in sensitivity analysis.

- *Each study site to describe the specimen collection (i.e. to include a description of the criteria and procedure for swabbing at the site level).*
- *Each study site to describe the specimen storage & transport procedures*
- *Each study site to describe the laboratory tests used & the selection of specimens and the procedures for genetic and antigenic characterisation*
- *Each study site to describe if the laboratory participates in QA/QC (Quality Assurance/Quality Control) schemes*

Sample size considerations

The minimal detectable VE or the smallest VE that can be detected as significantly greater than zero, for a range of samples sizes is given in Table below. The calculations are performed assuming 80% power, two-sided 95% confidence levels and overall vaccination coverages of 5%, 15%, 30% and 50%. It is assumed to have a 1:1 control per case allocation ratio. DRIVE recommends a minimum of 100 influenza positive cases.

As the optimal sample size strongly depends on the local vaccination coverage and brand distribution, site-specific sample size recommendations are formulated as part of the network expansion and site selection. Sample sizes smaller than recommended are allowed as capacity building is an ongoing activity within the DRIVE project. Indeed, the DRIVE objective is not to have robust IVE at site level but ensure a sufficient sample size which allow to increase the power of the pooled analysis.

A user-friendly web-application to perform sample size calculations for IVE studies has been

developed and is available from http://apps.p-95.com/drivesamplesize/.

	Minimum detectable VE			
Number of cases	**5% Coverage**	**15% Coverage**	**30% Coverage**	**50% Coverage**
100	NA	79.3	63.2	56.0
200	91.6	61.6	48.7	43.4
500	65.3	42.4	33.3	30.0
750	55.5	35.6	27.9	25.2
1000	49.2	31.4	24.6	22.2

Data management

Each study site is responsible for the data collection, data validation, and data management of their individual study. DRIVE has developed a generic data management plan (D4.8) and set up the necessary infrastructure for data collection and analysis of the pooled data. To consult such documents go to http://www.drive-eu.org/index.php/results/deliverables/.

- *Each study site to specify how data are collected (e.g. web-based, paper forms) and validated*
- *Each study site to specify procedures of data management.*
- *Each study site to provide a codebook that includes the variable names, variable descriptions, and the coding of variable values, if not following the DRIVE procedures/codebooks/tools.*
- *Each study site to provide any checks in place in the data entry system to avoid mistakes in data entry, and whether source data verification was conducted and how.*
- *Each study site to specify the data checking and cleaning process*

Summary and frequency tables as well as visual representations of appropriate variables will be used to find implausible or missing values within the dataset. Checks for inconsistencies will be carried out (e.g. date of respiratory specimen collection before date of onset of symptoms). Any changes to the data will be documented and stored separately from the crude database. Any additional recording of data during data cleaning phase will be documented. A guide and/or an example file for data cleaning will be provided if needed.

Quality control

The study data will be uploaded by the DRIVE research study sites to the DRIVE Research server using the DRIVE Electronic Study Support Application (DRIVE ESSA). Upon uploading TND data to the ESSA Environment, data quality checks and visualisations are automatically generated and a list with data quality issues can be downloaded by the study site. As such, potential data quality issues can still be solved by the study site before transferring the data to the DRIVE Central Analysis Environment.

The DRIVE ESSA performs 7 different types of quality checks, related to compliance with minimal data requirements, the presence of duplicated records, variable formats and implausible values, inconsistencies between variables and missing values. In addition to the quality checks, the DRIVE ESSA provides seven different data visualizations, summarizing the number of vaccinated subjects over time, the distribution of vaccine brands, the number of cases and controls over time, the age-gender pyramid and the distribution of covariates (sex, age, number of hospitalizations during the last 12 months, and presence of at least 1 chronic condition) among cases and controls.

Representativeness of subjects included in the study

- *Study teams to describe the potential limitations in terms of representativeness of the subjects included*

The study includes ILI and SARI cases. Health-seeking behaviour (referring to how individuals use health services: e.g. the decision to access healthcare, time from onset of illness to consultation, the type of healthcare provider consulted and the adherence to recommended treatment) may differ by country depending on the case management strategy (e.g. recommendation of seeing a GP first). In some cases, the management strategy will have an impact on the delay between onset of symptoms and hospitalisation. This, in turn, may have an impact on the time lag between onset and respiratory specimen collection, and may affect positivity rates between study sites.

Statistical analysis

This section describes the main principles for the study site level analysis. The details of adjustment for confounders and effect modifiers are attempted to be harmonised between the study sites. The variables to adjust for and the heterogeneity/homogeneity between the study sites will be optimised according to availability of data.

For two-stage pooling of data from several study sites, please also refer to the DRIVE D4.4 (Generic statistical analysis plan).

Demographics and baseline characteristics

The baseline characteristics of the study participants will be presented by study setting (i.e. primary care and hospitalized), and will be described and tabulated for cases and controls separately and for vaccinated and unvaccinated subjects within each group (by brand and overall).

Measure of effect

The crude (or unadjusted) IVE will be estimated as

$$VE = (1 - OR) \times 100\%,$$

where *OR* denotes the odds ratio, comparing the odds of vaccination among influenza-positive study participants by the odds of vaccination among influenza-negative study participants. The 95% confidence intervals will be obtained as well.

Confounder-adjusted IVE estimates will be obtained from multivariable logistic regression models, regressing the health outcomes on the confounders of interest. A parsimonious set of confounders will include a smooth function of age, and a smooth function of symptom onset date. For the sensitivity analysis, an extended set of confounders will be used additionally including sex, presence of at least one chronic condition, number of GP visits/hospitalizations, and COVID-19 infection status to allow further exploration of the impact of covariate adjustment on the IVE estimates.

Missing data

The analysis will be a complete case analysis, dropping records with missing information for the outcome, exposure of interest or any of the covariates.

For covariates for which the amount of missing data is substantial (>=10%), multiple imputation methods could be applied assuming that the missingness does not depend on unobserved variables.

Sensitivity analyses

When appropriate, sensitivity analyses may be conducted for the primary and secondary objectives to test the impact of different choices made (e.g. outcome definitions, exposure definitions, subset of the data, influenza testing methods (PCR, rapid test, POCT), time between symptom onset and swab (≤ 4 days vs 4-8 days), COVID-19 positivity and related variables (symptoms, treatments and co-morbidities)).

Adverse events reporting

This is a non-interventional epidemiological study for assessing the effectiveness of routine influenza vaccination. The organization conducting the study will follow local requirements as regards the submission of cases of suspected adverse reactions to the competent authority in the Member State where the reaction occurred.

Ethical evaluation and other relevant approvals

Each study site will comply with the relevant international, national and regional legal and ethics requirements and the declaration of Helsinki and ensures that the ethics committee of the institution has approved the study. Copies of the appropriate approvals from each site will be collected at the study site level and archived according with the local low, but at least for 5 years.
Informed consent will be required from all participants or legal tutors; the national ethics committees will specify whether oral or written consent will be required. The following information should be specified: who is responsible for the study, aim of the study, nature of processed data, purposes of processing, purpose of the use of the data, recipients of possible data transfers, rights of data subject & consequences of not accepting the informed consent.

The only exception is where the study is part of an ongoing routine program evaluation required by ministry of health or a requisite part of the public health institution's work, and would therefore fall outside the mandate for ethics committees. In these cases, a statement that no formal approval from ethics committee is required, is sufficient.

- *Each study site to describe the procedures to comply to the national ethics committee requirements and the type of informed consent needed as well as whether consent can be obtained for a legal tutor.*
- *Each study site to provide a copy of the ethical approval, Independent Review Board or equivalent, or a statement on why this is not needed.*

Dissemination of results

The study site will remain the owner of the data and may disseminate the study results according to their local practices. The data will be submitted to WP7 for pooled analyses across DRIVE sites. EFPIA members do not have access to this data. DRIVE will disseminate the results of its analyses according to its Communications plan (DRIVE D5.4).

Study reports

Each study site will write a report at the end of the season and submit it to DRIVE WP7. DRIVE WP7 will write a final report presenting the results of the pooled estimates.

Publications

Study sites may publish their own data independently from DRIVE. If DRIVE funds were used to collect the data, this should be acknowledged in the publications. Data should be published before the next influenza season.

Authorship of joint DRIVE publications follows the rules of International Committee of Medical Journal Editors (ICMJE).

Logistical aspects

Study sites

A study site is any entity that administers and conducts the individual studies according to the regulations and ethical codes of EU and the country and institutions involved. The study site collects data and provides it to DRIVE. EFPIA members do not have access to this data. Each study site must have a principal investigator responsible of all aspects of the individual study and data transfer to DRIVE WP7. Study sites may be local, regional or national; examples include GP and hospital networks, influenza surveillance schemes and public health institutes utilizing routine health care, social service and demographic databases.

Study leader

In each study site, a study leader (principal investigator) will coordinate the study at the study site level and act as focal point towards DRIVE. The WP7 of DRIVE is in charge of the pooled analyses across DRIVE study sites.

- ➢ *Each study site to introduce the study leader and the study team with brief CVs and Declarations of Interest.*

Standard operating procedures

Standard operating procedures (SOPs) developed and harmonised in DRIVE should be adapted to the individual studies and used by investigators during all the steps of the study for identification of study subjects, data collection, laboratory methods, data entry, monitoring, etc. as provided in DRIVE.

Potential systematic or major deviations from the SOP and study level SAP should be described for further development of the methodology and for interpretation of the results (see D2.1 Standard Operating Procedures and Templates: Guidance and Recommendations). DRIVE Quality Control & Audit Committee (QCAC) will further evaluate the quality of the studies and develop guidelines and methods for improving the quality.

- ➢ *Each study site to adapt DRIVE study SOP to be used by the study team, and provide a summary of systematic or other major deviations from them to WP7, to be stored in order to identify bias and potential confounders for pooling.*

Training

- ➢ *Each study site to describe the trainings to be organized.*

Archiving

Each study site will archive the data used for the analyses, the description of the data (metadata), the study-specific protocol including the analysis plan(s), a description of major deviations from the generic or study-specific protocols, SAP and SOPs, the ethical and other relevant approvals according to the EU level and local regulations, however at least for 5 years.

References

1. World Health Organization. A manual for estimating disease burden associated with seasonal influenza. Geneva, Switzerland, http://apps.who.int/iris/bitstream/10665/178801/1/9789241549301_eng.pdf(2015).

2. Uhart M, Bricout H, Clay E, et al. Public health and economic impact of seasonal influenza vaccination with quadrivalent influenza vaccines compared to trivalent influenza vaccines in Europe. Hum Vaccin Immunother; 12: 2259–2268, https://www.tandfonline.com/doi/full/10.1080/21645515.2016.1180490 (2016).

3. Nicoll A1, Ciancio BC, Lopez Chavarrias V, Mølbak K, Pebody R, Pedzinski B, Penttinen P, van der Sande M, Snacken R, Van Kerkhove MD. Influenza-related deaths--available methods for estimating numbers and detecting patterns for seasonal and pandemic influenza in Europe. Euro Surveill. 2012 May 3;17(18). pii: 20162.

4. Zhou H, Thompson WW, Viboud CG, et al. Hospitalizations Associated With Influenza and Respiratory Syncytial Virus in the United States, 1993–2008. Clin Infect Dis; 54: 1427–1436, https://academic.oup.com/cid/article-lookup/doi/10.1093/cid/cis211 (2012).

5. World Health Organization. Fact sheet Influenza (Seasonal), http://www.who.int/mediacentre/factsheets/fs211/en/ (2016).

6. Osterholm MT, Kelley NS, Sommer A, et al. Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis. Lancet Infect Dis; 12: 36–44, http://www.ncbi.nlm.nih.gov/pubmed/22032844 (2012).

7. Ramsay LC, Buchan SA, Stirling RG, Cowling BJ, Feng S, Kwong JC, Warshawsky BF. The impact of repeated vaccination on influenza vaccine effectiveness: a systematic review and meta-analysis. BMC Med. 2017 Aug 21;15(1):159. doi: 10.1186/s12916-017-0919-0.

8. McCartney M. What use is mass flu vaccination? BMJ; 349: g6182, http://www.ncbi.nlm.nih.gov/pubmed/25331457 (2014).

9. Committee for Medicinal Products for Human Use. Guideline on Influenza Vaccines – Non-clinical and Clinical Module. Eur Med Agency EMA/CHMP/VWP/457259/2014; 44: 1–31, http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2016/07/WC500211324.pdf (2016).

10. Kissling E, Moren A, Valenciano M. European Centre for Disease Prevention and Control Technical Document - Protocol for case-control studies to measure pandemic and seasonal influenza vaccine effectiveness in the European Union and European Economic Area Member Stateshttp://ecdc.europa.eu/en/publications/Publications/0907_TED_Influenza_AH1N1_Measuring_Influenza_Vaccine_Effectiveness_Protocol_Case_Control_Studies.pdf (2009).

11. World Health Organization. Evaluation of influenza vaccine effectiveness: a guide to the design and interpretation of observational studies, http://www.who.int/about/licensing. (2017).

12. de Lusignan S et al. Feasibility of Point-of-Care Testing for Influenza Within a National Primary Care Sentinel Surveillance Network in England: Protocol for a Mixed Methods Study. JMIR Res Protoc. 2019

Annex 1: Dataset requirement

DRIVE – Dataset for pooled data analysis (case-control studies)

Variable	Obligatory	Description	Additional info	Format	Values/coding	Example
idcountry	Obligatory	Country code defined in ISO 3166-1 alpha-2		2 letters text		UK
idstudy	Obligatory	Name of the study		Text		JorviTND
region	Optional	Region name		Text		Wales
idunit	Obligatory (for studies which include >1 GP offices / hospitals)	Identifier of the GP practice or hospital where the patient was seen		Text		JS123
setting	Optional	Type of unit (outpatient, e.g. GP practice, or inpatient, e.g. hospital)		Numeric (Categorical)	1=Outpatient 2=Inpatient 9999=No information	2
id	Obligatory	Patient identification number		Unique integer		101
sex	Obligatory	Sex		Numeric (Binary)	0=Female 1=Male	0
age	Obligatory	Age in years (at the onset of the symptoms)		Numeric		1984

agemonths	Obligatory for children <1 year of age	Age in months (only for children <1 year old)		Numeric		6
onsetdate	Obligatory	Date of symptoms onset		dd/mm/yyyy	Date within the study period	29/12/2017
swabdate	Obligatory	Date of swabbing		dd/mm/yyyy	Date within the study period	30/12/2017
visitdate	Obligatory	Date of visit to the GP or admission to the hospital	In hospital, the first point of contact (often, arrival at the emergency room)	dd/mm/yyyy	Date within the study period	30/12/2017
death	Optional	Has the patient died?	During hospitalization or within 30 days after discharge	Numeric (Binary)	0=Alive 1=Dead	0
deathdate	Optional	Date of death		dd/mm/yyyy	Date within the study period	99/99/9999
fever	Optional	Fever or feverishness	A measured fever of ≥38°C or temperature 37-38°C with patient-reported feverishness	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
headache	Optional	Headache		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
myalgia	Optional	Myalgia		Numeric (Categorical)	0=No 1=Yes 9999=No information	0

malaise	Optional	Fatigue/Malaise		Numeric (Categorical	0=No 1=Yes 9999=No information	1
suddenonset	Optional	Sudden onset of symptoms	Within 7 days before admission	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
cough	Optional	Cough		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
diffbreath	Optional	Difficulty breathing	Subjective evaluation of breathing difficulty by patient or caregiver, or any of the following: respiratory rate ≥25/min (adults) or SpO2 <90% (unless chronic) or PaO2 <8 kPa or respiratory acidosis	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
sorethroat	Optional	Sore throat		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
deterioration	Optional	Deterioration of general condition (asthenia, loss of weight, anorexia, confusion or dizziness)		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
pneumonia	Optional	Pneumonia		Numeric (Categorical)	0=No 1=Yes 9999=No information	1

anosmia	Optional	loss of the sense of smell		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
ageusia	Optional	loss of taste		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
ili	Obligatory (for outpatient)	Influenza like illness	Fulfilling the EU ILI case definition (or local adaptation)	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
covid1	Preferred	COVID-19 disease in the current season	Fulfilling the EU case definition	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
covid2	Preferred	COVID-19 disease in the previous season	Fulfilling the EU case definition	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
coviddatetest	Preferred	Date of COVID-19 test		dd/mm/yyyy	Date within the study period	29/12/2017

sari	Obligatory (for inpatient)	Severe acute respiratory infection	Fulfilling the I-MOVE+ SARI case definition (or local adaptation)			
hosp48h	Obligatory	Was the subject previously hospitalised <48 hours prior to ILI onset		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
contra	Obligatory	Any contraindication for influenza vaccination	Based on locally used criteria.	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
consent	Obligatory	Consent given		Numeric (Categorical)	0=No 1=Yes 9999=Not applicable	1
consentkin	Obligatory	Consent given by family member (or alternatively tutor, where applicable)		Numeric (Categorical)	0=No 1=Yes 9999=Not applicable	1

comm	Optional	Whether communication with the patient OR consent from next of kin was possible.		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
inst	Obligatory	Institutionalized	Living in a residence or nursing home (any such institution where nurse present 24/7)	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
prevflu	Obligatory	Did the patient have a previous lab-confirmed influenza in this season?		Numeric (Categorical)	0=No 1=Yes 9999=No information	0
labvirus1	Obligatory	Laboratory result: influenza virus type		Numeric (Categorical)	0=None 1=A 2=B 3=Other influenza not specified 4=Other virus 9999=No information	2

labsubtype1	Obligatory	Laboratory results: influenza virus subtype		Numeric (Categorical)	0=None 1=A(H1N1)pdm09 2=A(H3N2) 3=B Yamagata 4=B Victoria 5=Other influenza 9=Other virus 9999=No information	3
labvirus2	Optional	Laboratory results: influenza virus type (co-infection)		Numeric (Categorical)	0=None 1=A 2=B 3=Other influenza not specified 4=Other virus 9999=No information	2
labsubtype2	Optional	Laboratory results: virus subtype (co-infection)		Numeric (Categorical)	0=None 1=A(H1N1)pdm09 2=A(H3N2) 3=B Yamagata 4=B Victoria 5=Other influenza 9=Other virus 9999=No information	3

seasvaccany	Obligatory	Received influenza vaccination in current season		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
seasvaccbrand	Obligatory	Vaccine brand		Text		Vaxigrip tetra
seasvaccdate	Obligatory	Date of influenza vaccination in Current season		dd/mm/yyyy		11.1.2018
seasvaccn1	Preferred	Received influenza vaccination in previous season (season n – 1)		Numeric (Categorical)	0=No 1=Yes 9999=No information	
seasvaccn2	Preferred	Received influenza vaccination in season n – 2		Numeric (Categorical)	0=No 1=Yes 9999=No information	
seasvacckid1	Obligatory	Did the kid (< 9 years) receive 1st dose of influenza vaccination in current season?		Numeric (Categorical)	0=No 1=Yes 999=Not applicable 9999=No information	999
seasvacckid2	Obligatory	Did the kid (<9 years) receive 2nd dose of influenza vaccination in current season?		Numeric (Categorical)	0=No 1=Yes 999=Not applicable 9999=No information	999

seasvaccbrand1	Only if Seasvacckid1 is 1	Vaccine brand		Text		Vaxigrip tetra
seasvaccbrand2	Only if Seasvacckid2 is 1	Vaccine brand		Text		Vaxigrip tetra
seasvaccdate1	Only if Seasvacckid1 is 1	Date of 1st dose of influenza vaccination in the current season (only if Seasvacckid1=1)		dd/mm/yyyy	≥Date within the study period	11.1.2018
seasvaccdate2	Only if Seasvacckid2 is 1	Date of 2nd dose of influenza vaccination in the current season (only if Seasvacckid2=2)		dd/mm/yyyy	≥Date within the study period	11.1.2018
pneumovac	Optional	Received any pneumococcal vaccination	Any time.	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
pneumovaccdat	Optional	Date of pneumococcal vaccination	Latest dose.	dd/mm/yyyy		11.1.2018
chronic	Obligatory	Does the patient have at least one chronic disease?	Including obesity (BMI ≥30). Not including smoking or pregnancy.	Numeric (Binary)	0=No 1=Yes 9999=No information	1

liverdis	Optional	Chronic liver disease	Any of the following dg codes (ICD-10): B18, K70-74, K75.0-75.1, K75.3-75.9, K76-77 INCLUDING: Alcoholic liver disease, Toxic liver disease, Hepatic failure, Chronic hepatitis (viral & other), Fibrosis and cirrhosis of liver, Other inflammatory liver diseases, Other diseases of liver EXCLUDING: Clinically insignificant liver cysts	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
diabetes	Optional	Diabetes	Any of the following dg codes (ICD-10): E10-E14, O24 INCLUDING: Any form of diabetes, including sequelae & DM in pregnancy	Numeric (Categorical)	0=No 1=Yes 9999=No information	0

cardiovasc	Optional	Cardiovascular diseases	Any of the following dg codes (ICD-10): A52.0, B37.6, I01-02, I05-09, I11.0, I13.0, I13.2, I20-25, I26-28, I30-43, I44-46, I48, I49.0, I49.5, I50-52, I70-71, Q20-Q28 INCLUDING: all conditions of heart & large vessels that are chronic or likely to have chronic sequelae. Cardiovascular syphilis, endo-, myo- and pericarditis, rheumatic fever, chronic rheumatic heart diseases, congenital malformations, hypertensive (renal) diseases with heart failure, ischaemic heart diseases, diseases of pulmonary circulation, atherosclerosis, cardiomyopathies, most conduction disorders, heart failure, aortic aneurysms & dissecation, other heart diseases and their complications. EXCLUDING: uncomplicated hypertension, previous uncomplicated pulmonary embolism (with no lasting cardiac insufficiency), paroxysmal tachycardias, most cases of premature depolarization.	Numeric (Categorical)	0=No 1=Yes 9999=No information	1

cancer	Optional	Cancer	Any of the following dg codes (ICD-10): C00-97, D37-48, Z85, Z92.3, Z92.6. INCLUDING: All malignant neoplasms (both solid and haematologic) with potential to metastasize, either in treatment, active followup, or <5 years post curative treatment. EXCLUDING: Benign & in situ neoplasms. Basal cell carcinomas. Any cancer previously treated with curative intent & in complete remission for ≥5 years.	Numeric (Categorical)	0=No 1=Yes 9999=No information	0

immuno	Optional	Immunodeficiency or organ transplant	Any of the following dg codes (ICD-10): B20-B24, D80–84, D89, Z94 INCLUDING: HIV infections, immunodeficiencies & organ transplants. or iatrogenic: ≥2 week systemic treatment, in the 3 months preceding symptom onset, with any of the following: corticosteroid (≥20 mg prednisolone daily or equivalent), ciclosporin, tacrolimus, mycophenolate, methotrexate, azathioprine, TNF-α blockers and other biological or cytostatic drugs with immunosuppressive effect EXCLUDING: Disorders of the immune system which do not lead to immunosuppression (e.g. some autoimmune conditions).	Numeric (Categorical)	0=No 1=Yes 9999=No information	0

lungdis	Optional	Lung disease	Any of the following dg codes (ICD-10): A15-16, A19, A31.0, B33.4, E84.0, J40-47, J60-70, J80-84, J85-86, J90-91, J92.9, J93-94, J95-99 INCLUDING: TB (pulmonary, miliary but not that of other systems), atypical mycobacteria, cystic fibrosis, asthma, COPD, bronchiectasis and other chronic sequelae of infections, chronic lung diseases due to external agents, interstitial lung diseases, pleural diseases, respiratory failure. EXCLUDING: acute respiratory infections, lung cancer, diseases of pulmonary circulation, pleural plaques without asbestos, previous uncomplicated pneumothorax.	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
anemia	Optional	Anemia	Any of the following dg codes (ICD-10): D50-D64 diagnosed before the onset of symptoms. EXCLUDING: coagulopathies, uncomplicated hypersplenism, hepato/splenomegaly (D65-69, D70-77, D80-84, D86, D89)	Numeric (Categorical)	0=No 1=Yes 9999=No information	0

rendisease	Optional	Renal disease	Any of the following dg codes: (ICD-10): I12-13, M10.30, N00-19, N20.0, N25-27, N28.0, N28.9, Q63.9, Z90.5 EXCLUDING: Clinically nonsignificant kidney cysts	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
dement	Optional	Dementia	Any of the following dg codes (ICD-10): F00-03, F05.1, G30-31 EXCLUDING delirium w/o underlying dementia, hydrocephalus.	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
stroke	Optional	History of stroke	Any of the following dg codes (ICD-10): I61-64, I67.8, I69, G93.1 INCLUDING: both ischaemic and haemorrhaegic strokes and anoxic brain damage. Also counting previous episodes and clear ischaemic findings seen in cranial imaging (even if fully recovered / no symptoms).	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
rheumat	Optional	Rheumatologic diseases	Any of the following dg codes: ICD-10: M05–09, M13, M30–36, M45 INCLUDING rheumatoid diseases with presumed autoimmune origin and primarily musculoskeletal presentation. EXCLUDING: arthrosis, gout, scoliosis, infectious conditions etc.	Numeric (Categorical)	0=No 1=Yes 9999=No information	0

obesity	Optional	Obesity	BMI ≥30 or the dg codes (ICD-10): E66, E68 EXCLUDING: local adiposity and "other hyperalimentation" (=vitamin overdoses etc.)	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
childrisk	Optional	In children: Any perinatal or congenital risk factor?		Numeric (Categorical)	0=No 1=Yes 9999=No information	2
nhosp	Preferred	Number of hospitalizations in the last year	Any overnight stay in hospital. (One disease episode counts as one hospitalization even if a patient is moved from one unit to another)	Numeric	≥0 or 9999=No information	2
gpvisit	Preferred (for GP studies)	Number of GP consultations in the last year	Any consultation to nurse/GP/specialist in a primary care setting. Not counting follow-up visits for the same cause.	Numeric	≥0 or 9999=No information	5
antiviral_flu	Optional	Has the patient received an antiviral treatment for influenza within the 2 weeks before swabbing?		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
antiviral_covid	Optional	Has the patient received an antiviral treatment for COVID-19 within the 2 weeks before swabbing?		Numeric (Categorical)	0=No 1=Yes 9999=No information	1

statin	Optional	Statin use	At the time of vaccination.	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
pregnancy	Optional	Pregnancy	Any trimester at symptom onset.	Numeric (Categorical)	0=No 1=Yes 9999=No information	0

hcw	Optional	Is the patient a healthcare worker?		Numeric (Categorical)	0=No 1=Yes 9999=No information	0
siblings	Optional	(In children) Number of siblings		Numeric	≥0 or 9999=No information	2
bmi	Optional	Body Mass Index		Numeric	10 to 55 or 9999=No information	22,4
smoking	Optional	Smoking status (cigarettes, cigars, pipe, hookah). Not counting exclusively chew tobacco or snus.	Never-smoker: <100 cigarettes during their lifetime. Ex-smoker: has smoked ≥100 cigarettes over lifetime but has stopped ≥3 months ago. Occasional smoker: has smoked ≥100 cigarettes over lifetime and has still smoked in the 3 months preceding symptom onset, but not daily. Daily smoker: has smoked ≥100 cigarettes over lifetime and smokes daily.	Numeric (Categorical)	0=Never-smoker 1=Ex-smoker 2=Occasional smoker 3=Daily smoker 9999=No information	0

functstatus	Optional	Dependency / Patient has difficulty in at least 1 of these categories: bathing dressing eating going to the toilet stairs walk wheelchair user	Difficulty = needs help from others	Numeric (Categorical)	0=No 1=Yes 9999=Not applicable	0

Annex 2: Sample size considerations for case-control studies

Authors: Kaatje Bollaerts and Maria Alexandridou

For questions or feedback, please contact
e-mail: kaatje.bollaerts@p-95.com

This document provides sample size estimations for estimating overall and brand-specific influenza vaccine effectiveness (VE) using the case-control design. The minimal detectable VE as well as precision estimates are provided for various parameter settings and recommendations are formulated.

Minimal detectable vaccine effectiveness

The minimal detectable VE is the smallest VE that can be detected as significantly greater than zero in a given study using hypothesis testing. The minimal detectable VE for a case-control study is estimated as

$$VE_{MD} = 1 - {}_{(RR<1)}, \quad (1)$$

where $RR_{MD(RR<1)}$ is the minimal detectable relative risk (RR) if RR < 1, or

$${}_{(RR<1)} \cong 1 + \frac{-b - \sqrt{b^2 - 4(r+1)}}{2a}, \quad (2)$$

where

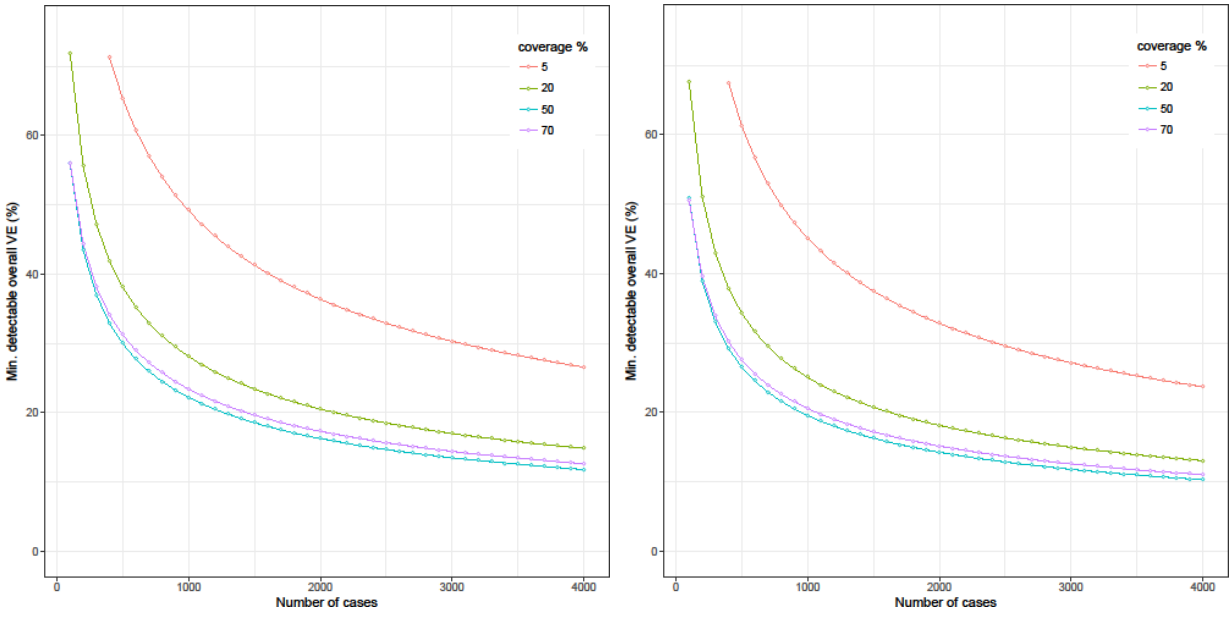
$$= r\gamma^2 - \frac{(1-\gamma)^2}{(z_{\frac{\alpha}{2}} + z_\beta)\,(r+1)}; \; b = 1 + 2r\gamma,$$

for 'cases to controls' ratio r, coverage γ, total number of subjects N, and where z_α and z_β are the standard normal z-scores for the type I and type II error rates (Woodward 2013).

We calculated the minimal detectable overall VE (1) with 80% power (1 – β) and a two-sided 95% confidence coefficient (1 – α/2) for case-control studies using 'cases to controls' ratio of 1:1, 1:2 and 1:4 with the number of cases varying from 100 to 4000, while assuming overall vaccination coverages of 5%, 20%, 50% and 70%.

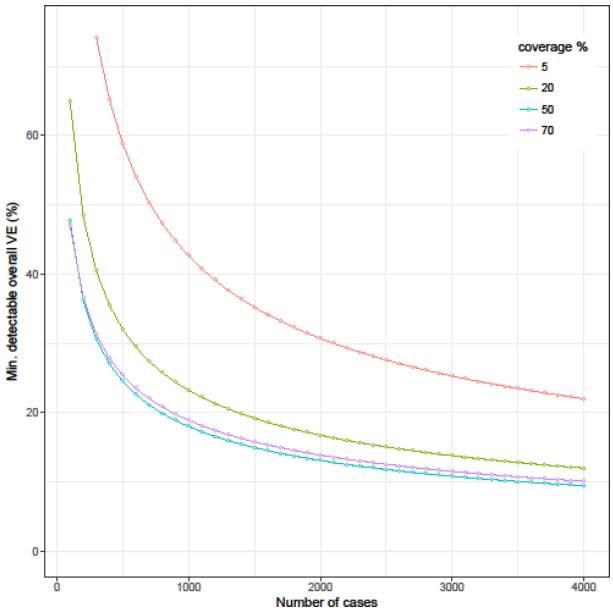
We additionally calculated the minimal detectable brand-specific VE, where cases/controls are considered exposed when they were vaccinated with the brand of interest and unexposed when they were unvaccinated. This means that subjects vaccinated with another brand are excluded from the analysis and that the same comparator group of unexposed subjects is used for the different brand-specific estimates. The minimal detectable brand-specific VE is calculated for the same settings above, additionally assuming that the brand of interest accounts for 10%, 20%, 40%, 60%, 80% and 90% of the overall vaccination coverage.

The results for the minimal detectable overall VE for the 1:1, 1:2 and 1:4 'cases to controls' ratios are given in Figure 1. These figures represent the minimal detectable VE by number of cases. The results for the minimal detectable brand-specific VE for the 1:1 'cases to controls' ratio and assuming overall vaccination coverages of 5%, 20%, 50% and 70% are given in Figure 2.



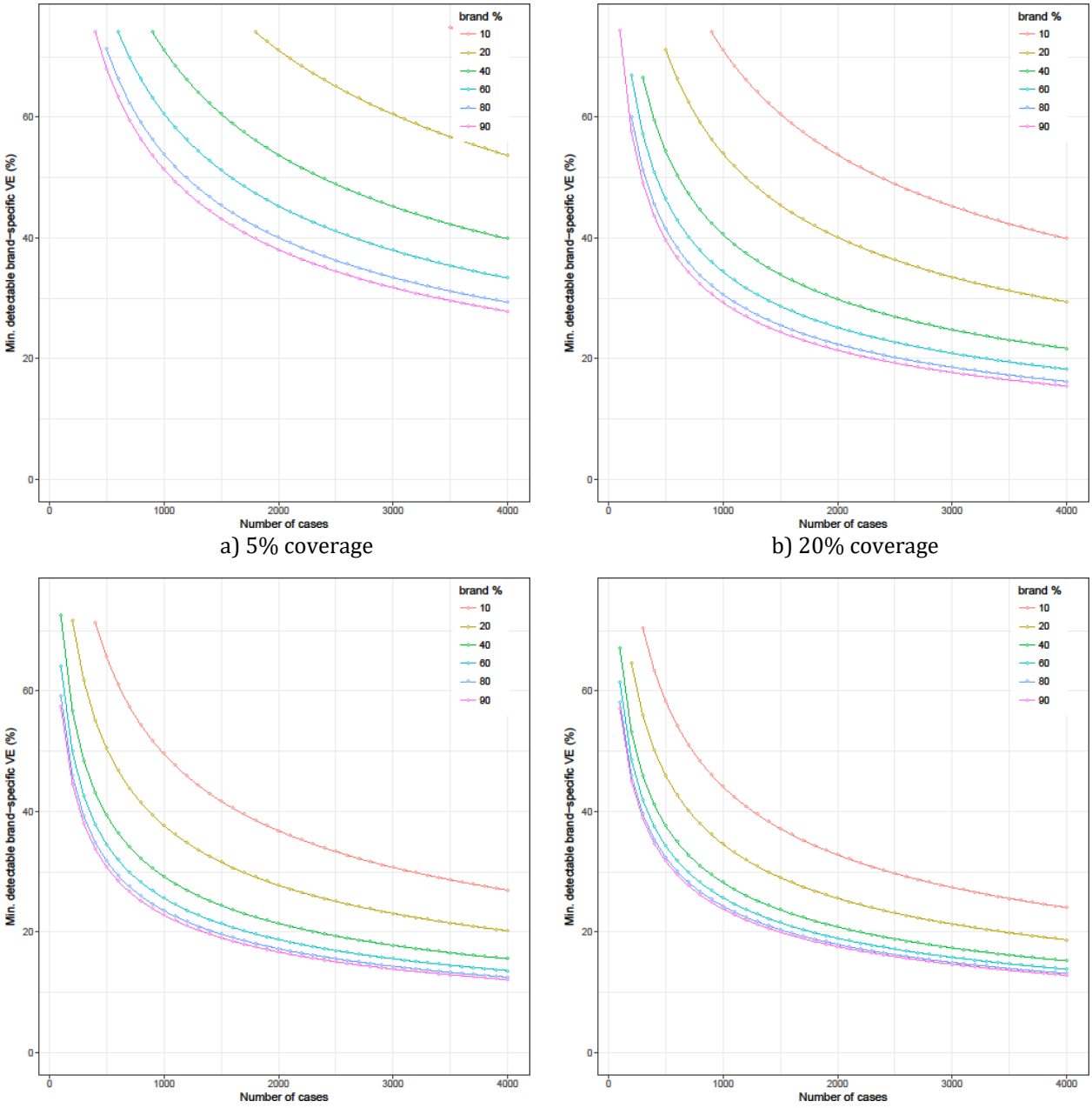
a) 1:1 cases to controls

b) 1:2 cases to controls



c) 1:4 cases to controls

Figure 1. Minimal detectable overall vaccine effectiveness for a case-control study (1:1, 1:2 and 1:4 cases to controls ratio) assuming vaccination coverage of 5%, 20%, 50% and 70% by number of cases.



a) 5% coverage

b) 20% coverage

c) 50% coverage

d) 70% coverage

Figure 2. Minimal detectable brand-specific vaccine effectiveness for a case-control study (1:1 cases to controls ratio) assuming 5%, 20%, 50% and 70% overall vaccination coverage with the brand of interest covering 10%, 20%, 40%, 60%, 80% and 90% of the overall coverage.

Precision

The precision refers to the level of sampling error. The standard error and consequently the width of confidence intervals (CI) are measures of precision. As the VE CIs are asymmetric, we express precision as the lower limit of the two-sided CI of the anticipated true VE, expressed in %. The precision can be derived starting from the anticipated true VE, the confidence coefficient (1 – α/2), the number of cases, the 'cases to controls' ratio 1:*r* and the overall vaccination coverage γ. Consider the notation as defined in Table 1, where N is the total number of subjects, N^+ the number of cases, N^-_d the number of controls, N^+_e the number of vaccinated subjects, N^-_e the number of unvaccinated subjects and where *r* is the number of controls per case and γ is the coverage.

Table 1: Cross-tabulation of exposure and disease in a case-control study

		Diseased		
		Yes (cases)	No (controls)	
Exposed	Yes	a	b	$N^+_e = N\,\gamma$
	No	c	d	$N^-_e = N(1-\gamma)$
		N^+_d	$N^-_d = r\,N^+_d$	N

Then, from the lower limit of the CI for VE estimates based on a case-control study, or

$$VE_{LL\,CI} = 1 - exp\left[\log(OR) + Z_{\alpha/2}\sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}\right], \quad (3)$$

where $OR = 1 - VE$ and where $Z_{\alpha/2}$ is the standard normal z-score, it follows that the precision is determined for given values for *a, b, c* and *d*. From anticipated values for *OR*, N^+_e, N^-_e and N^+_d, the

cell count *a* can be analytically derived as;

$$a = \left(\frac{1}{2}\right)\sqrt{\frac{x_1 + x_2 + x_3}{(OR-1)^2}} + \frac{N_e^+ OR + N_e^- + N_d^+ OR - N_d^+}{2(OR-1)},$$

$$b = N_e^+ - a$$

$$c = N_d^+ - a$$

$$d = N_d^- - b$$

where

$$x_1 = N_e^{+2} OR^2 + 2N_e^+ N_e^- OR - 2N_e^+ N_e^+ OR^2 + 2N_e^+ N_d^+ OR$$

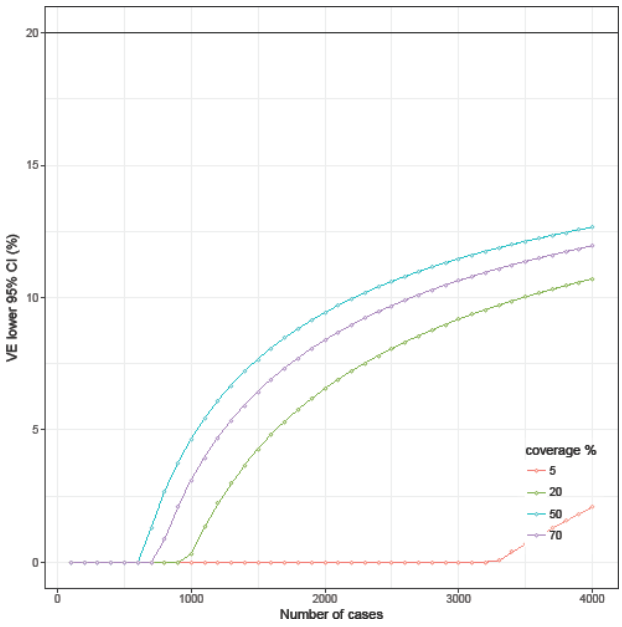
$$x_2 = N_e^{-2} + 2N_e^- N_d^+ OR - 2N_e^- N_e^+ - N_d^+$$

$$x_3 = N_d^{+2} OR^2 - 2N_d^{+2} OR + N_d^{+2}$$

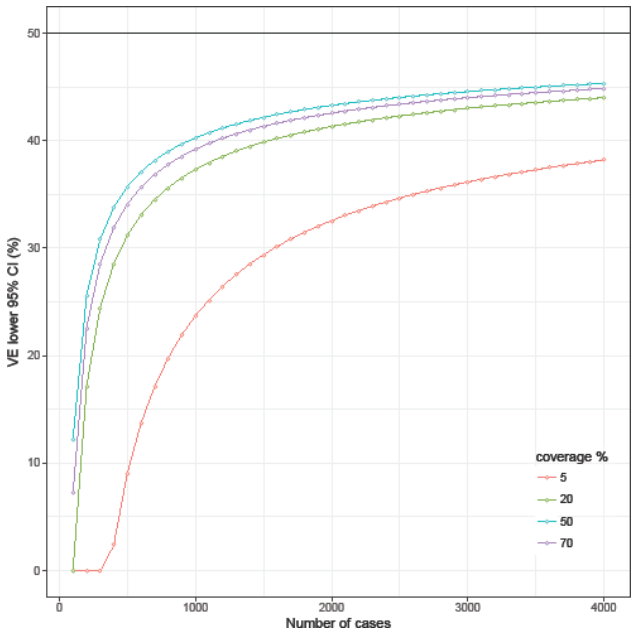
We calculated the precision of the overall VE based on a two-sided 95% CI for case-control studies using 'cases to controls' ratio of 1:1, 1:2 and 1:4 with the total number of cases varying from 100 to 4000, while assuming overall vaccination coverages of 5%, 20%, 50% and 70% and overall VE of 20%, 50% and 70%.

We additionally calculated the precision of the brand-specific VE, where cases/controls are considered exposed when they were vaccinated with the brand of interest and unexposed when they were unvaccinated. This means that subjects vaccinated with another brand are excluded from the analysis and that the same comparator group of unexposed subjects is used for the different brand-specific estimates. The precision of brand-specific VE is calculated for case-control studies using a 'cases to controls' ratio of 1:1 using the same settings as above, additionally assuming that the brand of interest accounts for 10%, 20%, 40%, 60%, 80% and 90% of the overall vaccination coverage.

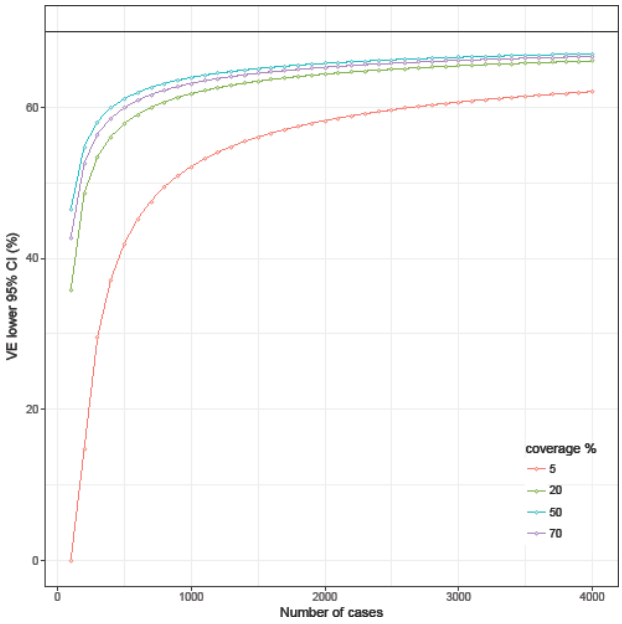
The results for precision of the overall VE using 'cases to controls' ratio of 1:1, 1:2 and 1:4 are given in Figure 3 to 5, respectively. These figures represent precision by number of cases. The results for the precision of brand-specific VE using 'cases to controls' ratio of 1:1 for anticipated true VE of 20%, 50% and 70% are given in Figures 6 to 8.



a) 20% VE

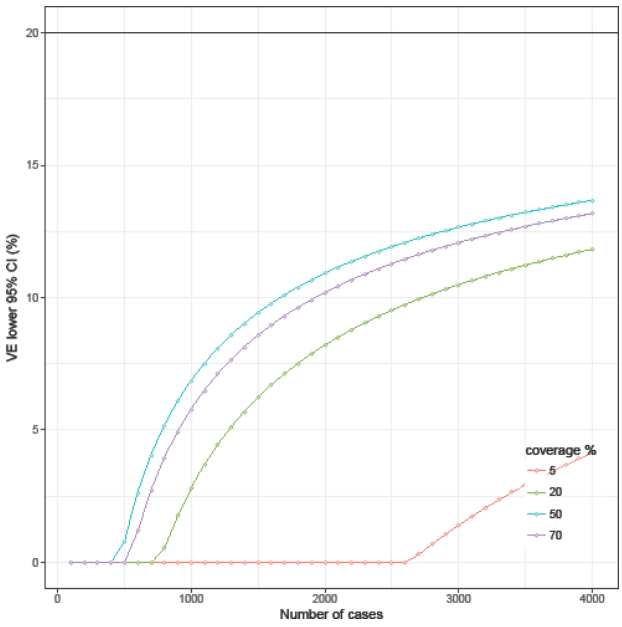


b) 50% VE

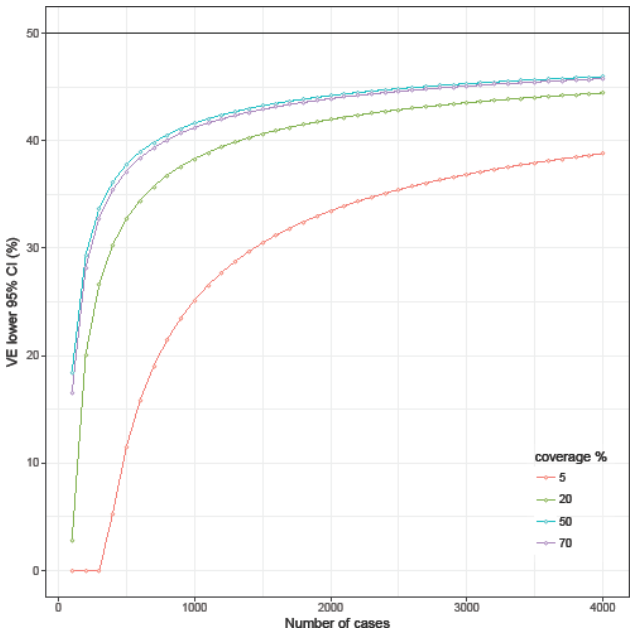


c) 70% VE

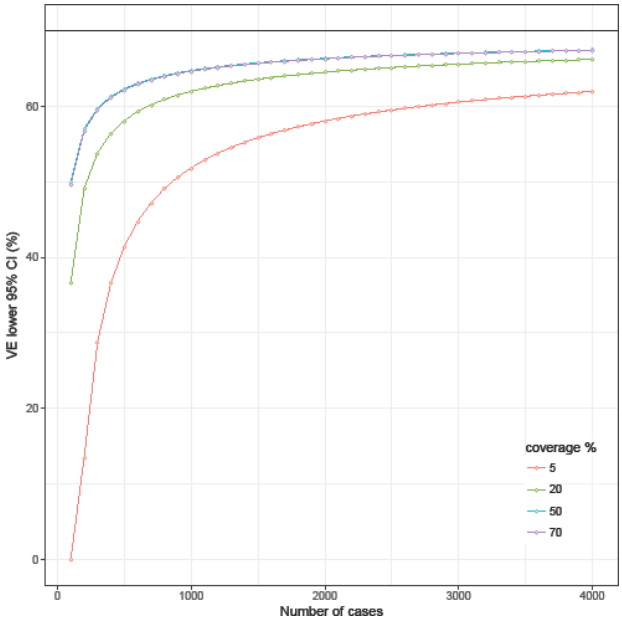
Figure 3. Precision of overall VE for a case-control study (1:1 case-control ratio) assuming overall vaccination coverage of 5%, 20%, 50% and 70%, and anticipated true VE of 20%, 50% and 70% (indicated with the black horizontal line), by number of cases.



a) 20% VE

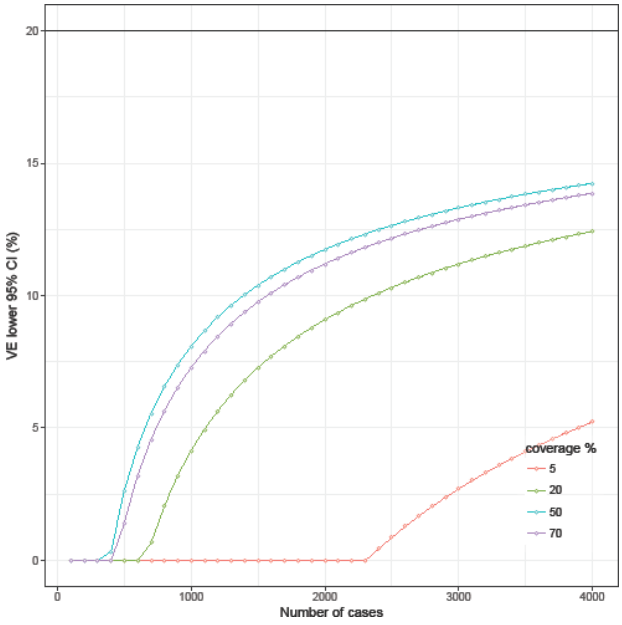


b) 50% VE

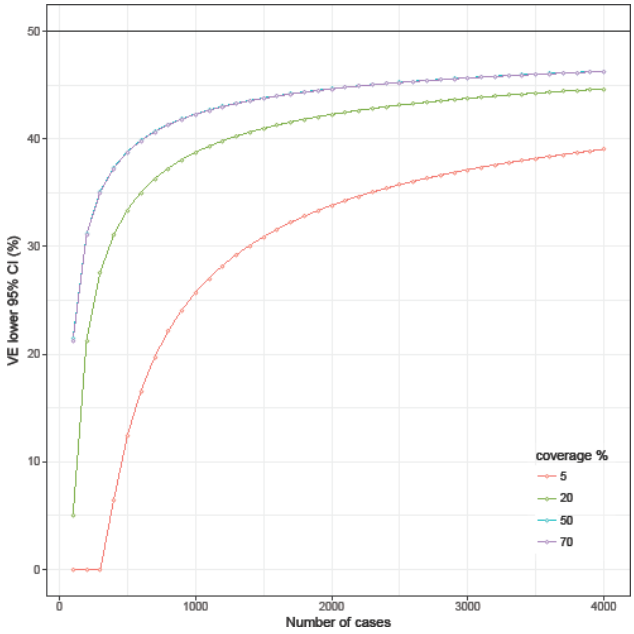


c) 70% VE

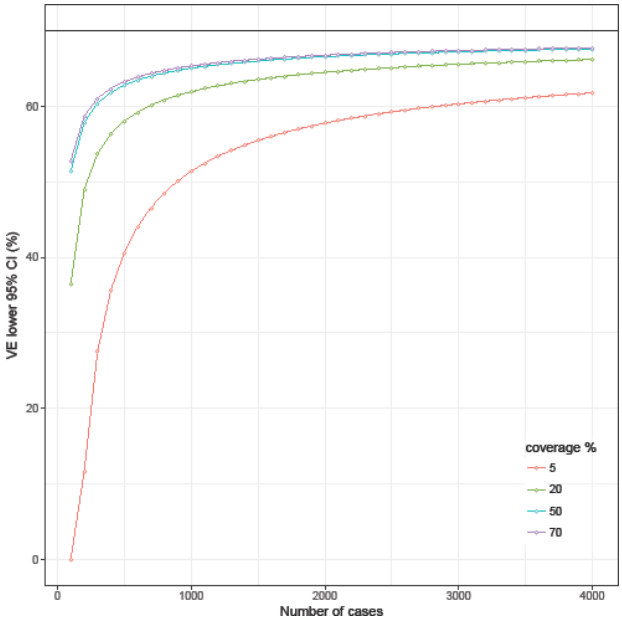
Figure 4. Precision of overall VE for a case-control study (1:2 case-control ratio) assuming overall vaccination coverage of 5%, 20%, 50% and 70%, and anticipated true VE of 20%, 50% and 70% (indicated with the black horizontal line), by number of cases.



a) 20% VE

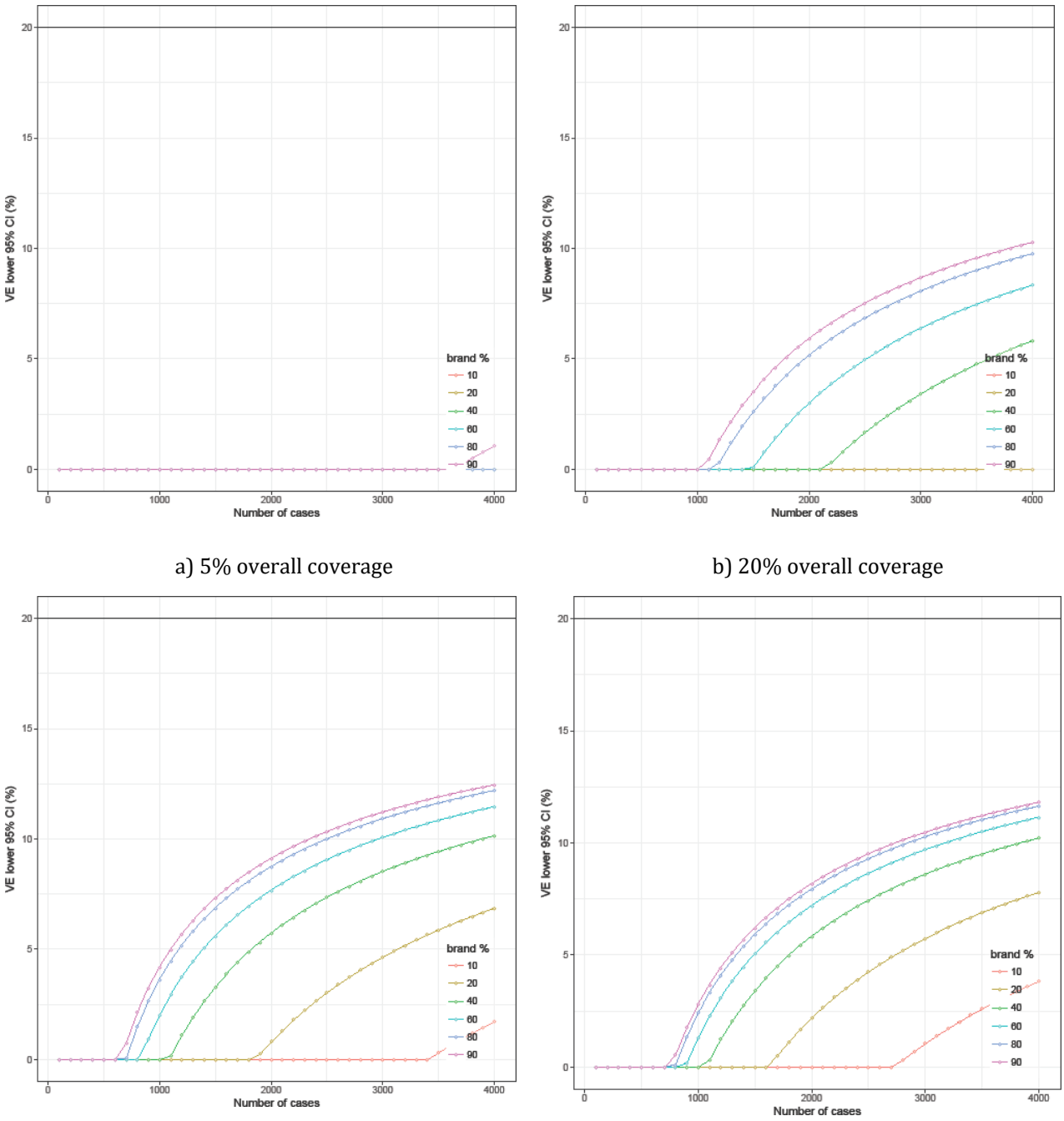


b) 50% VE



c) 70% VE

Figure 5. Precision of overall VE for a case-control study (1:4 case-control ratio) assuming overall vaccination coverage of 5%, 20%, 50% and 70%, and anticipated true VE of 20%, 50% and 70% (indicated with the black horizontal line), by number of cases.



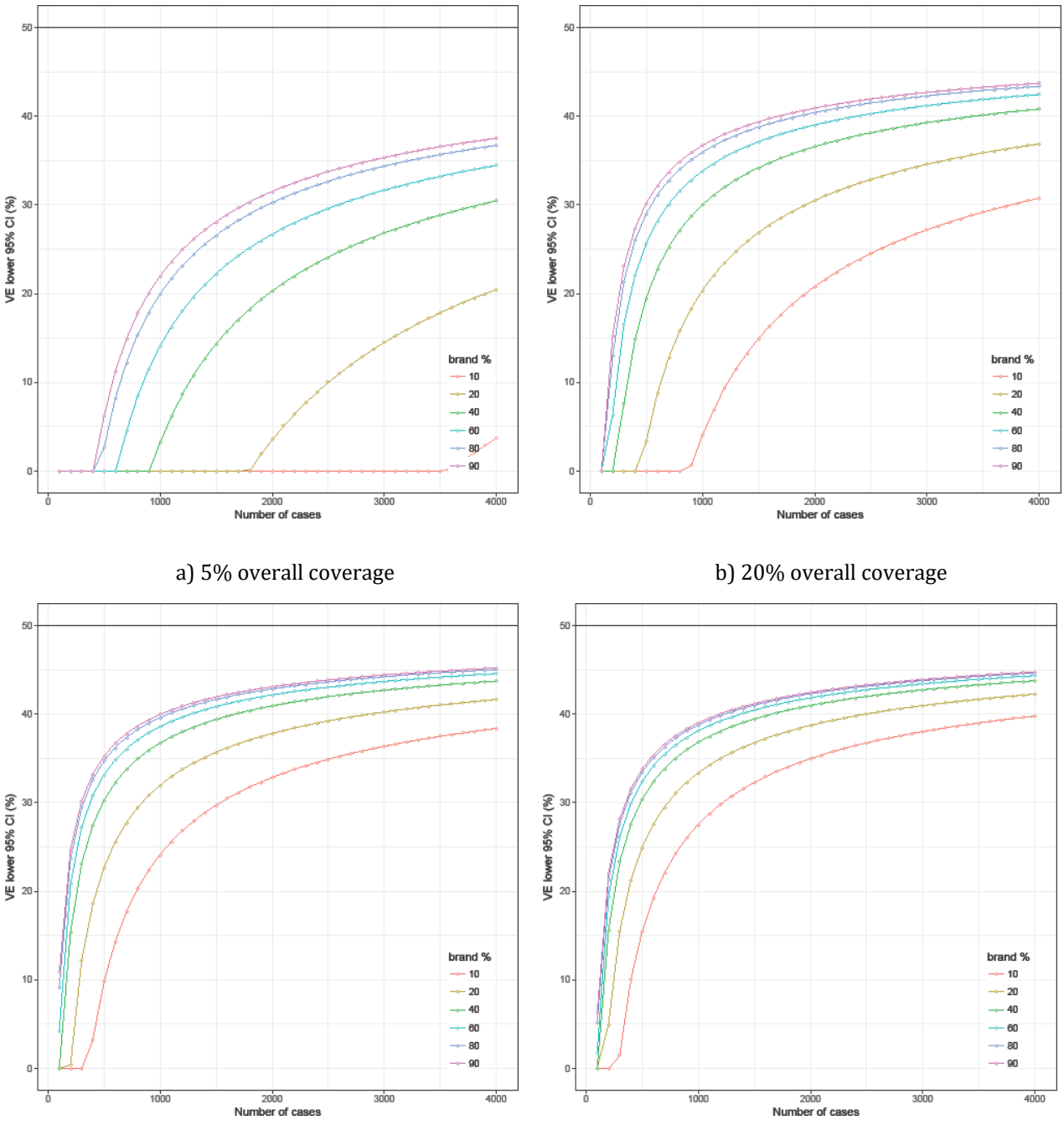
a) 5% overall coverage

b) 20% overall coverage

c) 50% overall coverage

d) 70% overall coverage

Figure 6. Precision of brand-specific VE for a case-control study (1:1 cases to controls ratio) assuming an anticipated true VE of 20% (indicated with the black horizontal line), overall vaccination coverage of 5%, 20%, 50% and 70% with the brand of interest covering 10%, 20%, 40%, 60%, 80% and 90% of the overall coverage, by number of cases.



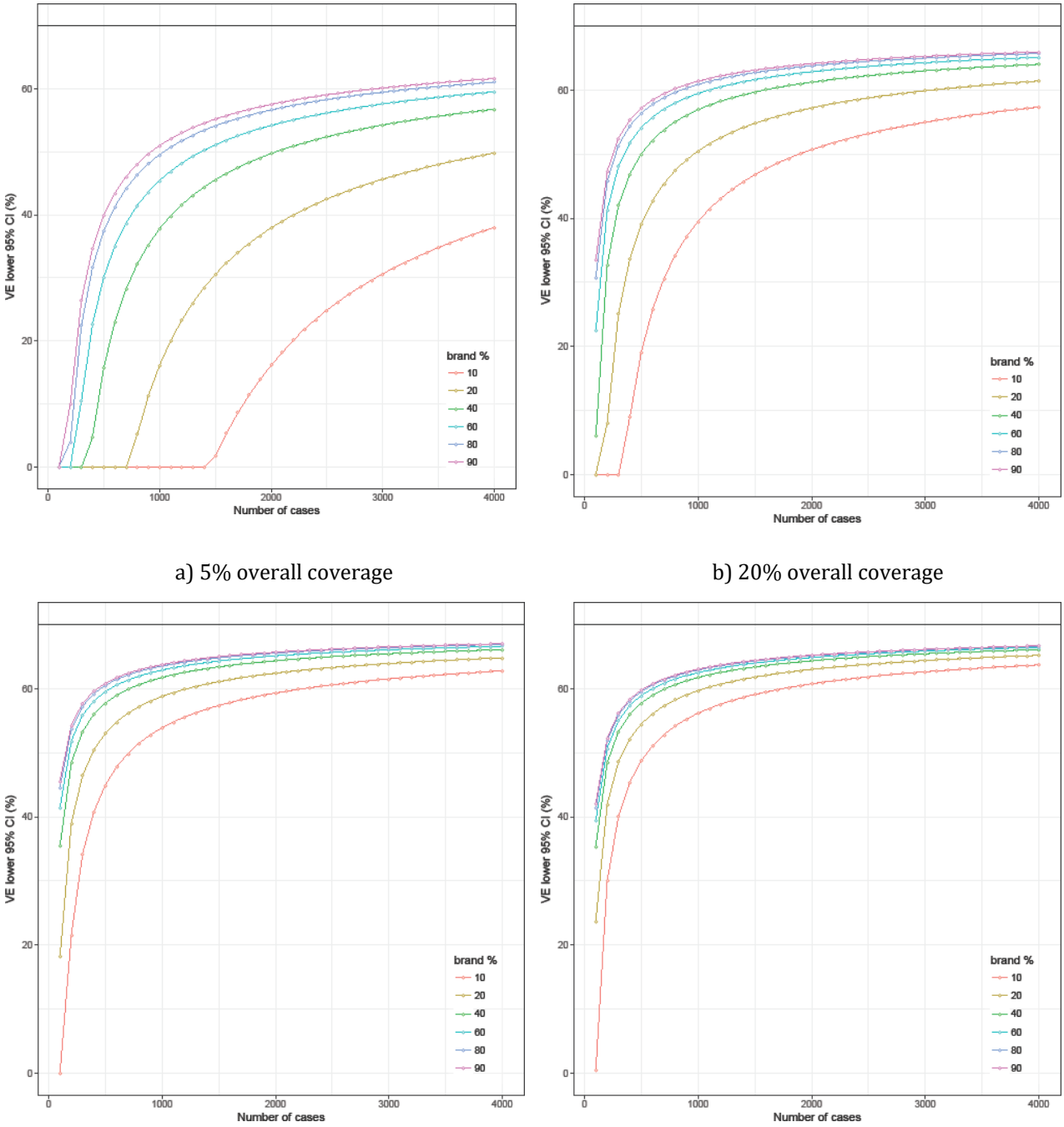
a) 5% overall coverage

b) 20% overall coverage

c) 50% overall coverage

d) 70% overall coverage

Figure 7. Precision of brand-specific VE for a case-control study (1:1 cases to controls ratio) assuming an anticipated true VE of 50% (indicated with the black horizontal line), overall vaccination coverage of 5%, 20%, 50% and 70% with the brand of interest covering 10%, 20%, 40%, 60%, 80% and 90% of the overall coverage, by number of cases.



a) 5% overall coverage

b) 20% overall coverage

c) 50% overall coverage

d) 70% overall coverage

Figure 8. Precision of brand-specific VE for a case-control study (1:1 cases to controls ratio) assuming an anticipated true VE of 70% (indicated with the black horizontal line), overall vaccination coverage of 5%, 20%, 50% and 70% with the brand of interest covering 10%, 20%, 40%, 60%, 80% and 90% of the overall coverage, by number of cases.

Concluding remarks and recommendations

We make the following observations and recommendations based on our sample size calculations for pooled case-control studies;

- We recommend case-control studies based on 500 cases or more. A case-control study with 500 cases and a 1:1 'case to control' ratio will result in 95% CIs of the overall VE with a lower limit of>30% given a true VE of 50% and an influenza attack rate of 5%, for coverages of > 20%.
- Case to control ratios of 1:2 or 1:4 yield slightly more accurate estimates compared to a 1:1 case to control ratio.
- A case-control study with 500 cases and a 1:1 'case to control' ratio will result in 95% CIs of the brand-specific VE with a lower limit of >25% given a true VE of 50% for brands covering 40% to 90% of the influenza vaccines when the overall vaccination coverage is 50% or more.
- A case-control study with 500 cases and a 1:1 'case to control' ratio will result in a minimal detectable overall VE of 30-40% for coverages >20%.
- A case-control study based on 500 cases and a 1:1 'case to control' ratio will result in minimal detectable brand-specific VE of 30-40%, for brands covering 40% to 90% of the influenza vaccines when the overall vaccination coverage is 50% or more.
- A case-control studies based on 1500 to 2000 cases and a 1:1 case to control ratio will result in a minimal detectable VE of 18-20% for an overall vaccination coverage of 20% or more. Improvements in accuracy both in terms of minimal detectable VE and precision will be minimal when increasing sample sizes further
- In case the VE is expected to be low (<20%), higher sample sizes are required to obtain VE estimates with acceptable precision.
- In case interest is in VE within subgroups, the sample size calculations should be done with respect to the subgroup-specific sample size.
- IMPORTANT: These are recommendations to support the design of case-control studies on (brand-specific) VE. Obtaining a minimum sample size is not a requirement for single site participation.

References

1. Woodward M. Epidemiology: study design and data analysis, 3d edition. London: Chapman & Hall, CRC Press; 2013.

Annex 3: Generic Statistical Analysis Plan for pooled analysis

DRIVE D4.4 Generic Statistical Analysis Plan: combining information on Influenza Vaccine Effectiveness across study sites

777363 - DRIVE

Development of robust and innovative vaccine effectiveness

WP4 – Framework for analysis and study reports

Lead contributor	Kaatje Bollaerts (P95)
	kaatje.bollaerts@p-95.com
Other contributors	Jos Nauta (Abbott), Margarita Riera (P95)

Due date	31 March 2018
Delivery date	19 April 2018
Deliverable type	R
Dissemination level	PU

Document History

Version	Date	Description
	07 11 2017	Outline - first draft
V1.0	29 01 2018	First draft
	22 02 2018	Comments Marga Riera and Jos Nauta
V1.1	26 02 2018	Second draft
	09 03 2018	Comments WP4 team members
V1.2	12 03 2018	First version, to share with SC
	26.03.2018	Comments SC: Seqirus, SP, GSK, IRD, THL
V1.3	29.03.2018	Final version
	18.04.2018	Addition of paragraph about sample size considerations and requirements following discussion with SC

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AD-MA	Aggregated data meta-analysis
CI	Confidence interval
DRIVE	Development of Robust and Innovative Vaccine Effectiveness
IPD-MA	individual participant data meta-analysis
IVE	Influenza vaccine effectiveness
OR	Odds ratio
RR	Relative risk
SAP	Statistical Analysis Plan
RRR	Relative Risk Ratio

BACKGROUND

The DRIVE consortium aims to enable the collaboration of different public and private stakeholders to perform annual brand-specific influenza vaccine effectiveness (IVE) studies for various influenza vaccines on the European market. To this end, IVE studies will be conducted at various study sites across Europe. In a second step, the site-specific data will be combined to obtain overall estimates at the European level. The purpose of this document is to provide guidance for writing the Statistical Analysis Plan (SAP) of combining and presenting information on IVE from different study sites. This document will be updated following the learnings from the pilot year 2017-2018.

There are two statistical approaches for pooling data: a one-stage or a two-stage pooling approach (1). The two-stage approach refers to the classical meta-analytical approach, also called aggregated data meta-analysis (AD-MA). In this approach, the patient-level or minimally aggregated data from each study are analysed separately in order to obtain the effect estimates of interest (here vaccine effectiveness estimates) and the corresponding confidence intervals (CIs). Then, in the second step, the effect estimates are combined by an appropriate meta-analysis model to obtain the meta-analytical (weighted averaged) estimate. The one-stage pooling approach analyses all the combined patient-level or minimally aggregated data from the different data sources in a single step. This approach is also called the individual participant data meta-analysis (IPD-MA).

We opt to pool data using the AD-MA approach, given the statistical equivalence of AD-MA and IPD-MA, given that many of the mentioned advantages of IPD-MA (i.e. transforming data to common sources or measures and standardizing analysis) can also be achieved through harmonization/standardization of the individual site-specific studies and given the additional complexity of performing IPD-MA when data are collected using different study designs (1). Within AD-MA, we prefer the use of random effects meta-analysis model, which assumes that the observed effect estimates can vary across study sites because of differences in the treatment effect in each study site (e.g. due to differences in population, in health care utilization, in circulating influenza strains) as well as sampling variability.

This document builds further upon or relates to the DRIVE generic study protocols for the analyses and presentation of data collected at a single study site, the DRIVE data management plan and the DRIVE report template (see Reference documents).

REFERENCE DOCUMENTS

[Here: refer to the generic study protocols for the analyses of the study site-specific data, the data management plan and the report template]

AGGREGATED DATA META-ANALYSIS

Objective(s)

To estimate seasonal IVE (%) through pooling site-specific estimates obtained as described in the site-specific protocols.

[Describe the primary and secondary objectives as per study protocol mentioned in Section 2 and for which pooling will be performed]

Effect measures

The effect measures for pooling are the study site-specific IVE estimates and their 95% confidence intervals (CIs).

Sample size considerations

[Sample size considerations for the primary objective(s) should be discussed in this section including the assumptions made for vaccination coverage, vaccine effectiveness and influenza attack rate. This section will be updated pending consultation with the DRIVE Ethics Advisory Board and EMA on the need to establish minimum sample size and/or minimum precision for the primary objective(s)].

Strategy for data synthesis

Inclusion criteria

We will pool seasonal IVE estimates from the individual study sites in line with the objectives as per study protocol (Section 3.1). Estimates that are not obtained following the study protocols will not be retained for the primary meta-analysis, but might be considered for inclusion as part of a sensitivity analysis (Section 3.4.6). Whenever there are two or more site-specific estimates retained, a meta-analysis will be performed.

Further pooling (e.g. incorporating IVE estimates which were not minimally adjusted for confounding

as per study protocol) might be considered upon lack of heterogeneity (see Sections 3.4.4 and 3.4.5).

Meta-analysis

For every objective listed in Section 3.1, a meta-analysis will be performed. First, the study site-specific IVE estimates will be back-transformed to the original relative risk (RR) estimates (in case of cohort studies) and odds ratio (OR) estimates (in case of case-control studies), which will be subsequently log-transformed, or

$$\log RR \text{ or } \log RR = \log(1-VE)$$

Then, standard inverse variance weighted random-effects meta-analysis of the log-transformed RR and OR estimates will be used to obtain the pooled estimate (2). The pooled estimate (and 95% CI) will then be back-transformed to obtain the pooled IVE estimate (and 95% CI), expressed in %.

Outlier and influence analysis

For every meta-analysis performed, the potential impact of outliers and influential estimates on the pooled estimate will be evaluated. Studentized deleted residuals *r* will be used to identify outliers in the meta-analysis. Site-specific IVE estimates will be considered outlying from meta-analysis when $|r| > 2.5$, where $|r|$ indicates the absolute value of the residual (3).

The standardized DFBETAs statistic will be used to identify influential estimates, examining the change in the averaged IVE from the random-effects model when excluding one site-specific estimate in turn. Site-specific estimates will be considered influential from meta-analysis when $|\text{DFBETAs}| > 2/\sqrt{n}$, where where |DFBETAs| indicates the absolute value of the DFBETAs statistics and *n* is the number of effect estimates (3).

Site-specific estimates that are outlying and influential, will be excluded from meta-analysis and the reason for being outlying will be investigated and documented.

Quantifying between-study heterogeneity

An indication for the heterogeneity among estimates from different study sites will be obtained by calculating I^2 according to Higgins et al (4). The I^2 statistic is to be interpreted as the proportion of total variation in the estimates of treatment effect that is due to heterogeneity between studies. Low, moderate and high levels of heterogeneity correspond to I^2 values of 25%, 50% and 75%

respectively. In case I^2 is high, it is worthwhile to explore sources of heterogeneity (Section 3.4.5).

Exploring sources of heterogeneity

In case of at least 5 site-specific IVE estimates, stratified analyses and meta-regression might be used to explore whether the magnitude of the IVE estimates are associated with design or other characteristics of the study site-specific estimates of interest (e.g. study design, adjustments for certain covariates). In stratified analyses, the meta-analysis (as in Section 3.3.3) will be repeated for each stratum of characteristics separately. In meta-regression, the meta-analysis (as in Section 3.3.3) will be extended with the site-specific study characteristics as predictor variables and relative risk ratios (RRRs) will be obtained (5). For example, assume the characteristic of interest is study design (cohort vs case-control studies). Then, the RRRs is to be interpreted as the ratio of the pooled IVE estimate of the case-control studies to the pooled IVE estimate of the cohort studies.

The permutation test as proposed by Higgins et al (6) will be used to assess the significance of a study characteristic while controlling the risk of false-positive results. If the study characteristic is not statistically significant in the meta-regression model, the study characteristic is unlikely a source of heterogeneity, and pooling across that study characteristic might be considered.

Sensitivity analysis

Sensitivity analysis in line with the study protocol will be performed.

Additional sensitivity analyses will be performed by including site-specific estimates that were excluded from the main meta-analysis models because 1) they were not obtained following the study-protocol (Section 3.3.1) or 2) they were identified as outlying and influential (Section 3.3.3).

Presentation of results

The site-specific IVE estimates (and 95% CIs) will be presented using a forest plot complemented with the pooled IVE estimate (and 95% CIs) as outlined in the report template. Estimates that were excluded from meta-analysis will included in the forest plot, but these estimates will be tagged as excluded. An example of a forest plot with pooled estimates by setting is given in Figure 1. This plot is generated using artificial data based on cohort designs.

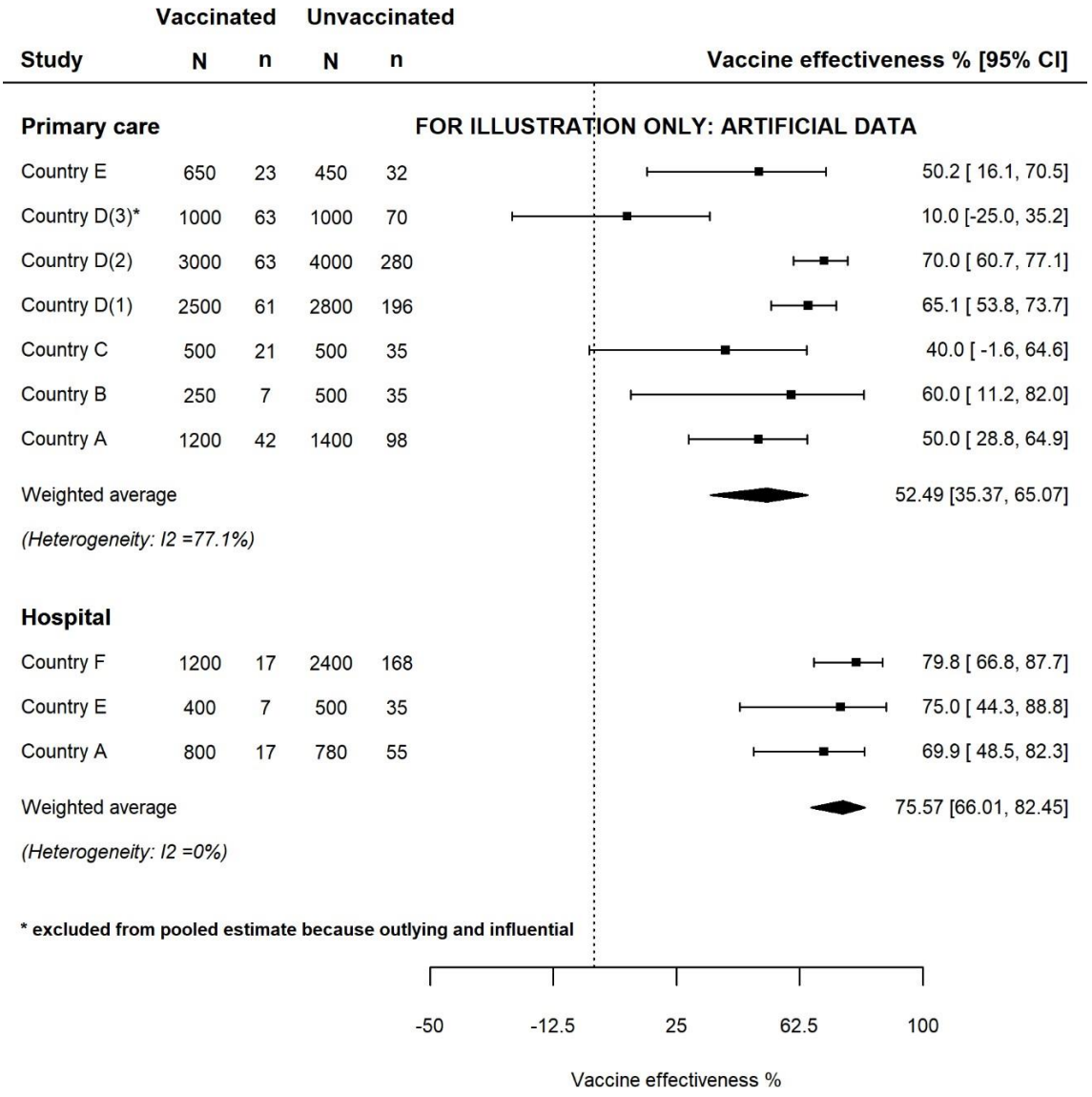


Figure 1: Forest plot and meta-analyses of influenza vaccine effectiveness, by health care setting. This plot is generated using artificial data based on cohort designs

REFERENCES

1. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. Stat Med 2017;36(5):855-875.
2. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-88.
3. Viechtbauer W, Cheung MW. Outlier and influence diagnostics for meta-analysis. Res Synth Methods 2010;1(2):112-25.
4. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21(11):1539-58.
5. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 2010;1(2):97-111.
6. Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. Stat Med 2004;23(11):1663-82.