D7.1 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	Caterina Rizzo (3 – Istituto Superiore di Sanità)
	caterina.rizzo@iss.it
Other contributors	Valeria Alfonsi (3 – Istituto Superiore di Sanità)
	Kaatje Bollaerts, Margarita Riera, Anke Stuurman (4 – P95)
	Topi Turunen (1 – FISABIO)

Due date	30 Sept 2018
Delivery date	31 Oct 2018
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.

Description of Work	Version	Date
	V1.1	05 Dec 2018

Document History

Version	Date	Description
V0.1	09/01/2018	First draft
V0.2	22/01/2018	Comments
V0.3	13/03/2018	Comments, 2nd round
V0.9	27/04/2018	Semifinal version
V1.0	04/06/2018	Version approved by Independent Scientific Committee
V1.1	05/12/2018	Updated codebook / minimum dataset requirement

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
MAH	Marketing Authorization Holder
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 IMI project 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.

The main objective of the 2017/18 pilot season is to test the different operational aspects of the DRIVE project, including governance, data collection, statistical analyses and dissemination of study results. Consequently, the number of study sites for this season is limited with narrow possibility to study the full range of vaccine brands used across Europe.

This generic protocol is intended to be adapted to the local procedures at each individual study site from season 2018/19 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.

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 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 data, that will provide product 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). While each of the study sites can be analysed 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. The details of each site-specific study will be provided in the study annexes (e.g. ethical committee clearance, study form used, data collection strategy, etc.).

Objectives

Primary objective

To measure seasonal IVE against medically attended (primary care/hospital) laboratory-confirmed influenza, by vaccine brand, then by vaccine type (e.g. by antigen preparation strategy, number of virus strains, adjuvant; see section Exposure), then by overall influenza vaccination.

Secondary objectives

To estimate IVE (brand-specific, type specific and total, if possible) against laboratory-confirmed influenza by:

- age group (6 months-14 years; 15-64 years; 65+ years, the age groups will be further defined when harmonising the study protocols between study sites according to availability of data)
- influenza virus type (A, B) and/or subtype (A/H1N1, A/H3N2) and lineage (B/Victoria, B/Yamagata)
    - risk groups / target groups for vaccination,
    - pregnant women
    - healthcare workers
    - any chronic condition (see annex 1)
    - specific chronic conditions (see annex 1)
- time since vaccination
- time window in the epidemic season (early, middle, end) (see study period section)
- previous influenza vaccinations (at least one previous season, preferably more)

To estimate IVE across:

- several seasons

The details of the analyses will be prescribed in the generic and study site specific SAP, updated annually according to the characteristics and structure of data available at the participating study sites.

Methods

Study design

➢ A multicentre study using data from several study sites

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

Study setting

The studies may take place in a primary care or a hospital setting. The study setting is defined by each study site depending on the available data.

- ➢ *Each study site to specify if the study is nested into the influenza surveillance scheme (the ILI sentinel 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*

Study period

The seasonal assessment will start when the influenza virus circulation begins (first virus detected at the national/study site level) in the country/region and will finish at the end of the influenza season (no cases detected during 2 consecutive weeks or equivalent).

- ➢ *Each study site to specify the assessment period: the definition of the beginning, peak and end of the influenza period at the study site according to the information provided by the local influenza surveillance system (including information on the type of virus circulating and virulence of the virus)*

For the joint analysis, a harmonised minimum period will be defined (e.g. from week 40 till week 20), but if needed, it will be extended to fully cover the vaccination campaign and the epidemic in each study site. Definition of shorter time periods (e.g. early, middle and end season) will be developed to take into account differences in influenza activity over time and probable development of immunity in unvaccinated population through encounters with the circulating viruses

For addressing the secondary objective of estimating influenza vaccine effectiveness over several seasons, multiple study periods will be combined.

Study population

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

- ➢ *Each study site to specify the study population and the case finding procedure, please see the* Case finding *section.*

Outcomes

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

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

- laboratory-confirmed influenza by clade (where possible).

Case definition

Influenza-like illness (ILI)

A case of influenza like illness (ILI) will be defined by the ECDC case definition 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 (SARI)

A case of severe acute respiratory infection (SARI) will be defined by the SARI - IMOVE+ 2017/2018 case definition as a hospitalised person with

- at least one systemic symptom or sign (fever or feverishness, malaise, headache or myalgia), or deterioration of general condition (asthenia or loss of weight or anorexia or confusion or dizziness)

AND

- at least one respiratory symptom or sign (cough, sore throat or 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.

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 an ILI patient in the study population, meeting the ILI - EU case definition for clinical criteria, with a respiratory sample negative for influenza.

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 (see laboratory testing section).

- 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.

Case finding

ILI and SARI patient identification

Patients will be identified among people who present at a healthcare provider (GPs or Hospitals) with influenza-like illness (ILI) or severe respiratory acute infection (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*

Inclusion criteria

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

Exclusion criteria

The ILI patient will not be enrolled in the study if she or he:

- Is less than 6 months of age at the time of recruitment
- has a contraindication for influenza vaccine
- 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)
- is institutionalised at the time of symptoms onset (lives in a residence for people who require continual nursing care and have difficulty with the required activities of daily living)
- had a respiratory specimen taken ≥ 8 days after ILI onset
- tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalisation

The SARI patient will not be enrolled in the study if she or he:

- Is less than 6 months of age at the time of recruitment
- has a contraindication for influenza vaccine
- was previously hospitalised < 48 hours prior to ILI onset
- had his/her ILI onset ≥ 48 hours after admission at the hospital
- 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)
- is institutionalised at the time of symptoms onset (lives in a residence for people who require continual nursing care and have difficulty with the required activities of daily living)

- had a respiratory specimen taken ≥ 8 days after ILI onset
- tested positive for any influenza virus in the current season before the onset of symptoms leading to the current hospitalisation

Note: a patient can be selected several times as long as he/she does not have a previous laboratory confirmed influenza for the current season

Exposure (vaccination)

Exposure of interest

The exposure of interest is vaccination with any influenza vaccine (seasonal or pandemic) 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.

The vaccine type specific VE may be assessed e.g.

- by strategy used for influenza antigen preparation (live attenuated, inactivated, subunit, split virion),
- by number of vaccine virus strains contained in the different vaccines available (trivalent, tetravalent)
- by adjuvant (adjuvanted, non-adjuvanted)
- by vaccine dose (one dose, two doses; 0,25 ml, 0,5 ml)
- by manufacturing process (egg-based, cell-based)

The vaccine types selected to primary and potential sensitivity analyses will be specified in the generic study site level SAP, which will be updated annually, if needed.

Vaccination status ascertainment

The sources of information for the vaccination status may include:

- vaccination registry
- consultation of the patient's vaccination card
- interview with the patient's GP
- interview with the patient's pharmacist
- data from the patient's insurance company showing evidence of pharmacy delivery or reimbursement of influenza vaccine during the current influenza season

An interview of the patient and/or their relatives alone is not a preferred method of vaccine status ascertainment but may be performed. When vaccination status is positive according to any of the above sources but not recalled by patient, they will be coded as vaccinated. When positive vaccination status is indicated only by recall or is otherwise ambiguous, the vaccination status will be coded as "potentially vaccinated".

➢ *Each study site to describe the precise way of vaccination status 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 previous 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 (see section *Vaccination status ascertainment*)
- **partially vaccinated** 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
- **potentially vaccinated** if the positive vaccination status is based on recall alone and cannot be confirmed by registers, or is otherwise ambiguous.

A child aged < 9 years who has not been fully vaccinated (see above) during the previous 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 (see section *Vaccination status ascertainment*)
- **partially vaccinated**
    - during the first 14 days after the second record of injectable vaccination or the first record of LAIV vaccination during the current season
    - after the first record of injectable vaccination until >14 days have elapsed since the second record of vaccination during the current season
- **unvaccinated** until the first vaccination record during the season
- **potentially vaccinated** if the positive vaccination status is based on recall alone and cannot be confirmed by registers, or is otherwise ambiguous.

The *partially* and *potentially* vaccinated groups will be excluded from primary analysis; their significance will be assessed in sensitivity analyses.

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).

The minimum set for a pooled analysis is marked with an asterisk(*). If available and relevant, the other determinants may be used in individual study site analyses, and if possible, they will may be harmonised between the study sites for pooled analysis, by developing guidelines for harmonization according to availability of data and included in the generic study level SAP.

- Age*
- Sex*
- Number of healthcare 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*

- Any chronic underlying conditions or if possible to define (like chronic pulmonary disease, cardiovascular disease, metabolic disorders, renal disease, treatment-induced immunosuppression and disease-induced immunosuppression, medically attended obesity .*
- Influenza vaccination in previous influenza seasons (at least one)*
- Contraindication to influenza vaccination
- Pregnancy
- Use of influenza antivirals
- Use of statins
- Pneumococcal vaccination
- Socio-economic status or applicable proxy
- Smoking behaviour or parental smoking behaviour (for subjects ≤18 years)
- (For children) Perinatal and congenital risk factors (e.g. birth weight and/or maturity at birth, perinatal factors, inborn errors of metabolism, relevant malformations and congenital syndromes)
- (For children) Number of siblings
- (For children) Adherence to the local childhood vaccination programme

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

The list will be updated based on results of DRIVE D2.2: Systematic review of the sources of confounding, bias and strategies to manage their impact in influenza vaccine effectiveness studies, due June 2018.

Sources of information

Data will be collected using a standardised questionnaire/data collection form (see data collection section). The source(s) for the current and previous vaccination status and for collecting general data may include:

- hospital medical records
- consultation of the patient's vaccination card
- data from the patient's insurance company showing evidence of pharmacy delivery or reimbursement of influenza vaccine
- interview with patient or his/her family
- interview with patient's GP (according with rules for Vaccination Status Ascertainment)

- interview with patient's pharmacist
- vaccination register

➢ *Each study site to define the sources of information used for each variable collected*

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*

Laboratory testing

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

Laboratory confirmation should be done through one of the following laboratory tests: reverse transcription-polymerase chain reaction (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).

➢ *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 (see Annex 4 for an example of results presentation)*

➢ *Each study site to describe if the laboratory participates in QA/QC (Quality Assurance/Quality Control) schemes*

Sample size considerations

This section gives sample size considerations and formulates recommendations. These recommendations are meant to support the design of the case-control studies on IVE. Obtaining a minimum sample size is not a requirement for study participation. Details on the sample size calculations based on the minimal detectable VE as well as precision are given in Annex 3.

DRIVE recommends case-control studies based on 500 cases or more. However, studies with smaller sample sizes might still contribute to the power of the pooled analyses, provided that the study site is able to optimally harmonise its protocol with the other study sites to minimize the

between-study heterogeneity. In case VE estimates with unacceptable large CIs are obtained, it might be considered to not report these estimates.

Figure 1 presents the precision of the overall VE for number of cases varying from 200 to 4000 subjects, when assuming a true VE of 50%, a 'cases to controls' ratio of 1:1 and a total vaccination coverage of 5%, 20%, 50% and 70%. The number of cases per control is likely to vary (by definition of test-negative design). The calculations are based on an anticipated true VE of 50% as this is a conservative choice, requiring larger sample sizes compared to assuming lower/higher VE values. A case-control study based on 500 cases and a 1:1 'cases to controls' ratio will result in 95% CIs of the overall VE with a lower limit larger than 30% given a true VE of 50%, for coverages of >20%.

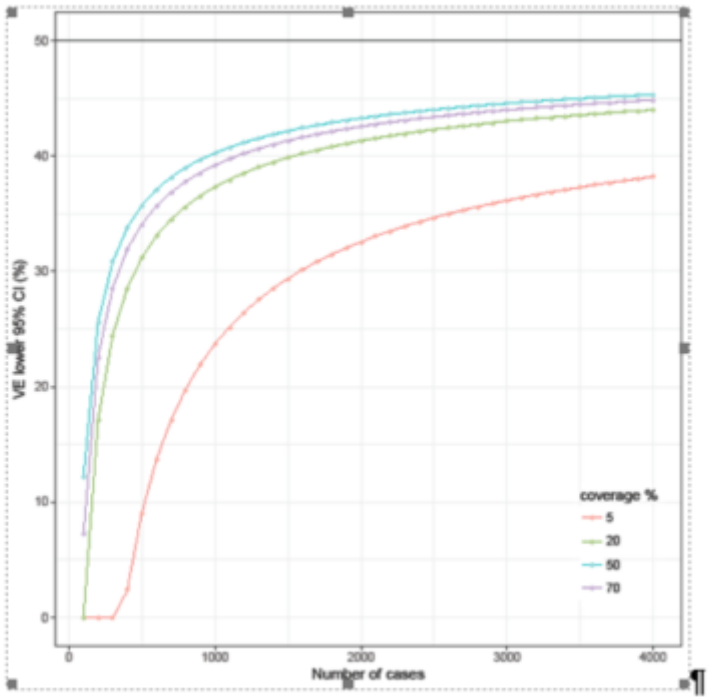


Figure 1. Precision of the overall VE expressed as the lower limit of the 95% CI, assuming a true VE of 50%, a 'cases to controls' ratio of 1:1 and a total vaccination coverage of 5%, 20%, 50% and 70%.

Figure 2 presents the precision of the brand-specific VE for number of cases varying from 200 to 4000 for the same parameter settings as above and additionally assuming the brand of interest accounts for 10%, 20%, 40%, 60%, 80% and 90% of the total vaccination coverage. A case-control study based on 500 cases and a 1:1 'cases to controls' ratio will result in 95% CIs of the brand-specific VE with a lower limit larger than 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.

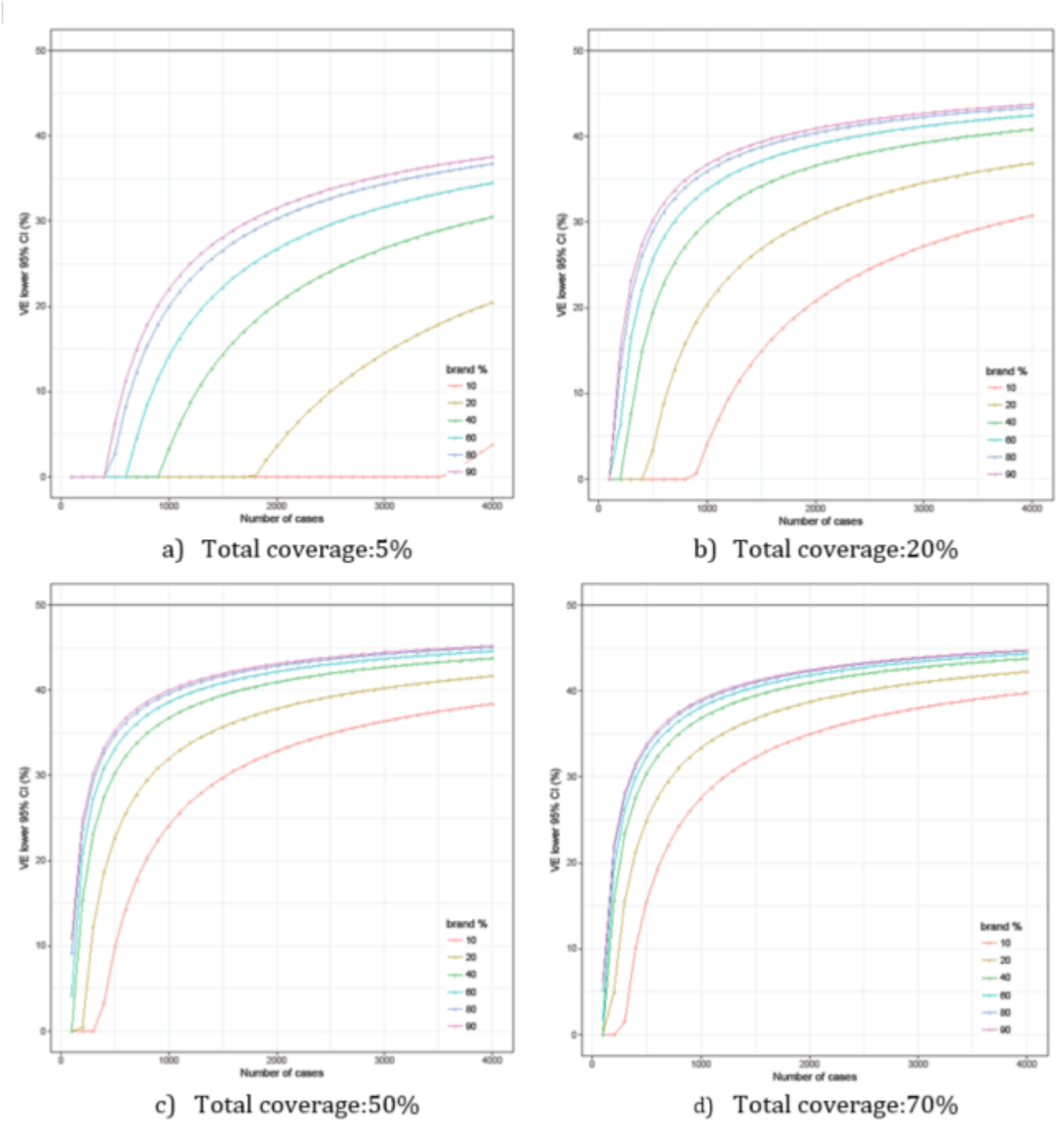


Figure 2. Precision of the overall VE expressed as the lower limit of the 95% CI, assuming a true VE of 50%, a 'cases to controls' ratio of 1:1, a total vaccination coverage of 5%, 20%, 50% and 70% and that the brand of interest accounts for 10%, 20%, 40%, 60%, 80% and 90% of the total vaccination coverage.

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 (task 4.2.1) and set up the necessary infrastructure for data collection and analysis of the pooled data (task 4.2.2). 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). Ideally, these checks will be included in the electronic questionnaire in order to avoid inconsistencies in the data entry. These values will be checked against the questionnaires or queried with the hospitals. 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.

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 compliance 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. Beside the collection of dates of onset/admission/respiratory specimen collection, health-seeking behaviour and case-management strategy should be described for each study and it should be noted how it may affect the VE estimates.

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 amount of variables to adjust for, and the heterogeneity/homogeneity between the study sites will be optimised according to availability of data. For one-stage and/or 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 described and tabulated for cases and controls separately and for vaccinated and unvaccinated subjects within each group (by brand, type and overall). The baseline characteristics of the cases and controls will be compared using the Fisher's exact test (in case of nominal variables for the baseline characteristics), Mann-Whitney test (in case of ordinal or non-normal continuous variables) or Student's t-test (in case of normal variables).

Measure of effect

The crude (or unadjusted) brand-specific 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 brand-specific IVE estimates will be obtained from multivariable logistic regression models, regressing the health outcomes of interest on exposure status, age, sex and the confounders of interest. In case of effect modifiers, an interaction term between exposure and the effect modifier will be included in the regression model or stratified regression analyses will be performed.

Missing data

Subjects with missing data in the exposure (e.g. missing date at vaccination) or health outcome variables (e.g. missing data at symptom onset) will be excluded.

For each covariate, the amount and possible reason for missing data will be described. For covariates for which the amount of missing data at the study site level is not substantial (<15%), we will introduce an additional missingness category.

For covariates for which the amount of missing data is substantial (>=15%), multiple imputation methods will be applied assuming that the missingness does not depend on unobserved variables. A sensitivity analysis will be carried out comparing the IVE estimates based on the multiple imputation approach with the IVE estimates based on a complete case analysis (e.g. omitting records with missing covariate information from the analysis).

Addressing confounding & bias

Observational influenza vaccine effectiveness studies are prone to several sources of confounding and other types of bias. Please also refer to section *Potential confounders and effect modifiers* and

DRIVE D2.2: Systematic review of the sources of confounding, bias and strategies to manage their impact in influenza vaccine effectiveness studies.

- *Negative confounding* refers to biases that reflect the fact that high risk groups (people more likely to develop severe complications) will be more likely to be vaccinated and therefore reduce VE. If negative confounding is present, the VE will be underestimated. Adjustment for potential negative confounding factors documented in the study (e.g. presence of chronic diseases) will minimise negative confounding.
- *Positive confounding* refers to biases that reflect a 'healthy vaccine effect'. People with a healthy lifestyle will be more likely to accept vaccination, thus leading to an increase of measured VE. Or, similarly, people being in a state of "extreme frailty" will not be offered vaccination. If positive confounding is present, VE will be overestimated.

Thus, it is important to collect information on both the frailty and the healthcare seeking behaviour adequately and to balance possible differences between the vaccinated and the unvaccinated in the study population.

As the data are collected directly form GPs it is difficult that misclassification (information bias) might occur. In any case, during the DRIVE collaboration process, the potential confounders and biases as well as strategies to reduce their impact in VE studies will be identified by a systematic literature review (DRIVE D2.2), and the protocol may be updated accordingly.

Sensitivity analyses

When appropriate, sensitivity analyses may be conducted to test different outcome definitions, different exposure definitions or exclude a subset of the data (e.g. the different influenza testing methods PCR vs rapid; swab taken >= 4 days after symptom onset; underlying swabbing practice, etc.).

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 also be submitted to WP7 for European pooled and/or meta-analyses. 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 pool estimates.

Both study site- and consortium level reports are to follow the template provided by DRIVE D4.3: Report templates.

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.

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 (responsible 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 and/or meta-analysis across several 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. Guidelines of definitions for the study variables will be included in the generic study site level analysis plan (SAP), for harmonisation of the methods between the study sites (Annex 2).

Potential systematic or major deviations from the SOP and generic study level SAP should be described for further development of the methodology and for interpretation of the results. DRIVE WP 2 and WP 3 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 organised*

Changes to the protocol

After further evaluation of the characteristics of the data available in the study sites, the protocol will be further developed to define the minimum data set to provide crude VE estimates and datasets to provide adjusted VE estimates. The aim of DRIVE is to develop methods and receive sufficient data to reach the highest possible accuracy in controlling for confounding and other bias. However, also less optimal datasets may be valuable in improving the precision of the VE estimates and in analysing the nature and impact of bias in observational study designs.

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. Margaret McCartney: 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).

Annex 1: Minimum dataset requirement

DRIVE – Minimum dataset for pooled data analysis (case-control studies) v4 16 Oct 2018

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 years old. Else should not be provided.)		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
ili	Obligatory (for outpatient)	Influenza like illness	Fulfilling the EU-ILI case definition	Numeric (Categorical)	0=No 1=Yes 9999=No information	1

sari	Obligatory (for inpatient)	Severe acute respiratory infection	Fulfilling the I-MOVE+ SARI case definition			
respinfection	Optional	Does the patient have a suspected respiratory infection?		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
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: 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: 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: 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 2017-2018		dd/mm/yyyy		11.1.2018
seasvaccn1	Optional	Received influenza vaccination in previous season (season n – 1)		Numeric (Categorical)	0=No 1=Yes 9999=No information	
seasvaccn2	Optional	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	Obligatory	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	Obligatory (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	Optional	Has the patient received an antiviral treatment 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	Obligatory	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: Chronic conditions and risk factors

Potential list of chronic conditions and risk factors to be considered:

Chronic Conditions

Presence of any chronic disease

Lung disease

Heart disease

Diabetes

Renal disease

Hematologic disorders and hemoglobinopathies

Neoplasia

Cirrhosis

Diseases leading to a reduction in antibody production

Immunodeficiency

Chronic inflammatory disease and intestinal malabsorption syndrome

Diseases associated with an increased risk of aspiration of respiratory secretions i.e. neuromuscular diseases

Risk Factors

Obesity

The child goes to kindergarten

Close contact of a-risk individual who cannot be vaccinated

Patient belongs to professional category for which vaccine is recommended

Person is currently pregnant or delivered in the previous 6 months

Hypercholesterolemia or hypertension

Smoking habit

Statin use

Requires assistance to walk

Requires assistance to bathe

Requires assistance to eat

Severity (Proxy to evaluate the health condition prior to the enrolment in the study)

Number of hospitalisations previous year for the chronic disease

Number of GP consultations previous year

Annex 3: 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

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

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

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

where

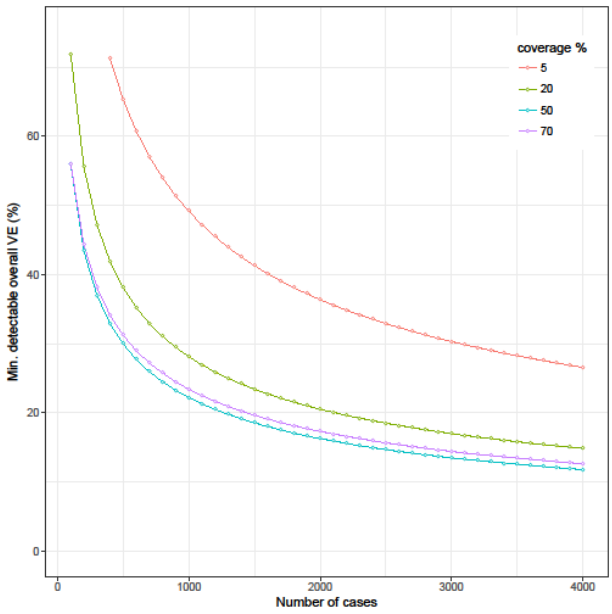
$$a = r\gamma^2 - \frac{Nr\gamma(1-\gamma)}{\left(z_{\frac{\alpha}{2}} + z_\beta\right)^2 (r+1)} \quad ; b = 1 + 2r\gamma,$$

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).

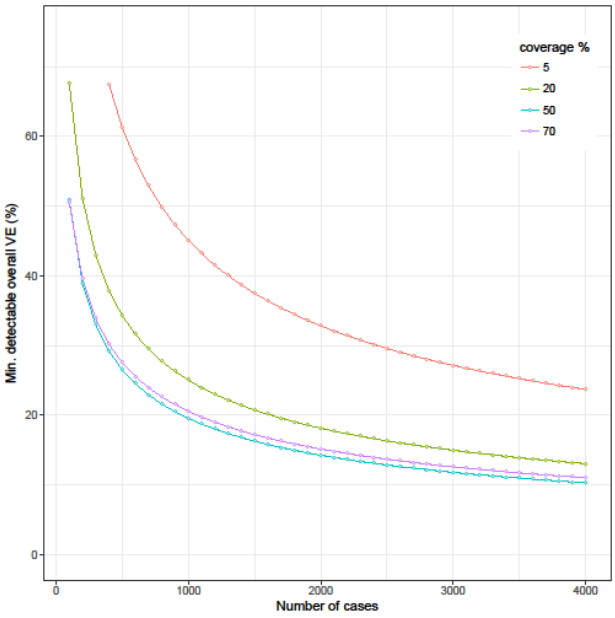
We calculated the minimal detectable overall VE (1) with 80% power (1 – β) and a two-sided 95% confidence coefficient (1 – α/2) for case-control studies using 'cases to controls' ratio of 1:1, 1:2 and 1:4 with the number of cases varying from 100 to 4000, while assuming overall vaccination coverages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

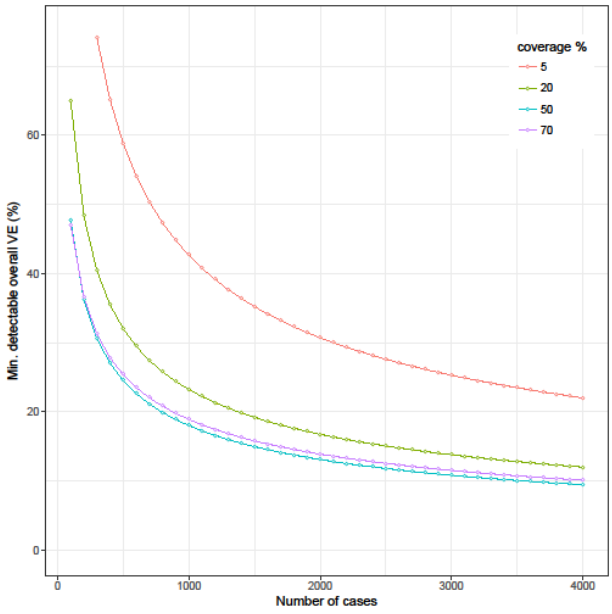
The results for the minimal detectable overall VE for the 1:1, 1:2 and 1:4 'cases to controls' ratios are given in Figure 1. These figures represent the minimal detectable VE by number of cases. The results for the minimal detectable brand-specific VE for the 1:1 'cases to controls' ratio and assuming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



a) 1:1 cases to controls

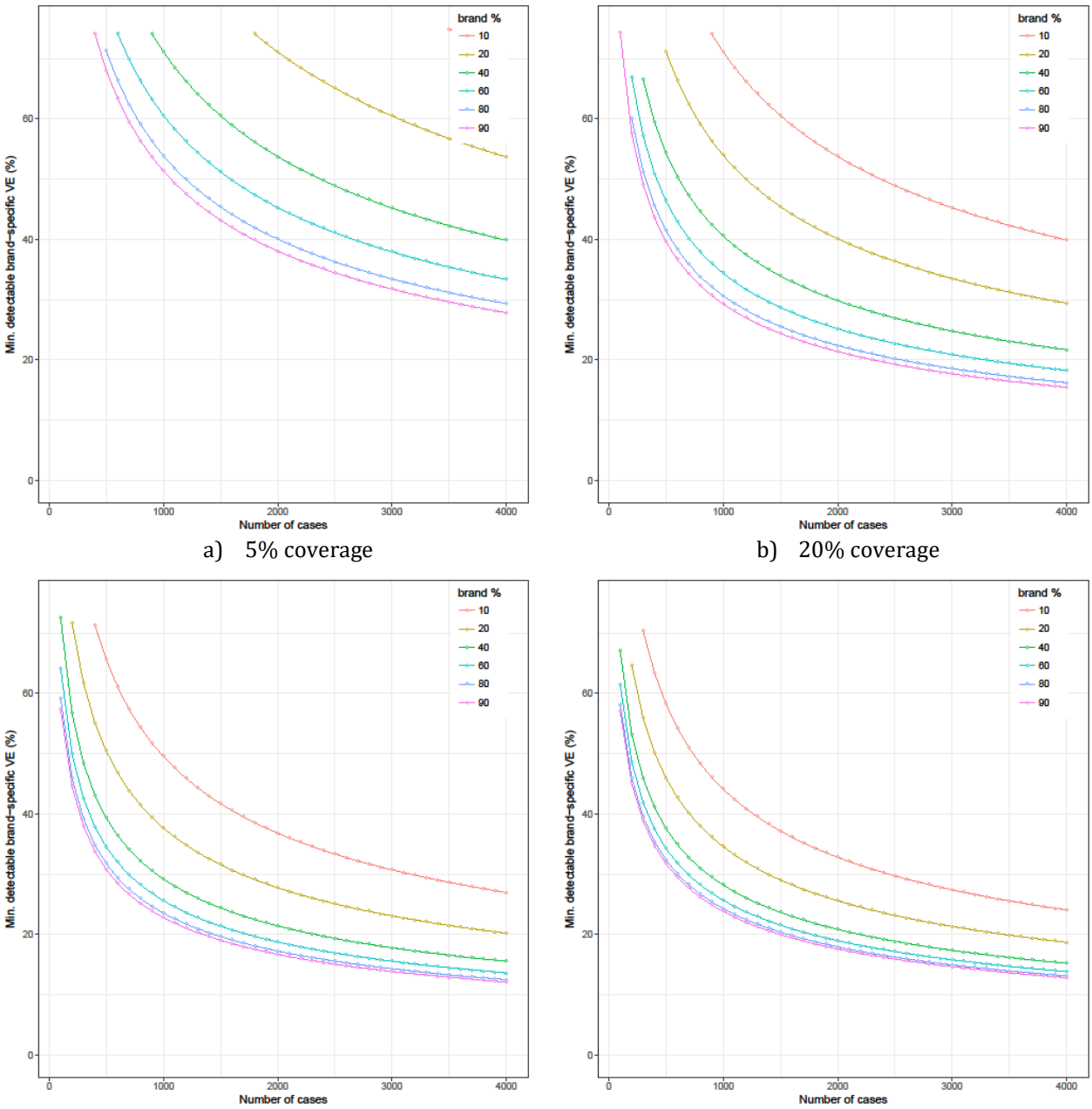


b) 1:2 cases to controls



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_d^+ 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

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

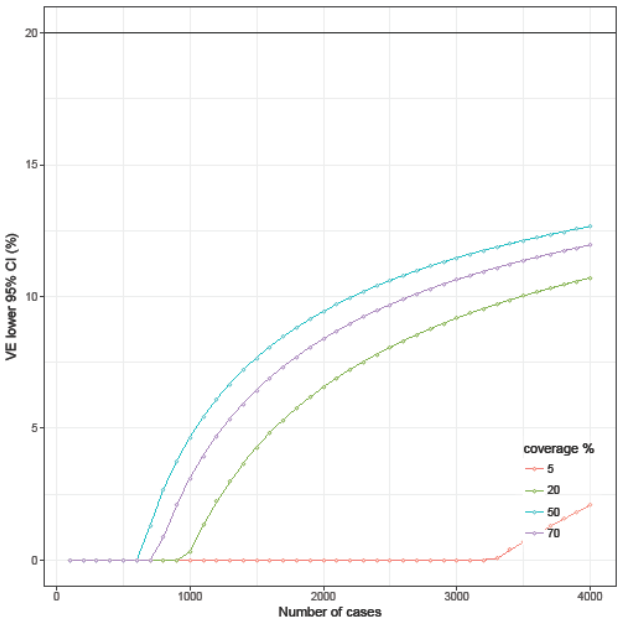
$$x_2 = {N_e^-}^2 + 2N_e^- N_d^+ OR - 2N_e^- N_d^+$$

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

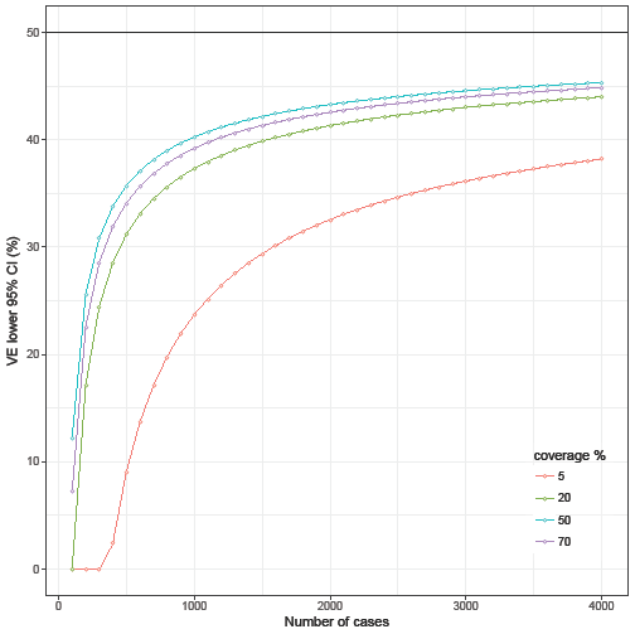
We calculated the precision of the overall VE based on a two-sided 95% CI for case-control studies using 'cases to controls' ratio of 1:1, 1:2 and 1:4 with the total number of cases varying from 100 to 4000, while assuming overall vaccination coverages of 5%, 20%, 50% and 70% and overall VE of 20%, 50%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

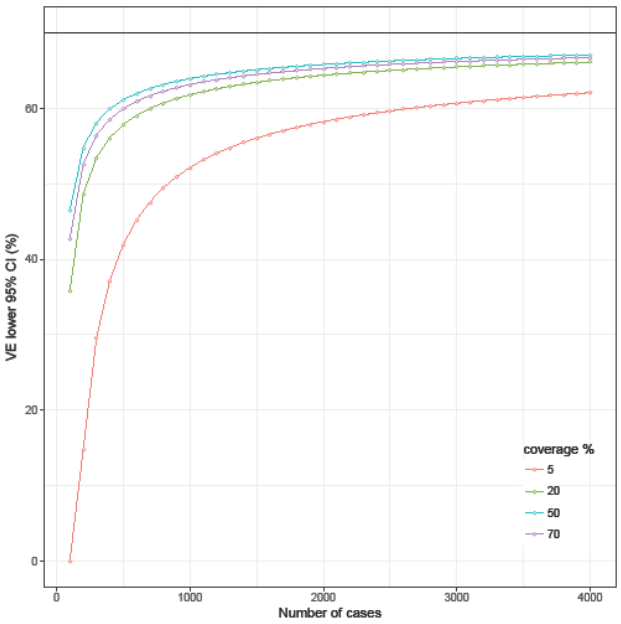
The results for precision of the overall VE using 'cases to controls' ratio of 1:1, 1:2 and 1:4 are given in Figure 3 to 5, respectively. These figures represent precision by number of cases. The results for the precision of brand-specific VE using 'cases to controls' ratio of 1:1 for anticipated true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



a) 20% VE

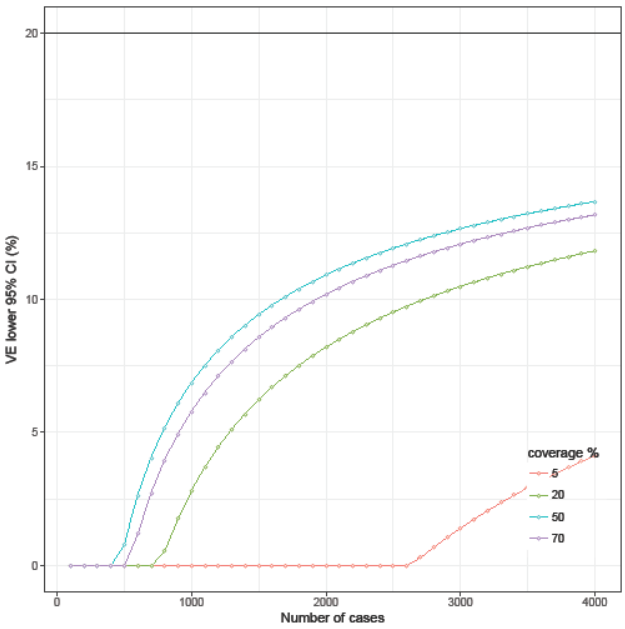


b) 50% VE

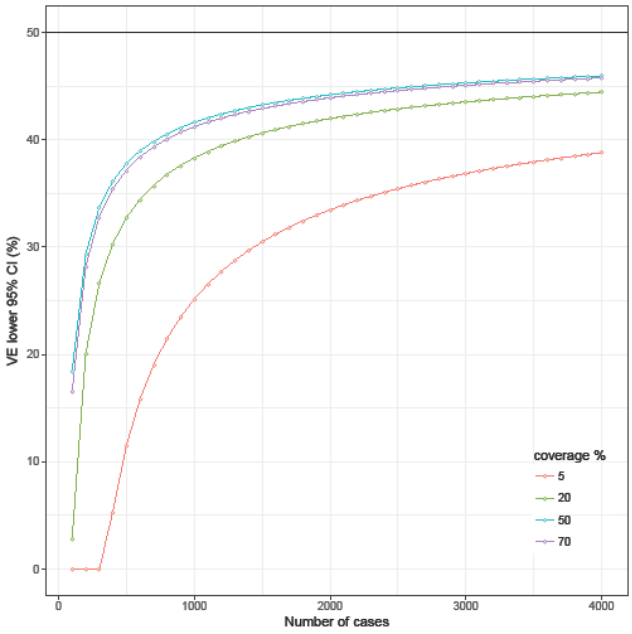


c) 70% VE

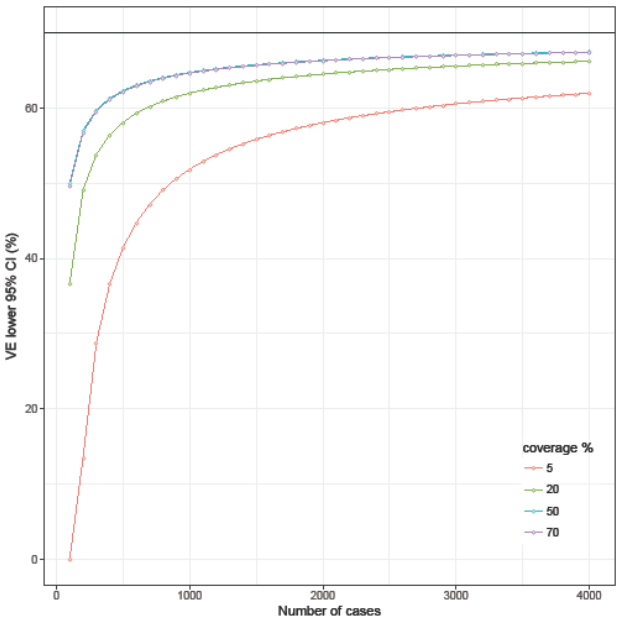
Figure 3. Precision of overall VE for a case-control study (1:1 case-control ratio) assuming overall vaccination coverage of 5%, 20%, 50% and 70%, and anticipated true VE of 20%, 50% and 70% (indicated with the black horizontal line), by number of cases.



a) 20% VE

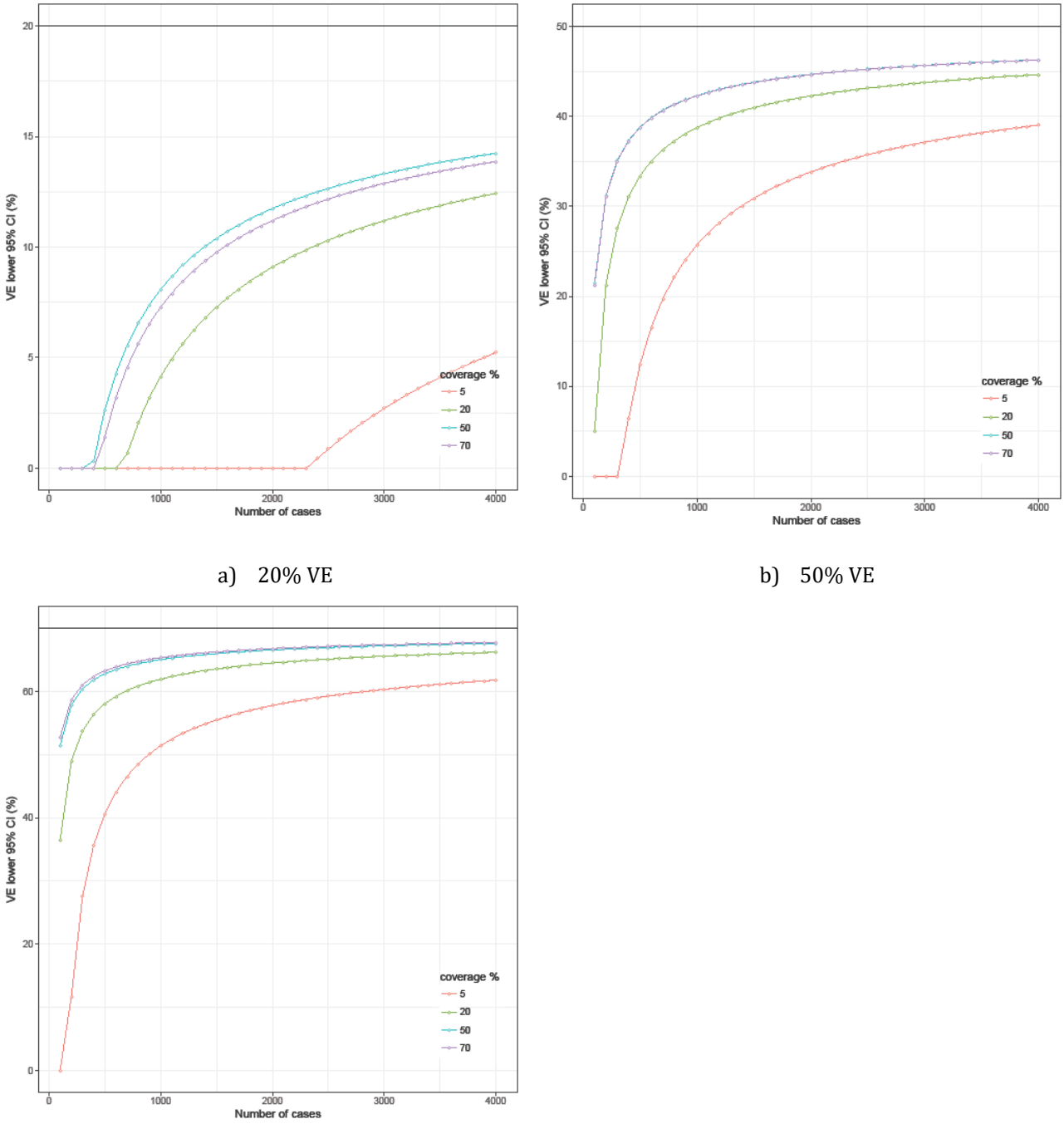


b) 50% VE



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.

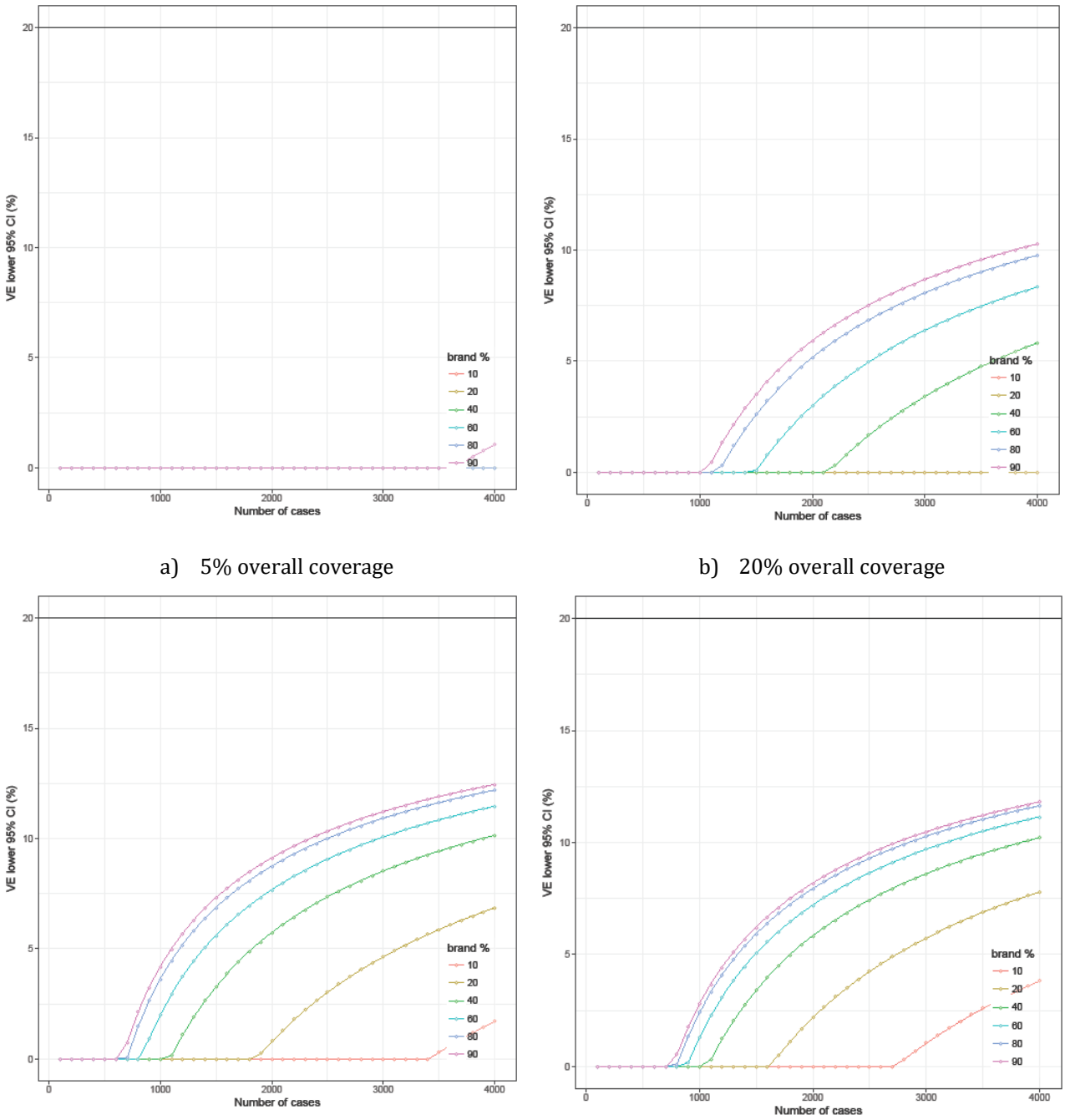


a) 20% VE

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.



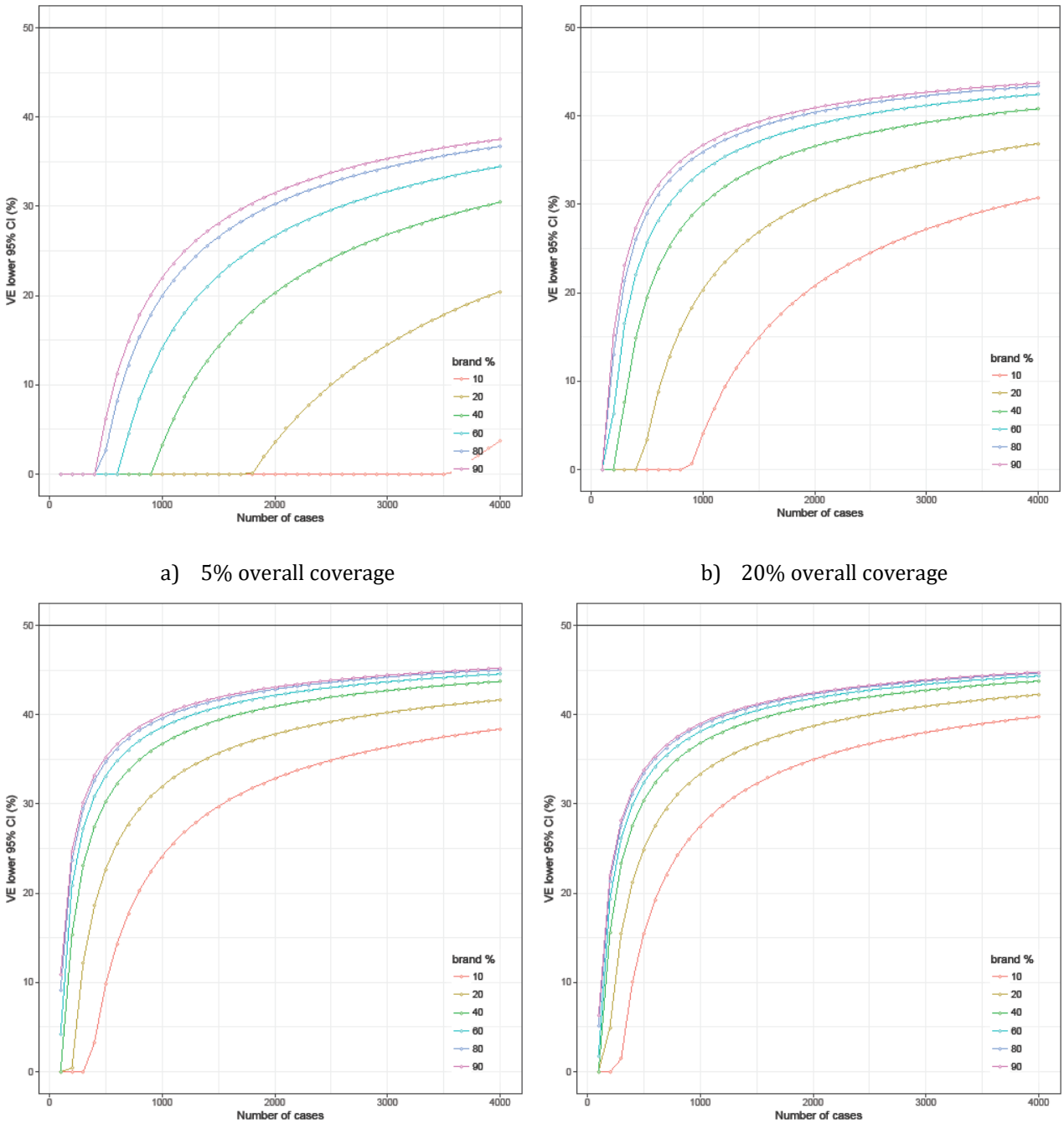
a) 5% overall coverage

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.



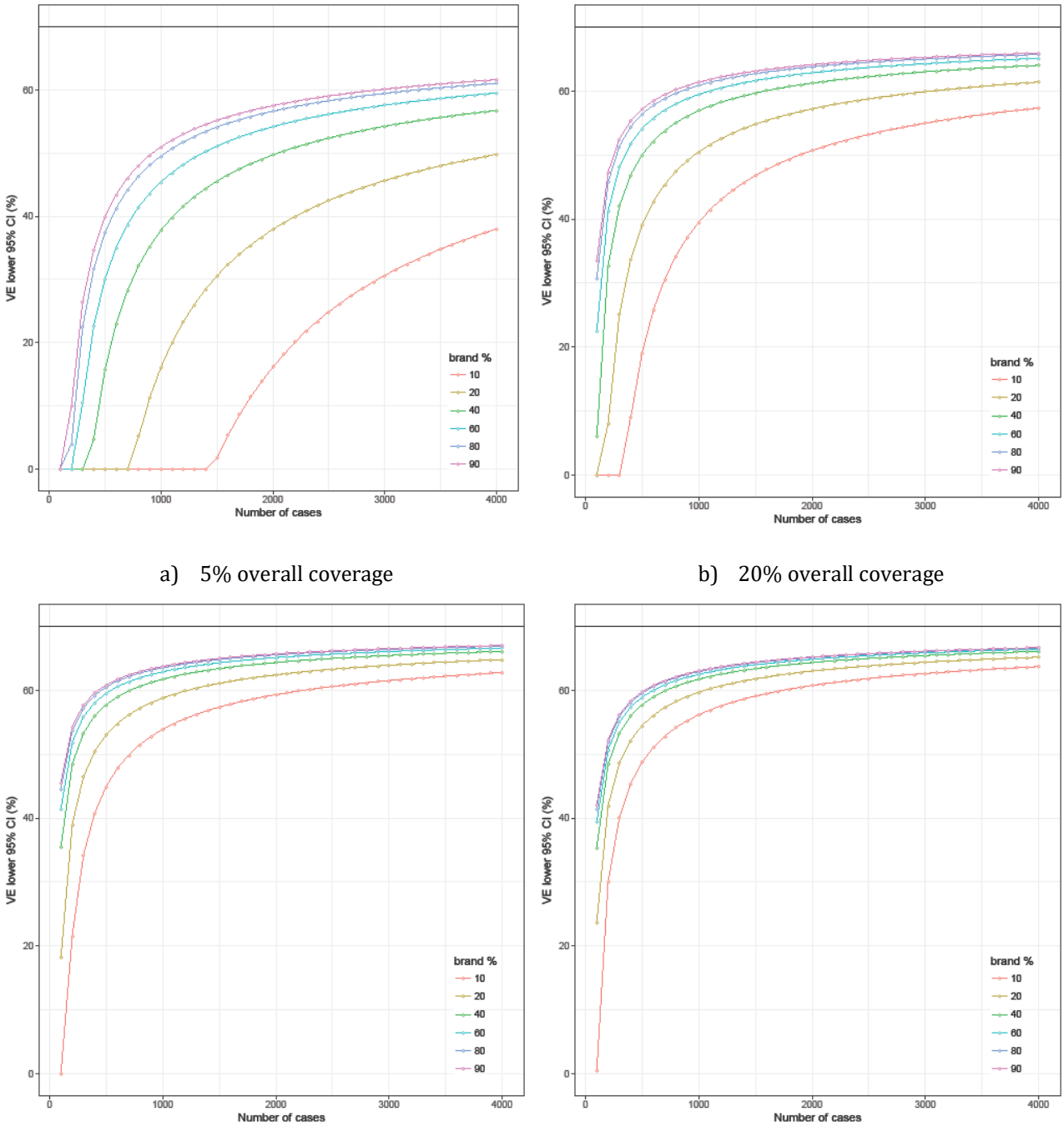
a) 5% overall coverage

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 single-site 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 study participation.

References

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

Annex 4: 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.

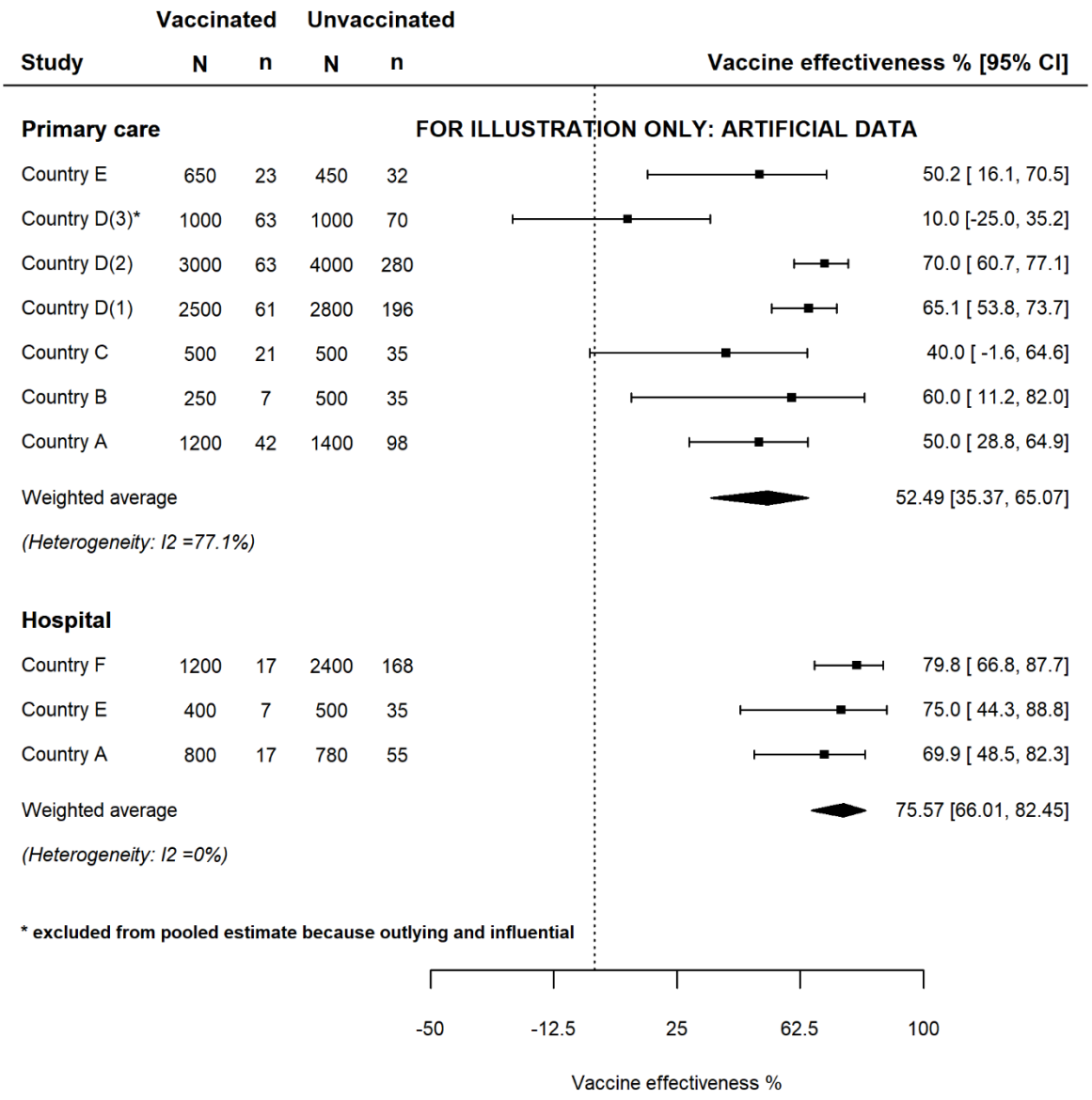


Figure 1: Forest plot and meta-analyses of influenza vaccine effectiveness, by health care setting. This plot is generated using artificial data based on cohort designs

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.