D7.3 Report on feasible, novel and innovative approaches for measuring influenza VE

777363 – DRIVE

Development of robust and innovative vaccine effectiveness

WP7 – Influenza Vaccine Effectiveness Pilot Studies

Lead contributor	Anke Stuurman (3 - P95)
	Anke.stuurman@p-95.com
Other contributors	Ritva Syrjänen, Ulrike Baum, Hanna Nohynek (6 - THL) Topi Turunen, Ainara Mira-Iglesias (1 - FISABIO) François Simondon (2 - IRD) Marga Riera, Kaatje Bollaerts (3 - P95) Miriam Levi, Roberto Bonaiuti, Alfredo Vannacci (4 - UNIFI) Uy Hoang, Harshana Liyanage, Simon de Lusignan (8 - SURREY) Bruno Lina (10 - UCBL) Caterina Rizzo (7 - ISS)
Due date	30 JUN 2018
Delivery date	31 DEC 2018
Deliverable type	R
Dissemination level	PU

Description of Work	Version	Date
	V1.1	31 DEC 2018

Document History

Version	Date	Description
V0.0	26 MAR 2018	First Draft
V0.1	18 APR 2018	First Draft – compiled comments
V0.2	03 MAY 2018	Second Draft
V0.3	07 JUN 2018	Third draft
V0.4	12 JUN 2018	Fourth draft – for ISC review
V0.5	09 NOV 2018	Fifth draft – incorporating ISC feedback and additional chapter on novel approaches
V0.6	20 NOV 2018	Sixth draft – incorporating WP7 feedback
V1.0	07 DEC 2018	Incorporating ISC feedback
V1.1	31 DEC 2018	Clarified Chapter 9, added references

Publishable Summary

Currently, measuring the effectiveness of influenza vaccines relies on observational studies which compare the occurrence of influenza in vaccinated and unvaccinated populations or the odds of vaccination in cases and non-cases. The different types of observational studies used include cohort, case-control (notably, the test-negative design) and screening method study designs.

A key question in assessing influenza vaccine effectiveness (IVE) is how to balance the inputs in terms of resources to the accuracy and generalizability of the IVE estimates. Many of the traditional observational study designs are relatively costly to establish and maintain, yet remain susceptible to bias and may not provide reliable information on all the desired outcomes. DRIVE aims to improve existing systems and explore novel and innovative approaches to measure IVE in order to promote robust IVE assessment and improve the utilization of existing data sources and new technologies. This report presents the results of an initial mapping of existing and potential innovative methods. We describe novel diagnostic methods, participatory approaches, ways to capture data on outcomes of specific interest, novel designs, non-traditional data sources, and relatively unexplored methods to control for confounding in IVE studies. We describe the potential approaches, identify the most promising ones, describe if they can be integrated in traditional data collection and how one might validate them and recommend prioritization for novel methods to be explored, within DRIVE.

List of abbreviations

AGE	Acute gastroenteritis
ARI	Acute respiratory illness
CKD	Chronic kidney disease
DRIVE	Development of robust and innovative vaccine effectiveness
ECDC	European Center for Disease Prevention and Control
HCW	Healthcare worker
ICU	Intensive care unit
ILI	Influenza-like illness
ITU	Intensive therapy unit
IVE	Influenza vaccine effectiveness
LLDB	Large linked databases
NPHI	National Public Health Institute
OWL	Web Ontology Language
RCGP	Royal College of General Practitioners
RCT	Randomized controlled trial
RIDT	Rapid influenza diagnostic test
RIMT	Rapid influenza molecular test
RSC	Research and Surveillance Center
RWE	Real-world evidence
THL	Finnish National Institute for Health and Welfare
TND	Test-negative design
VE	Vaccine effectiveness
WHO	World Health Organization

Table of Contents

1 Background and objective

The impact of influenza vaccines can be measured in many ways. While randomized studies can be considered a gold standard to determine how well a vaccine works, they are often not ethically feasible with influenza vaccines given existing vaccination recommendations nor methodologically feasible given variation in susceptibility and effectiveness between time, place and populations. The DRIVE project relies on observational studies which compare the occurrence of influenza in vaccinated and unvaccinated populations without randomization, or the odds of vaccination in cases and non-cases. Different types of observational studies include cohort, case-control (notably, the test-negative design) and screening method studies; for the characteristics of each, please refer to DRIVE D4.1: Framework for analysis of influenza vaccine effectiveness studies.

One of the key questions in assessing influenza vaccine effectiveness (IVE) is how to balance the inputs in terms of resources to the accuracy and generalizability (including over time and place) of the IVE estimates, and other factors such as timeliness, acceptability, applicability. Many of the traditional observational study designs are relatively costly to establish and maintain yet remain susceptible to bias and may not provide reliable information on all the desired outcomes. DRIVE aims to improve existing systems and explore novel and innovative approaches to measure IVE in order to promote robust IVE assessment and improve the utilization of existing data sources and new technologies.

This report presents the results of an initial mapping of existing and potential innovative methods that may be of interest to DRIVE. We describe novel diagnostic methods, participatory approaches, ways to capture data on outcomes of specific interest, novel designs, non-traditional data sources, relatively unexplored methods to control for confounding in IVE studies, and the use of ontologies for case identification. We describe the potential approaches, identify the most promising ones, describe if they can be integrated in traditional data collection and how one might validate them and recommend prioritization for novel methods to be explored within DRIVE.

This document will evolve over time. Potentially useful ideas will be collected here throughout the project and prioritized. Some of these ideas will be selected to be worked out in more detail, including a cost-assessment. Which ideas merit further exploration will be based on methodological feasibility and the extent to which we expect them to address a challenge faced

by DRIVE. We expect to have a clearer picture of the most important challenges that need to be overcome after the results of the first non-pilot study year.

2 Novel and innovative testing methods

2.1 Rapid influenza diagnostic tests

RT-PCR, a technique with high specificity and sensitivity, is considered the gold standard for laboratory confirmation of influenza. However, RT-PCR is expensive since it requires well-equipped laboratories and trained personnel; therefore there is an interest to explore less expensive alternatives.

Rapid influenza diagnostic tests (RIDT) are usually based on antigen-detection assays. RIDTs are cheaper than RT-PCR and can be used as bed-side test. On the down side, RIDTs have a low sensitivity and their use in clinical surveillance studies has demonstrated that there is a risk for a high number of false-negative results, which may lead to unneeded use of antibiotics, with adverse effects on resources and antimicrobial resistance trends. However, it is possible to correct for misclassification resulting from poor sensitivity in the statistical analysis if the test characteristics (sensitivity, specificity) are known [1].

Furthermore, rapid influenza molecular tests (RIMT) have been recently developed, however their use as bed-side tests remains limited because they need a specific device with a power supply to carry out the molecular amplification. An approach using RIMTs in primary care is described below.

2.2 Rapid near patient molecular diagnostic assays in primary care

2.2.1 The method

To estimate influenza vaccine effectiveness in primary care using rapid near patient molecular diagnostic assays.

2.2.2 What the method adds

Rapid diagnostic testing for influenza promises to influence clinical decision making due to the rapid availability of the results and hence improve patient outcomes [2]. Furthermore, as they

are cheaper than traditional testing methods, they can reduce the cost of IVE research.

It remains to be explored how rapid diagnostic molecular assays could best be integrated into diagnostic and management protocols for influenza in primary care.

2.2.3 Is this already being done?

In Europe, very rapid molecular diagnostic assays have recently been licensed for use. It can provide a result in 10 minutes and has excellent specificity, a sensitivity almost equivalent to that of the conventional PCR, and match favorably against influenza A and influenza B culture [3, 4]. However, it has only been available for purchase by hospitals and health care clinics and there is no evidence of its use in primary care, so far [5]. Some of these RIMT use a very small device, may be transportable by clinicians, and may be of use at the first point of contact for people with influenza like symptoms.

2.2.4 Potential synergies with other groups

None.

2.2.5 Can this be integrated into existing DRIVE data collection?

Yes, the data from these rapid diagnostic tests can be used to estimate influenza vaccine effectiveness in primary care across all age groups. However, these rapid diagnostic tests do not determine the subtype of the Influenza A viruses, nor the lineage of the influenza B viruses.

2.2.6 Pros and cons

The advantages of this method would include:

- The ability to obtain rapid, accurate diagnostic confirmation of the presence or absence of an influenza A/B infection, provided the negative and positive predictive value meet predefined minimum criteria
- Provide cheaper information to support clinical research including studies on vaccine effectiveness and real-world trials in primary care
- The opportunity to combine rapidly collected research information on diagnosis with changes in clinical practice/ patient management

The disadvantages would include:

- Logistical issues of implementing rapid diagnostic testing into busy clinical/ primary care practice workflows
- Impossibility to assess IVE by subtype and lineage, unless additional sample is collected and sent to specialized laboratory
- Quality issues when GPs administer the rapid test.

Thus, initially, this method would need to be limited to enthusiastic primary care pilot practices prepared to undertake intensive sampling to prove feasibility and cost compared with traditional diagnostic laboratory testing.

2.2.7 Validation

Validation of the feasibility and cost compared with diagnostic laboratory testing (RT-PCR) would be required. Validation could be achieved through double testing of patients or random allocation of rapid tests vs. RT-PCR.

2.3 Self-swabbing

2.3.1 The method

Swabs could be taken by symptomatic patients themselves and be sent for diagnostic testing.

2.3.2 What the method adds

Swabs used for diagnostic testing are typically taken by trained healthcare workers. Self-swabbing would allow influenza status determination in patients without needing to see a healthcare worker. In addition, self-swabbing would allow swabbing immediately after the onset of symptoms meeting the criteria, thus avoiding time lags in sampling leading to false negatives. Self-swabbing could be implemented in the context of other participatory methods of data collection (see 6.1, 6.3 and 6.4).

2.3.3 Is this already being done?

Self-swabbing has been tested within participatory epidemiology studies and been found feasible to obtain valid samples for analysis [6]. Personal experience in using self-swabbing for the detection of influenza viruses during antiviral treatment has taught us the results were pretty good and the infection could be documented properly for the vast majority of the self-swabs analysed (close to 98%) [7]. However, there is a need for training, or for the implementation of an e-learning source that would explain the procedure for self-swabbing. Such a resource should be developed.

2.3.4 Potential synergies with other groups

Please refer to section 6.1.4.

2.3.5 Can this be integrated into existing DRIVE data collection?

This approach would not fit within the current DRIVE data collection but could be a part of other participatory approaches proposed in this report (see sections 3.4 and 6.1).

2.3.6 Pros and cons

The advantages of this method would include:

- No need for the patient to see a healthcare worker to obtain a swab
- Capture of influenza cases that do not seek health care
- Rapid swabbing after symptom onset

The disadvantages of this method would include:

- Accuracy of the sampling technique, the reason(s) for sampling and the timing of sampling
- Need to ensure that clinical case definitions are met for those to be swabbed, in the absence of a healthcare worker
- Need for training or e-learning source that would explain the procedure for self-swabbing

2.3.7 Validation

In a subset of patients, swabs could be taken by both the patient and a healthcare practitioner,

and results between the two swabs compared.

3 Populations of special interest

In some populations, IVE data is scarce though of interest. Here we describe three such populations: pregnant women, health care workers, and persons with specific chronic conditions. These populations could be targeted for participation in clinical cohort studies due to their frequent contact with healthcare.

3.1 Pregnant women

Influenza vaccination is recommended for pregnant women (WHO recommendations, ECDC recommendations, national recommendations), although not all countries have accepted this recommendation currently, and in countries where it is recommended, the coverage is low [8]. The need for IVE data in pregnant women is described in the ECDC specific report on pregnant women and influenza [9]. Special information is needed for these groups because two persons (the mother and the unborn/newborn child) are at risk and may potentially be protected through vaccination at the same time; furthermore, pregnancy is associated with specific immunological changes which cannot be studied in other groups.

Pregnant women often have special and frequent contacts with healthcare and may have special interest to promote the wellbeing of their unborn child. This may offer the possibility to implement special study designs, such as serologic studies or self-swabbing.

3.2 Healthcare workers

Healthcare workers (HCWs) are a group of special interest because of their close contact with patients. Due to this close contact, HCWs experience a high infection pressure, especially those working in units where infectious diseases are treated. At the same time, this close contact means HCWs who are infected with influenza have a high possibility to transmit influenza to patients, especially to those vulnerable to severe disease[10, 11] . Therefore, assuming that a vaccinated healthcare worker will not only not fall ill, but will also not transmit if infected (asymptomatically), IVE estimates in HCWs are of interest both from an occupation health and from a patient safety point of view.

In Finland, the new Communicable Disease Act states that unvaccinated HCWs are allowed

to treat vulnerable patient groups only under exceptional circumstances. Because of this (at least moral) engagement, knowledge of the effects of repeated vaccination on the vaccine effectiveness and on prevention of the total individual burden of influenza over the years is important to ensure the rights of this group of HCWs.

HCWs are often in daily contact with health care organisations (researchers, laboratories), and many of them have special education and interest for clinical procedures; therefore, implementation of specific study designs may be possible among them.

An example of a study that could be conducted in these populations in described below.

3.3 Persons with specific chronic conditions

Patients with specific chronic conditions, who are recommended influenza vaccination regardless of age, are of interest for IVE studies. Information on the conditions may be available through electronic medical records (e.g. ICD codes) that can be directly linked with vaccination and infectious disease databases for VE estimation, or may be proxied from e.g. registers recording pharmaceutical purchases.

Examples of conditions of interest include

- Chronic pulmonary conditions (e.g. asthma, chronic obstructive pulmonary disease)
- Chronic cardiovascular diseases (e.g. coronary artery disease, chronic heart failure)
- Chronic neurological and neuromuscular conditions
- Diabetes
- Renal insufficiency
- Chronic liver diseases
- Chronic immunodeficiencies

3.4 Clinical cohort study with novel diagnostic approaches in specific populations

3.4.1 The method

Although very resource demanding in the traditional approach, a clinical cohort study may be feasible among specific populations with close contacts to health care organisations e.g. HCWs or pregnant women (see chapters 3.1 and 3.2), or persons with specific chronic

conditions, if novel diagnostic approaches are used. Once established, the same cohort may also be used to carry out immunological studies for improving tools for interpreting the VE results.

After enrolling the cohort, background data and the vaccination status will be assessed and updated during the follow-up. The cohort will be equipped with an influenza test kit and instructed how to self-swab for influenza (although this has been used, there may be a needed to include validation in a sub-set of the cohort), whenever pre-defined symptoms occur. Symptoms could be identified by regular follow-up (e.g. weekly calls by a study GP who checks the symptoms, or daily text messages followed by a call from a study GP in case symptoms are reported) and upon case confirmation the GP could instruct the participant to take a self-swab. and how to send the specimen to the study laboratory [6]. In addition, the participants may be instructed how to perform a rapid bed-side test after self-swabbing and to send a mobile phone picture of the result to the study staff. In addition to SMS, this could be done e.g. via an internet-based platform, through which also additional information on symptoms, missing swabs or changes in the background factors could be provided.

This design can be complemented with immunological sampling before and after the influenza season, potentially also after vaccination and/or influenza infection. If combined with covering self-swabbing, this would give information on the seroconversion after natural virus infection, and its effect to the VE estimate during the season and towards its end, to give information on to which degree natural clinical and subclinical infections explain the observed decrease in the VE estimate towards the end of the season. Including specimens for cellular immunity would further increase knowledge of the reciprocal roles of pre-existing immunity and immunity acquired from vaccination and encounter with circulating viruses.

3.4.2 What the method adds

The method allows for the collection of data to calculate IVE in groups of special interest.

Capture of influenza cases that do not seek medical care
More sensitive criteria for swabbing can be used with moderate expenses
Rapid swabbing after symptom onset

Furthermore, this type of data can provide tools for interpreting the VE estimates, e.g. to

explore the role of immunity to natural infection, and the lowering of the VE towards the end of the epidemic.

3.4.3 Is this already being done?

A clinical immunological cohort study (humoral and cellular immunity) is ongoing among vaccinated HCWs in Finland starting 2017-18, but self-swabbing and unvaccinated individuals are not included.

3.4.4 Potential synergies with other groups

This method could be complemented with enrolling and instructing a larger cohort for participatory epidemiology, i.e. web-based collection of background and vaccination data and announcement of symptoms and other clinical features of defined outcome events (see also chapter 6.1).

3.4.5 Can this be integrated into existing DRIVE data collection?

The self-swabbing could perhaps be piloted and validated in a TND study. The results could be pooled as a (nested) TND study, if the same information is collected, the background information is updated at the time of self-swabbing and the individuals with the same swabbing criteria can be identified.

3.4.6 Pros and cons

The advantages of this method would include:

- Easy contact to potentially motivated, important target groups
- Capture of influenza cases that do not seek health care
- Rapid swabbing after symptom onset

The disadvantages of this method would include:

- Accuracy of the sampling technique, the reason(s) for sampling and the timing of sampling
- Verification of vaccination status and background factors may be difficult
- Needs personnel for recruitment and training
- Taking serology samples may be considered interventional

3.4.7 Validation

The following aspects could be implemented to validate the swab and the vaccination status:

- If reliable vaccination databases are not available, a study specific vaccination card could be given at recruitment, before vaccinations have started and asked to be returned, signed by the vaccinator (or with statement of non-vaccination at the end of the epidemic; a negative status is still impossible to verify, if e.g. inter-linkable vaccine registry is not available)
- Re-sampling at the laboratory (probably possible only later when the initial phase of the disease is over)
- Validating against RT-PCR

4 Outcomes of specific interest for novel and innovative methods

4.1 Inclusion of severely ill patients

An important reason for influenza vaccination is to prevent severe illness, such as disease requiring treatment in intensive care units (ICUs) and disease in elderly which prohibits them from returning to independent life after hospitalization. However, most severely ill people may be excluded from clinical studies because they often cannot give consent, which introduces bias in the study population. Even if severely ill people are able to consent, they are often unable to give all the information required for the study, such as date of symptom onset and vaccination status, which introduces information bias. These issues have to be examined from an ethics and study conduct point of view.

The VE to prevent severe influenza cases may be assessed in population-based cohort studies using secondary data from inter-linkable routine health care registers which include vaccination data and data on e.g. reasons and length of hospitalizations and causes of death. However, information on laboratory confirmation may be lacking.

Two examples to tackle these challenges are described below.

4.2 Consent from next of kin

4.2.1 The method

One way to circumvent the difficulty of asking consent from a severely ill patient is to ask consent from the next of kin.

4.2.2 What the method adds

By asking consent from next of kin, more severely ill patients can be included in prospective clinical studies and test-negative design studies in hospital settings.

4.2.3 Is this already being done?

THL (Finland) has received a positive feedback from the Ethics Committee to ask consent from the next of kin and used it in a TND VE study for elderly. This increased the rate of severely ill cases included in the study to some degree.
In England, TND studies do not necessarily require ethics approval when they are conducted as part of routine surveillance.

4.2.4 Can this be integrated into existing DRIVE data collection?

Yes, this can be integrated into existing prospective epidemiological studies.

4.2.5 Pros and cons

The advantages of this method would include:

- Consent, which is difficult to obtain from severely ill patients, can still be obtained

The disadvantages of this method would include:

- Next of kin may not always be available
- The importance of the validation of the data provided by the next of kin from medical records is increased
- More efforts may be needed for the justification of the approach in the ethical evaluation process

4.3 Systematic swabbing in hospitals

4.3.1 The method

To improve data on severely ill patients in cohort studies, National Public Health Institutes (NPHIs) could promote enhanced and systematic swabbing in the ICUs and/or among hospitals (e.g. for elderly or risk groups), as well as systematic collection of vaccination data.

4.3.2 What the method adds

Systematic swabbing and collection of vaccination data would improve the data retrieved from registers, enabling the calculation of more accurate IVE estimates, but removing the need for individual consent.

4.3.3 Is this already being done?

THL has received a positive opinion from the ethics committee to swab the most severely ill patients for national influenza surveillance purposes without individual consent. However, this has currently not been implemented for IVE assessment.

4.3.4 Potential synergies with other groups

Collaborations with hospitals are crucial for this approach to succeed.

4.3.5 Can this be integrated into existing DRIVE data collection?

This could be integrated into existing DRIVE data collection in those sites where data is also used for national or regional surveillance.

4.3.6 Pros and cons

The advantages of this method would include:

- Individual consent, which is difficult to obtain from severely ill patients, may no longer necessary (depending on the interpretation of legislation)
- Data for patients beyond those with a primary suspicion of influenza disease would be swabbed (e.g. those with cardiac attack)

The disadvantages of this method would include:

There may be administrative and /or economical hinders for implementing an enhanced systematic swabbing.

4.4 Active monitoring of the systematic swabbing may be needed? Cases that do not seek medical care

Most studies IVE studies take place in, or are based on data from, general practice, emergency departments or hospitals, and therefore estimate IVE against medically-attended outcomes. Cases that do not seek medical care are not captured. These cases may be included in IVE studies through participatory epidemiology (see section 6.1).

4.5 Non-specific influenza outcomes to estimate influenza VE against influenza

4.5.1 The method

When studying influenza VE it is essential to accurately identify cases of the vaccine-preventable disease, often making laboratory confirmation a requirement for influenza VE studies. However, laboratory tests make influenza VE studies expensive. These costs might get substantial when large sample sizes are needed, such as for studies aiming to estimate brand-specific influenza VE. Therefore, DRIVE has a strong interest in alternative ways of estimating influenza VE.

We propose to estimate influenza VE against influenza based on less specific influenza outcomes, such as influenza like illness (ILI). However, ILI is caused by several respiratory pathogens, including influenza but also others like RSV and parainfluenza. This implies outcome misclassification, with many of the ILI cases not being influenza. However, data from laboratory surveillance might be used to predict the ILI seasonal trends [12] based on which estimates of the positive predictive value (PPV) of influenza given ILI can be obtained. The PPV will vary over the influenza season, with the highest PPV values expected during the peak of the influenza epidemic. Therefore, it is important to estimate PPV as a function of time. These time-specific PPV estimates can then be used to correct the VE estimates based on ILI for outcome misclassification. Several methods to correct for outcome misclassification exist,

including probabilistic bias analysis [13], Bayesian methods [14], likelihood-based methods [15] as well as imputation methods [16]. Alternatively, one might consider estimating influenza VE based on ILI, using only ILI cases that occurred during the peak of the influenza season, when misclassification is expected to be low and PPV highest. These alternative ways of estimating influenza VE against influenza should be validated through benchmarking the results with results obtained using laboratory-confirmed influenza.

4.5.2 What the method adds

Less expensive way of estimating influenza VE compared to a traditional TND, as swabbing will only be necessary for a subset of the ILI patients to calculate the PPV over time.

In addition, targeting non-specific outcomes gives more information on the value of vaccination in reducing the total disease burden caused by influenza than estimating the VE against laboratory confirmed influenza only, provided that the higher incidence and the costs of non-specific disease entities are considered when interpreting the presumably lower VE estimates against non-specific than specific outcome events/episodes [17]. This approach is very important from a public health point of view.

4.5.3 Is this already being done

ILI has been used as an outcome in a number of clinical trials (such as those summarized by the Cochrane Collaboration [18]) and burden of disease studies [19]. THL uses ILI (defined as certain ICD-10 codes) as one outcome of their population-based database IVE studies.

4.5.4 Potential synergies

Many public health institutes have GP sentinel surveillance to monitor ILI complemented with laboratory surveillance. In population-based cohort studies using secondary data from existing databases, changing from non-specific endpoints to the specific ones causes only minor to moderate increase in the resources needed.

4.5.5 Can this be integrated into existing DRIVE collection

Yes, data on ILI and laboratory surveillance of the geographical area are required. Addressing

non-specific endpoints is one of the secondary objectives of the population-based database cohort study protocol (D7.2).

4.5.6 Pros and cons

The advantages of this method include:

- Less expensive since it would not require systematic laboratory testing
- Results of high relevance for public health

The disadvantages of this method include:

- Results need to be interpreted carefully, and taking into consideration influenza laboratory surveillance data for the season
- No type or strain-specific data will be generated.

4.5.7 Validation

This approach would need to be validated against laboratory-confirmed outcomes.

5 Novel and innovative design

5.1 Adaptive design

5.1.1 The method

An adaptive design refers to a process in which earlier findings influence on later stages of the process. Adaptive designs help researchers to reduce the overall amount of collected data needed for the analysis. This approach is usually followed in clinical trials and DRIVE has the possibility to explore its validation and relevance in observational studies for influenza vaccine effectiveness estimation.

Although a Bayesian perspective is not especially necessary for adaptive designs, is the most natural way to follow this design. Bayesian approaches constantly update probabilities and data can be explored as being collected or performing interim analysis. These analyses are useful to compare exposure groups at any time before the end of the recruitment process. The frequency and timing of interim analysis should be specified in advance when conducting it following classical (frequentist) sequential methods due to the likelihood of type I error increases as one looks at the data more frequently. However, Bayesian analyses can be performed at any point of the study [20].

There are different types of adaptive design. What we can perform with the data collected by DRIVE is a kind of adaptive design called *internal pilot study* or *sample size reassessment*. The calculation of an appropriate sample size to achieve the desired level of statistical power is always implicit when planning a study. Power calculations require specifying an effect size and estimating “nuisance” parameters, for example, the overall incidence of the outcome. In observational studies, the rate of the exposure must be estimated as an additional source of randomness. A poor estimate of any of those parameters will produce an erroneous sample size calculation [21].

Internal pilot designs use a revised “nuisance” parameter estimate at an interim stage of the study that allows the re-adjustment of the final sample size. This approach selects a sample size sufficiently large to achieve the desired power without using unnecessary resources [21].

5.1.2 Is this already being done?

Although several groups have been publishing about adaptive designs [22-25], there are not many references that relate adaptive design and observational studies [21, 26].

5.1.3 Pros and cons

Main drawbacks for this design could be found when establishing the timing and frequency of the interim analysis if the analysis is done following classical statistics. In the Bayesian approach we have to choose an appropriate prior distribution that could be non-informative but which we can update with our data.

Frequentist approach estimates if the difference between exposure status is statistically significant whereas Bayesian methodology estimates the magnitude of this difference. Moreover, previous knowledge can be incorporated in the Bayesian approach and researchers can obtain an indication of how the new information modifies their previous belief. Frequentist statistical inferences are designed to be done only at prespecified interim analyses or the final analysis. By contrast, one can perform a Bayesian analysis at any point in the study without incurring any statistical penalty for repeated analyses. As Bayesian approach permits constant monitoring of data, investigators have the flexibility to refine the study [20].

Internal pilot designs reduce the risk of poor sample size estimation leading to better resource utilization [21].

5.2 Case-only studies

5.2.1 The method

To estimate influenza vaccine effectiveness in primary care and hospitalized cases using the self-controlled case series study's methodology.

5.2.2 What the method adds

The case-series method is applicable to routinely collected data at primary care or hospital level, so there is a great interest in order to solve the point related to the sample size of the

studies.

5.2.3 Is this already being done?

There is only one study reported in the literature using the case series method to estimate influenza vaccine effectiveness, during the 2009 pandemic, in Germany [27]. A second recent study has been conducted in UK evaluating the influenza vaccine effectiveness in reducing antibiotics prescription in pre-school children [28].

5.2.4 Potential synergies with other groups

Not applicable.

5.2.5 Can this be integrated into existing DRIVE data collection?

Yes, participating sites can contribute with secondary data.

5.2.6 Pros and cons

The advantages of this method would include:

- Is applicable to routinely collected data at primary care or hospital level (database at GPs level, at hospital level and/or notifications of confirmed influenza cases)
- Use of secondary data
- Increased sample size
- Obtaining VE by brand, age and target group for vaccination
- Regulators are familiar with the methodology as it has been used widely to explore vaccine safety

The disadvantages would include:

- Lack of adjustment for time-varying confounding
- Because of the timing of vaccination often before or early during the influenza epidemic, the individual control time during an epidemic may be short
- It has to rely on the unvaccinated cases to model the seasonal effect of influenza. There is some question over whether this is OK, might seasonal effects be different for the unvaccinated population?

- For cases who have influenza very early in the season, are they more or less likely to be vaccinated?
- How long do you assume that the vaccine is effective for? It is possible to have an indefinite length risk (or effectiveness) window.
- Because risk (effectiveness) windows are long it's not very powerful, so ideally a large sample size is needed.
- Results may be difficult to interpret
- In countries with low influenza vaccine coverage, difficult to have sufficient sample size
- Thus initially this method would need to be well explored from a methodological point of view as it has been used rarely to estimate Influenza VE.

5.2.7 Validation

This method would need to be well explored from a methodological point of view as it has been used rarely to estimate Influenza VE.

5.3 Analysis of adverse events where vaccine failure is treated as an adverse event

5.3.1 The method

As all medical products, no vaccine is perfectly safe or effective. Vaccine safety is monitored through pharmacovigilance systems where reporting of adverse events are regularly analysed. Systems like the Vaccines Adverse Events Reporting System (VAERS) organised by CDC with regular weekly analyses allow detection of possible safety signals [29]. In Europe these activities can be carried out through the European EudraVigilance database, which contains all suspected adverse drug reactions (ADRs), for authorised medicines in the European Economic Area (EEA), included those related to vaccines [www.adrreports.eu].

Events of (confirmed or suspected) influenza vaccine failure can be reported as adverse events. Confirmed vaccination failure is defined as the occurrence of the specific-vaccine preventable disease in a person who is appropriately and fully vaccinated taking into account the incubation period and the normal delay for the protection to be acquired as a result of immunization. Reports of suspected vaccine failure are also collected. An increase in vaccine failure reports could point to a decreased degree of protection provided by the vaccine

(although this may also be due to new diagnostic or different testing protocols). Furthermore, the use of vaccine failure reports in combination with other data sources could be explored for vaccine effectiveness studies.

5.3.2 What the method adds

Evidence related to vaccine lack of effectiveness could trigger signals that could be further investigated in specific studies with the complementary aim of better characterizing all potential variables associated to vaccination failure, both host- and vaccine-related:

1) Host-related:

(a) immunodeficiency (leading to suboptimal or even absent immune response after vaccination);

(b) age-related maturation and senescence of immune responsiveness;

(c) insufficient or suboptimal immune response (other than a defined immunodeficiency) to one or more antigenic vaccine components or vaccine strains or serotypes; this may or may not be measurable by standard laboratory tests such as serum antibody tests;

(d) interference due to other infectious agents (e.g., wild type enterovirus infection causing interference with the immune response to oral poliomyelitis vaccine);

(e) waning immunity;

(f) suboptimal health status (e.g., underlying disease, nutrition);

(g) immunological interference (e.g., maternal antibodies, administration of immunoglobulins);

(h) pre-existing infection with pathogen targeted by the vaccine or immunization during incubation period (after exposure to pathogen);

(i) immunosuppressive therapy.

2) Vaccine-related:

(a) vaccine is not 100% efficacious against included antigens;

(b) incomplete coverage of strains, serotypes, genotypes, antigenic variants or escape mutants that can cause a vaccine-preventable disease;

(c) antigenic interference or other vaccine-vaccine interactions in case of co-administered vaccines;

(d) manufacturing-related (e.g., batch variations, quality defect).

5.3.3 Is this already being done?

To the best of our knowledge, no previous study reports such an analytic methodology. Indeed, we believe that this method could represent an innovative strategy to study vaccines effectiveness. Although influenza-vaccine effectiveness evaluation starting from pharmacovigilance information could be very complex (i.e., lack of information on serologic post-immunization tests, poor quality of data included in pharmacovigilance reports, poor quality of data included in pharmacovigilance databases, poor evidences from observational studies), combining such data could increase robustness of DRIVE results regarding vaccine effectiveness.

5.3.4 Potential synergies with other groups

Synergy with marketing authorization holders would be required as they report all safety events to regulatory agencies.

5.3.5 Can this be integrated into existing DRIVE data collection?

It has been decided that DRIVE will generally not consider safety events. However, the safety event “vaccine failure” reported under European pharmacovigilance could be further examined.

5.3.6 Pros and cons

Advantages

- Links with other approaches such as participatory

Disadvantages

- No controls
- No or unstandardized laboratory confirmation
- Very incomplete reporting likely for such a frequent event as influenza vaccine failure
- High likelihood of bias
- Not fully in the scope of DRIVE, rather complementary to DRIVE
- Difficulty in evaluating vaccine effectiveness starting from pharmacovigilance observational data
- Approach may lead to delays due to difficulty in obtaining data

5.3.7 Validation

Feedback from pharmacovigilance databases on how lack of effectiveness are reported as adverse events for influenza vaccines.

Past vaccine failure report from years with known vaccine effectiveness would need to be analysed, to determine if there is indeed a signal to be observed in years of lower vaccine effectiveness.

It remains to be determined if vaccine effectiveness could be calculated by combining vaccine failure reports with other data source, such as number of vaccine doses administered, and through which methods (e.g. related to screening method).

6 Novel data sources

6.1 Participatory surveillance

Participatory approaches are characterized by active involvement of the study population. Here we describe the online participatory disease surveillance platform Influenzanet; we look at the added value of collecting detailed data albeit on a limited number of subjects; and describe populations that could be of specific interest for IVE studies through clinical cohort studies with active participation of the participants.

6.1.1 The method

The wide-spread use of internet in the general population has allowed for the development of online participatory disease surveillance, the largest of which in Europe is Influenzanet. On other continents similar initiatives exist such as Flu Near You in the United States [30] and Flutracking in Australia [31].

Influenzanet is a “system to monitor the activity of influenza-like-illness (ILI) with the aid of volunteers via the internet” [32]. It was launched in 2003 in the Netherlands and Belgium, and over the past 15 years it was expanded to include 11 European countries - The Netherlands, Belgium, Portugal, Italy, United Kingdom, Sweden, France, Spain, Ireland, Denmark and Switzerland (Figure 1) [33]. Participation is open to anyone residing in the countries where Influenzanet is implemented; in the 2015-2016 season there were over 36,000 participants [33]. Data is collected on various medical, geographic and behavioural questions at registration, after which participants receive weekly reminders to report any symptoms. Participants receive a weekly reminder to report any symptoms. Cases are identified using the ECDC ILI case definition [32].

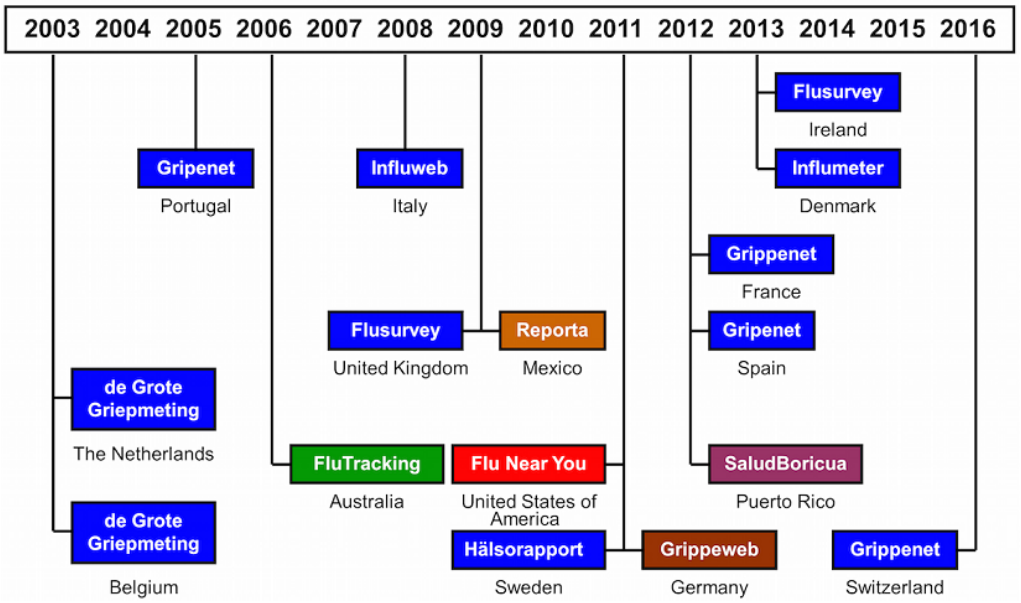

Figure 1. Timeline of Influenzanet (in blue) and other participatory surveillance systems for influenza-like-illness. Source: Koppeschaar et al. 2017 [33]

Koppeschaar et al. have summarized previous research on representativeness of Influenzanet participants compared to the general population and Influenzanet vs. traditional ILI. Research conducted in different countries and years showed varying results in terms of Influenzanet participants representativeness of the general population. Quantifying the differences between participants and the general population allows to correct for them in the analyses [33]. ILI incidence as determined through Influenzanet and traditional surveillance systems were found to be positively correlated (including ILI rise, peak and decline), although Influenzanet incidence rates were systematically higher in all countries [33].

6.1.2 What the method adds

Capture of ILI cases that do not seek medical care

Traditional ILI surveillance is based on cases that seek medical care. Internet-based ILI surveillance allows to capture those cases that do not seek medical care and are therefore missed in traditional surveillance.

Furthermore, Influenzanet questions participants on their health care-seeking behaviour following the report of symptoms. This allows for the quantification of the proportion of ILI cases that seek medical care. Of note is that large between-country variations in healthcare- seeking behaviour have been observed [33].

Estimation of IVE against ILI

Influenzanet collects the self-reported vaccination status of participants, therefore IVE against ILI can be calculated by comparing ILI incidence in vaccinated and non-vaccinated participants. This has previously been done by using Influenzanet data from Flusurvey (UK) [34], Grippenet (France) [35] and De Grote Griepmeting (Netherlands) [36]. Some studies have found IVE estimates broadly comparable to those found in other studies [34, 35]. Furthermore, stratification VE estimates against ILI have been produced by risk groups such as presence of chronic condition and age >60 years [33].

Close to real-time estimates

Participants receive weekly reminders to report any symptoms; therefore, ILI incidence can be followed close to real time. This allows for rapid assessment of IVE [34].

Pilot study on virological confirmation of ILI

For test-negative designs and cohort studies undertaken within DRIVE, laboratory-confirmed influenza is the primary outcome. Currently, Influenzanet only collects information on ILI and not on laboratory-confirmed influenza.
However, Wenham et al. conducted a pilot study aiming to obtain virological confirmation of influenza infection among Flusurvey participants and concluded self-swabbing in an online cohort study is methodologically feasible [6]. Virological swabbing kits were sent to pilot study participants, and if they reported ILI they were requested to take a self-swab and return this to a laboratory for multiplex respiratory virus PCR testing [6]. All returned samples contained human DNA, consistent with the correct use of the swab. Twenty out of 21 participants who completed an evaluation form report that undertaking the swab was easy or very easy. The authors note that the cost and logistics of distributing kits may make self-swabbing impractical for routine influenza surveillance, but may prove useful for ad hoc surveys or as a supplementary tool to traditional surveillance systems [6].

6.1.3 What else could be done

Collection of product-specific vaccination information

As the aim of DRIVE is to estimate product-specific IVE, collection of details on the vaccine received will be important. In a small number of European countries, such as Finland and Slovenia, only one influenza vaccine product is procured for the adult population (see DRIVE D3.1), therefore product-specific vaccination status can be inferred from the vaccination status

(NB: neither of these two countries participates in Influenzanet). However, for most countries this will not be feasible. For this reason, an alternative way of collecting the information of the vaccine received by participants will have to be developed.

(Selectively) implement virological confirmation to calculate virologically-confirmed VE
A pilot study has shown the methodological feasibility of obtaining self-swabs of Influenzanet participants with ILI[6]. Exploring possibilities of implementing this at a sufficiently large scale to enable the calculation of VE against virologically-confirmed influenza would be of interest.

6.1.4 Potential synergies

This approach could be complemented with the self-swabbing approach. Preliminary contacts have been made in Italy with Influenzanet coordinators [37] to explore their availability in including the evaluation of VE against virologically-confirmed influenza as already tested in UK [6]. There is also the potential of implementing self-swabbing in some other EU countries participating in Influenzanet.

6.1.5 Can this be integrated into existing DRIVE data collection?

Could be part of a TND study or clinical cohort study, if their protocol is adapted to the DRIVE requests, including the collection of validated data from their vaccine clinic on the specific vaccine brand used and on the date of the vaccine. Influenzanet can be invited to submit a proposal through the DRIVE tender process.

6.1.6 Pros and cons

The advantages of this method would include:

- Ability to capture people who do not seek medical care
    - More complete picture of number of cases

  Elimination of between-country bias caused by differences in healthcare-seeking behaviour
- Ability to analyse ILI data close to real-time
- Less costly and less resource intensive than traditional TND studies

The disadvantages of this method would include:

- Issues of representativeness of underlying study population and case severity

- Outcome:
  - Currently only ILI is measured, with no virological confirmation
  - Implementing virological confirmation may be expensive and logistically challenging. To be explored if it is more expensive than TND.
- Vaccination status:
  - Self-report, no verification of vaccination status against medical records
  - Currently no info on vaccine brand used
- Influenzanet is present in 10 European countries (in some of which they collaborate with the public health institute), but not in all countries of interest to DRIVE

6.1.7 Validation

- Validation of brand-specific vaccination status will be required.

TND and cohort studies estimating influenza VE are considered the golden standard. Pilots using participatory surveillance with brand-specific VE against virologically-confirmed influenza should be performed in countries where cohort or TND studies are also being conducted to validate this approach.

6.2 Syndromic surveillance

Traditional influenza surveillance and vaccine effectiveness studies use data from people who attend formal medical care (potentially missing a large proportion of the influenza disease burden) and may be subject to delay in obtaining the data. Various proxy indicators have been proposed to complement the picture provided by the traditional systems.

6.2.1 Insurance claims data

This has been described by Viboud et al. [38] and seems to provide good alignment with the timing and amplitude of proven influenza activity.

- Advantages: Good spatial resolution & potentially timeliness, preliminary studies report good correlation with traditional surveillance, information is individual level
- Disadvantages: Reliant on a certain kind of healthcare financing/reporting system, may be costly (if using proprietary databases), only tracks visits to formal medical care

- Feasibility for IVE studies: possible if vaccination data available (e.g. by linking with a vaccination register).

6.2.2 School absenteeism

Association has been noted between the school-reported absence prevalence and ILI surveillance in the UK [39]. An alternate system using students' personal smart cards and telephone queries was described in Hong Kong [40]. It has been proposed that school absenteeism data would be useful in guiding public health interventions such as school closure in the event of a severe epidemic.

- Advantages: Reaches a population that may otherwise missed by traditional surveillance
- Disadvantages: Not all areas have feasible tracking systems, presumably high heterogeneity in systems, highly unspecific outcome, children not a targeted group of influenza vaccination in many countries
- Feasibility for IVE studies: In order to be of use, the system would need to be highly automated. The low specificity of the outcome might mean that IVE is difficult to detect.

6.2.3 Over-the-counter medication sales.

OTC sales of medicines (e.g. cough remedies) have been linked to the volume of urgent care centre visits [41] and communitywide epidemics including annual influenza epidemics [42].

- Advantages: Captures a population that is symptomatic but not necessarily visiting formal care
- Disadvantages: Large variance in data availability, only population-level data, highly unspecific outcome
- Feasibility for IVE studies: Poor (difficulty in linking with vaccination data). This is only useful for indicating when the influenza season starts.

6.2.4 Social media & search engine data.

Google Flu Trends was an attempt to use search engine queries to track influenza-like illness. It provided very timely data with good correlation with CDC ILI surveillance (with which it was

calibrated) but suffered some major failures in 2009 and 2012/13. Twitter is another web platform used in influenza surveillance [43]; its use in health research has been described in a systematic review [44]. In 2013, CDC launched a competition to use social media to predict influenza[45] and several teams' contribution is included on the FluSight website [46]. Other (mainly US-based) disease-tracking apps include Doctor Reports Illness Tracker [47] and Sickweather [48].

Advantages:

- Potentially very timely, low-cost, captures people not visiting formal care. Emerging technologies such as machine learning and AI may add to the effectiveness of these approaches.

Disadvantages:

- Only population-level data, highly unspecific outcome (e.g. search queries not necessarily related to actual disease episode)
- Feasibility for IVE studies: Poor (difficulty in linking with vaccination data)

In conclusion, many proxy indicators may be more suited to disease surveillance and prediction than studying IVE (unless they can be accommodated in hybrid systems, please see below). In particular, linking with vaccination status information is a question to address.

6.3 Enriching data with external sources: hybrid systems

6.3.1 The method

Hybrid systems that utilize both traditional and novel "big data" approaches have been discussed as a way to improve infectious disease surveillance [49]. The rationale of using the two side-by-side is in linking the potential timeliness and scale of digital media with the more specific disease confirmation by traditional surveillance systems.

Traditional means such as physician-based sentinel surveillance systems can provide individual-level information on various unspecific (ILI) or specific (LCI) outcomes but are relatively costly to maintain. Where electronic health registers are available, their information can be used in a more cost-effective way, but both approaches only reach people who attend formal medical care.

A hybrid system would combine data from the traditional studies with other sources (e.g. the proxies outlined under 6.3 and participatory epidemiology) to improve the accuracy or representativeness of IVE estimates [44].

6.3.2 Pros and cons

The data from nontraditional sources may be inherently more complex and noisy than traditional individual-level health data and may require a lot of pre-processing. In order to estimate vaccine effectiveness, a link to vaccination status information (either individual or population level) would be necessaryI didn't find more recent data than this publication from 2017. No info on the influenzanet website On the other hand, they may be more wide-ranging and affordable given the use of existing data and crowdsourcing.

6.3.3 Is it already being done?

To our knowledge, hybrid systems have not been used in VE studies. There are applications that aggregate data from various sources for the purpose of infectious disease surveillance (e.g. Healthmap [44]).

6.3.4 Validation

Any new methods must be validated if they are to be routinely used for public health purposes; this may be hampered by a circular problem (evidence is not generated because systems are not implemented; systems are not implemented due to lack of evidence). Methods using hybrid systems may also run into novel privacy concerns that need to be taken into account.

7 Novel and innovative statistical methods

7.1 High-dimensional propensity score adjustment to control for confounding in large register-based studies

7.1.1 Propensity score

Rosenbaum and Rubin defined the propensity score as 'conditional probability of assignment to a particular treatment given a vector of observed covariates' [50]. It translates into that each study subject in an observational study can be assigned an estimated probability of being exposed conditional on the measured covariates affecting the exposure and outcome of interest. Consequently, the propensity score can take only values from 0 to 1. It is often estimated using logistic regression but also other data mining techniques can be applied [51].

The propensity score is a balancing score that can be utilized to control for confounding. The conditional distribution of the observed covariates given the propensity score is the same for the exposed and the unexposed subjects [50]. In other words, in a group of subjects with similar propensity scores the covariate distribution of the exposed is similar to the covariate distribution of the unexposed. Therefore, stratification, matching, or regression adjustment for the propensity score or its subclasses are potential techniques to control for confounding caused by the observed covariates. However, although the balance on the covariates might be greater than one would expect from randomized treatment assignment, the propensity score does not allow to control for confounding caused by unmeasured or imperfectly measured covariates [52].

7.1.2 Algorithm for variable selection

The propensity score model is aimed at prediction not explanation. Thus, the significance of covariates included in the model is not important. It has been recommended to select all variables associated with both the exposure and the outcome as well as all variables related solely to the outcome. Variables that are mainly affecting the exposure but not or only little the outcome (e.g. instrumental variables) should not be included in the propensity score model though [53].
Unfortunately, there are confounding factors that cannot be measured in observational studies.

General examples are a person's frailty or healthcare seeking behaviour. However, a sufficiently large set of measurable proxies might form a good surrogate for those confounders. Schneeweiss *et al.* described and demonstrated the use of a generic algorithm that empirically identifies and automatically selects such proxies from healthcare claims databases [54].
The important steps are

- the identification of empirical candidate covariates,
- the assessment of recurrence,
- the prioritisation of covariates,
- the selection of covariates,
- and estimation of the propensity score.

In a nutshell, the idea is to identify the most frequent diagnoses, procedures, and prescriptions recorded in the available databases. For each of these, three binary variables are created that indicate the within-patient frequency (e.g. once, sporadic, or frequent [54]) of the respective diagnosis, procedure, or prescription. Subsequently, the algorithm evaluates the potential of each binary indicator variable to confound the exposure-outcome relationship and includes the most prioritised ones in the propensity score model. In addition to the automated selection, further covariates such as age, sex, and other demographics can be manually forced into the model based on the study-specific background.

7.1.3 Application in influenza vaccine effectiveness research

Propensity score adjustment is often used in pharmacoepidemiological studies and it seems to be a valuable method worth to be also considered in influenza vaccine effectiveness research [55-58] [59, 60]. Estimates originating from cohort studies are prone to confounding bias, which is difficult to control for because differences in infection pressure or healthcare seeking behaviour can hardly be measured directly. Therefore, in the presence of routine healthcare register data, high-dimensional propensity score adjustment might be a solution.

The propensity score could be easily added as a covariate independently precalculated by each study site in accordance with the available data sources into the DRIVE data collection. However, the validity of this method in infectious diseases epidemiology must be investigated first. One of the fundamental assumptions on which the propensity score definition is based on is the Stable Unit Treatment Value Assumption [50]. As the (influenza) outcome of one study subject is not necessarily unaffected by the treatment (vaccination) assignment of the other subjects, this assumption does not hold in settings with coverage levels close to or above the

herd immunity threshold. Additionally, the propensity score method only leads to unbiased estimates, when truly all confounders are included in the model. The presence and effect of unobserved confounders might violate the statement of strongly ignorable treatment assignment [50, 52]. How these issues influence the applicability of propensity score adjustment in studies estimating influenza vaccine effectiveness must still be understood.

Ultimately, the propensity score methodology is designed for binary exposures. In general, vaccination against influenza can be modelled as a binary variable. In practise, the exposure is sometimes also classified into more than two levels holding information on the time since vaccination, the number of doses received, the credibility of the vaccination status' data source, or the vaccine brand. Such exposure variables could be simplified into binary ones, although part of the information would in consequence be lost. Alternatively, a subgroup analysis could be conducted in which only two exposure levels are allowed and all study subjects with other exposure levels are excluded potentially introducing other biases.

7.2 Using negative control outcomes to detect residual confounding

7.2.1 Negative control outcomes

Negative control outcomes are alternative outcomes used in observational studies to detect residual confounding in the estimates of the effect between the exposure and outcome of interest. They must be neither directly nor indirectly be influenced by the exposure of interest. Furthermore, all observed and unobserved confounders of the actual association of interest must have the same effect on the negative control outcome as they have on the outcome of interest [61].

7.2.2 Detection of and control for residual confounding

If there appears to be an effect between the exposure of interest and the negative control outcome after adjustment for all measured covariates, the estimates of the effect between the exposure and outcome of interest are confounded. Tchetgen proposed a control outcome calibration method to correct for unobserved confounding using negative control outcomes [62].

7.2.3 Application in influenza vaccine effectiveness research

Negative control outcomes seem to be a valuable tool to check for residual confounding in influenza vaccine effectiveness estimates based on cohort studies. However, it is very difficult to define such alternative outcomes. The best attempt to start with might be to conduct a study in a season with a confirmed vaccine antigen mismatch. The negative control outcome could be influenza caused by the virus type not included in the vaccine.

Such an analysis could be easily integrated into DRIVE as it only requires a confirmed vaccine antigen mismatch season that has already been observed in the past (e.g. 2014/15) and will possibly also occur in the future. The general protocols would not need any changes or adaptations. In case other negative control outcomes fulfilling all the criteria mentioned in 7.2.1 can be found (Therefore, the suitability of RSV-positive laboratory tests should be investigated!), they should be included in the DRIVE protocols.

Instead of aiming for the correction of confounded estimates, it might be also worth to simply revise the adjustment for confounders using the propensity score methodology (see 5.2 High-dimensional propensity score adjustment to control for confounding in large register-based studies). By repeating the two steps of improving the propensity score and assessing the magnitude of residual confounding, the bias due to unobserved confounders could possibly be minimised. Further research would be crucial.

8 Ontologies

8.1 Ontological approach for identifying influenza cases across heterogenous data sources

Ontologies form a time-consuming but appropriate method for formalizing the specification of key variables used in real world evidence (RWE) studies. Ontologies may well become one of the established tools of RWE studies [63]. DRIVE could develop ontologies for key variables to improve the comparability and transparency of the same variables recorded across health systems.

8.1.1 Introduction to ontologies:

An ontology in information science is defined either as *a set of concepts and their relationships* or as the *formalization of a specification* [64]. Using the former definition such concepts need to be defined explicitly and their relationships constructed through shared conceptualization by domain experts. This is definition used in this section. Ontological approaches can be used to define key variables such as cases and also outcome measures within DRIVE.

Ontologies have been used in a range of setting where there may be inconsistency in key variables. For example, the definition of a case of acute gastroenteritis (AGE) can be challenging to detect from routine data. Not every case of diarrhea represents as case of AGE, but very often it will do, our ontological approach allows such dilemmas to be dealt with transparently. In AGE, it was found that variability in case definition may account for the difference in incidence and prevalence reported by different data sources [65]. Exactly the same challenges may be faced with ILI in DRIVE. Ontologies can help to formalize these differences in case definitions.

The semantics of the concepts within a domain can be modelled using an ontology language such as Web Ontology Language (OWL) [66]. These ontologies are independent of the actual representation of the concepts (e.g. coded clinical data) and therefore, can be used to align data originating from heterogeneous data sources.

8.1.2 Recommended three step methods:

Ontologies can enhance the case identification from routine health data sources. A three step process an ontological approach for identifying cases is described below (Figure 2) [67].
It is suggested to use this within DRIVE as it separates the conceptual (ontological stage) from the coding layer, where clinical codes are selected to identify the case or outcome measure specified. Importantly this measure includes a logical data extract step. This final step is important because it allows the investigator to document the effect off using different conceptual elements and combinations of codes. This latter step makes final code selection transparent.

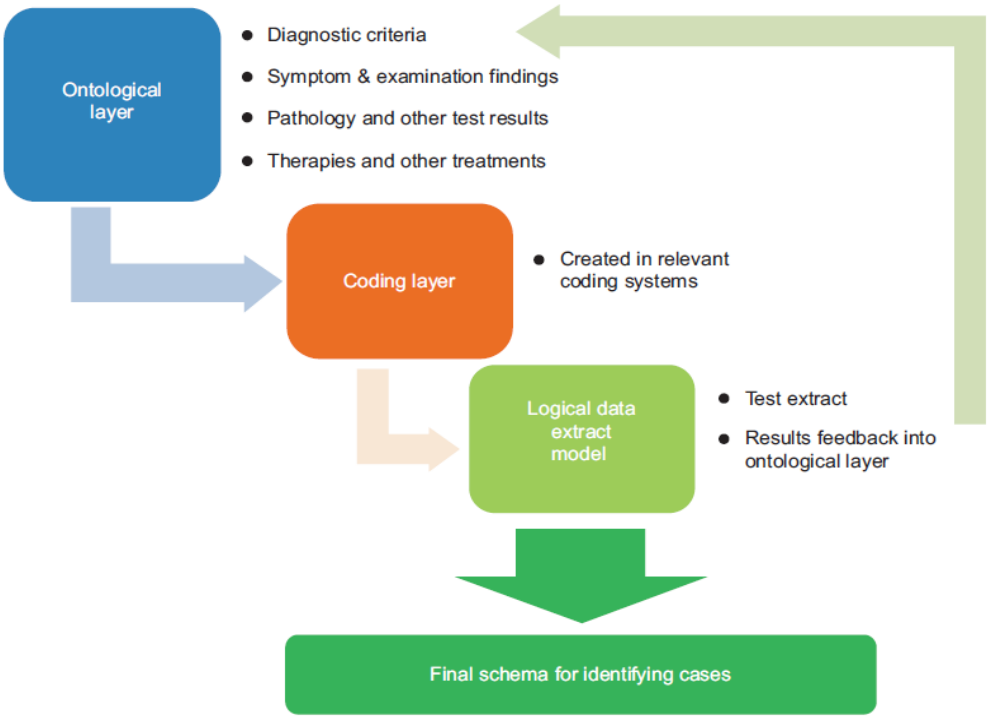

Figure 2. Three step process for ontological case identification

The three-step process in detail:

1. **Ontological layer**: Implementing the three-step process begins with the creation of an ontology defining concepts required to describe a case (e.g. Influenza ontology defines the range of concepts used to define an influenza case). The ontology will typically include concept categories such as diagnostic criteria, symptoms/ examination findings, pathology/ test results (e.g. positive virology) and therapy (e.g. exposure to flu vaccine, which may challenge the likelihood of the diagnosis, or anti-viral therapy).
2. **Coding layer:** Concepts in the ontology are then mapped to clinical coding systems used to represent the data within the data sources that require case identification. During code mapping, it may not be possible to map all ontological concepts to the codes in the target clinical coding system (e.g. therapeutic codes do not exist within ICD-10). As a result several levels of semantic equivalence between ontological concepts and the codes that they are mapped to can be defined: (1) **Direct mapping** (concept can be directly mapped to specific code(s)), (2) **Partial mapping** (concept can be mapped to a code in the coding system which is incompletely or partially representative), or (3) **No clear mapping** (concept cannot be mapped to any code(s)).
3. **Logical query layer:** Mapped ontology can be used as the basis for developing a case identification algorithm implemented within the data source. The algorithm will need to

limit false positive cases that can arise due to miscoding (code selected which lack specificity for concept), misclassification (coding indicates the right concept but had incorrect detail), misdiagnosis (coding indicates an indirect) or no coding due to: a) information not being known to anyone anywhere b) information not known to the provider but known elsewhere in the health system c) information known to provider but not coded entry recorded (e.g. free text entry or information in hospital letter).

The comprehensive definition of ontological cases allows case identification at various degrees of certainty based through the implementation of the algorithm. The algorithm and the ontology can be used to define what constitutes of definite cases (case ascertainment with a high degree of certainty - using concepts related to diagnosis), probable cases (case ascertainment with a moderate degree of certainty - using concepts related to a pattern of symptoms and signs) and possible cases (case ascertainment with a low degree of certainty - using concepts related to lab tests without clear indication of result).

8.1.3 What the method adds

This ontological approach allows to more comprehensive case identification from real world data sources. The method improves consistency of case identification across multiple sites that use different terminologies to represent data. The ontological case identification approach is capable of handling situations where different sites have varying availability of data (e.g. data sources in some sites may have diagnostic data but not prescribing data).

An ontology is flexible and offers substantial advantages over other approaches. For example, an ontology can take into account temporal changes in clinical concepts that may reflect influenza drift and may be relevant to data quality [68]. It can also be used to draw in other concepts which may improve sensitivity and specificity to maximize case ascertainment. For example, the University of Surrey has used ontologies to maximize identification of ethnicity through including concepts like language spoken [69].

Additionally, code mapping across different health systems can more accurately take account of contractual or health systems reasons why particular codes are used (or not used).

8.1.4 Is this already being done?

The ontological approach has been successfully used by the Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) for a number of case identification process such as chronic kidney disease (CKD) and influenza immunization during

pregnancy in surveillance activities in England [70, 71] in addition to the case definition for AGE set out above.

The RSC started their journey through difficulties they had in identifying cases of diabetes in routine data. Diagnosis codes, may be missing, therapy (e.g. on insulin) may indicate the patient had the condition, as might blood tests showing raised glucose [72, 73]. A range of different concepts can all imply the patient has this diagnosis, though individual codes have different mappings and diagnoses end up being definite, probably or possible [74]. Through several iterations we eventually adopted a more formally ontological approach [75].

8.1.5 Can this be integrated into existing DRIVE data collection?

This could be integrated into DRIVE data collection during future flu seasons for national or regional surveillance. Ontologies would be useful to consistently define ILI, acute respiratory illness (ARI) and vaccine exposure.

8.1.6 Pros and cons

The advantages of this method would include:

- Improved identification/definition of key variables – case definition, key exposures and outcome measures.
- Consistency and transparency in case and other key variable identification across multiple sites that use different clinical coding systems
- Ability conduct more accurate case identification when availability of coded data is limited and not sufficient to ascertain with high degree of certainty using conventional methods
- They can be revised. Ontologies can be revised, or concepts included, or not included depending on their impact on case or other variable ascertainment.

The disadvantages of this method would include:

- The time-consuming nature of ontologies. They are appropriate for case definitions, key exposures and outcome measures, but not more widely.
- Justification for revising the elements used in an ontology risks criticism.

8.1.7 Validation

Validation can be performed internally and externally. Internal validation requires detailed record examination to explore if cases are false positives, and looking at high risk individuals for false negative. External validation can be carried out by comparing the sensitivity and specificity of ILI case ascertainment against conventional case identification methods or published information. Publication of an ontology on the Bioportal website also allows others to comment, modify and use the ontology. Our CKD ontology is an example of this, see: https://bioportal.bioontology.org/ontologies/CKDO.

9 Summary and recommendations

The table below summarizes all the methods discussed in this document, with the proposed prioritization for further exploration and potential implementation within DRIVE in three categories:

High priority (1)

Medium priority (2)

Low priority (3)

The priorities were assigned by the authors of this document based on a synthesis of the pros and cons identified above. This priority-setting is considered preliminary and not prescriptive in relation to methodological innovations to be considered by DRIVE. Future priority-setting of innovative approaches (e.g. in upcoming calls for tenders) may build upon, expand or modify the principles outlined here.

Method	**Priority**
Rapid near patient molecular tests	1
Clinical cohort study with novel diagnostic approaches in specific populations	3
Consent from next of kin	2
Systematic swabbing in hospital	1
Non-specific influenza outcomes to estimate influenza VE	1
Adaptive design	3
Case-only studies	3
Analysis of adverse events where vaccine failure is treated as an adverse event	3
Participatory surveillance	2
Syndromic surveillance	3
Hybrid systems	3
High-dimensional propensity scores to control for confounding in large register-based studies	1
Using negative control outcomes to detect residual confounding	2
Ontological approach for identifying influenza cases across heterogenous data sources	1

10 References

[1] De Smedt T, Merrall E, Macina D, Perez-Vilar S, Andrews N, Bollaerts K. Bias due to differential and non-differential disease- and exposure misclassification in studies of vaccine effectiveness. PloS One. 2018;(in press).

[2] Brendish NJ, Malachira AK, Armstrong L, Houghton R, Aitken S, Nyimbili E, et al. Routine molecular point-of-care testing for respiratory viruses in adults presenting to hospital with acute respiratory illness (ResPOC): a pragmatic, open-label, randomised controlled trial. 2017;5:401-11.

[3] Merckx J, Wali R, Schiller I, Caya C, Gore GC, Chartrand C, et al. Diagnostic accuracy of novel and traditional rapid tests for influenza infection compared with reverse transcriptase polymerase chain reaction: a systematic review and meta-analysis. 2017;167:394-409.

[4] Public Health England. Point of Care Tests for Influenza and other Respiratory Viruses. London: Public Health England. Accessible: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/762344/point_of_care_tests_for_influenza_and_other_respiratory_viruses.pdf; 2018.

[5] Howick J, Cals JW, Jones C, Price CP, Plüddemann A, Heneghan C, et al. Current and future use of point-of-care tests in primary care: an international survey in Australia, Belgium, The Netherlands, the UK and the USA. 2014;4:e005611.

[6] Wenham C, Gray ER, Keane CE, Donati M, Paolotti D, Pebody R, et al. Self-Swabbing for Virological Confirmation of Influenza-Like Illness Among an Internet-Based Cohort in the UK During the 2014-2015 Flu Season: Pilot Study. Journal of medical Internet research. 2018;20:e71.

[7] Lina B, Boucher C, Osterhaus A, Monto AS, Schutten M, Whitley RJ, et al. Five years of monitoring for the emergence of oseltamivir resistance in patients with influenza A infections in the Influenza Resistance Information Study. 2018;12:267-78.

[8] ECDC. Seasonal influenza vaccination in Europe. Vaccination recommendations and coverage rates in the EU Member States for eight influenza seasons: 2007–2008 to 2014–2015. Stockholm: ECDC; 2017.

[9] ECDC. ECDC scientific advice on seasonal influenza vaccination of children and pregnant women Stockholm: ECDC; 2012.

[10] Hayward AC, Watson J. Effectiveness of influenza vaccination of staff on morbidity, and mortality of residents of long term care facilities for the elderly. Vaccine. 2011;29:2357-8.

[11] Riphagen-Dalhuisen J, Burgerhof JG, Frijstein G, van der Geest-Blankert A, Danhof-Pont M, de Jager H, et al. Hospital-based cluster randomised controlled trial to assess effects of a multi-faceted

programme on influenza vaccine coverage among hospital healthcare workers and nosocomial influenza in the Netherlands, 2009 to 2011. Eurosurveillance. 2013;18.

[12] Bollaerts K, Antoine J, Van Casteren V, Ducoffre G, Hens N, Quoilin S. Contribution of respiratory pathogens to influenza-like illness consultations. Epidemiology & Infection. 2013;141:2196-204.

[13] Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data: Springer Science & Business Media; 2011.

[14] Gustafson P. Measurement error and misclassification in statistics and epidemiology: impacts and Bayesian adjustments: CRC Press; 2003.

[15] Lyles RH, Tang L, Superak HM, King CC, Celentano DD, Lo Y, et al. Validation data-based adjustments for outcome misclassification in logistic regression: an illustration. Epidemiology (Cambridge, Mass). 2011;22:589.

[16] Cole SR, Chu H, Greenland S. Multiple-imputation for measurement-error correction. International journal of epidemiology. 2006;35:1074-81.

[17] Palmu A, Jokinen J, Nieminen H, Syrjänen R, Ruokokoski E, Puumalainen T, et al. Vaccine effectiveness of the pneumococcal Haemophilus influenzae protein D conjugate vaccine (PHiD-CV10) against clinically suspected invasive pneumococcal disease: a cluster-randomised trial. The Lancet Respiratory Medicine. 2014;2:717-27.

[18] Demicheli V, Jefferson T, Al-Ansary LA, Ferroni E, Rivetti A, Di Pietrantonj C. Vaccines for preventing influenza in healthy adults. The Cochrane Library. 2014.

[19] Nichol KL, Heilly SJ, Greenberg ME, Ehlinger E. Burden of influenza-like illness and effectiveness of influenza vaccination among working adults aged 50–64 years. Clinical Infectious Diseases. 2009;48:292-8.

[20] Brown AM, Schriger DL, Barrett TW. Outcome Measures, Interim Analyses, and Bayesian Approaches to Randomized Trials: Answers to the September 2009 Journal Club Questions. Annals of emergency medicine. 2010;55:216-24. e1.

[21] Gurka MJ, Coffey CS, Gurka KK. Internal pilots for observational studies. Biometrical Journal. 2010;52:590-603.

[22] Todd S, Whitehead A, Stallard N, Whitehead J. Interim analyses and sequential designs in phase III studies. British journal of clinical pharmacology. 2001;51:394-9.

[23] Briel M, Lane M, Montori VM, Bassler D, Glasziou P, Malaga G, et al. Stopping randomized trials early for benefit: a protocol of the Study Of Trial Policy Of Interim Truncation-2 (STOPIT-2). Trials. 2009;10:49.

[24] Jitlal M, Khan I, Lee S, Hackshaw A. Stopping clinical trials early for futility: retrospective analysis of several randomised clinical studies. British journal of cancer. 2012;107:910.
[25] Schoenfeld DA, Meade MO. Pro/con clinical debate: It is acceptable to stop large multicentre randomized controlled trials at interim analysis for futility. Critical Care. 2004;9:34.
[26] Mehtälä J, Auranen K, Kulathinal S. Optimal designs for epidemiologic longitudinal studies with binary outcomes. Statistical methods in medical research. 2015;24:803-18.
[27] Uphoff H, an der Heiden M, Schweiger B, Campe H, Beier D, Helmeke C, et al. Effectiveness of the AS03-adjuvanted vaccine against pandemic influenza virus A/(H1N1) 2009–a comparison of two methods; Germany, 2009/10. PLoS One. 2011;6:e19932.
[28] Hardelid P, Ghebremichael-Weldeselassie Y, Whitaker H, Rait G, Gilbert R, Petersen I. Effectiveness of live attenuated influenza vaccine in preventing amoxicillin prescribing in preschool children: a self-controlled case series study. Journal of Antimicrobial Chemotherapy. 2017.
[29] CDC. Chapter 21: Surveillance for Adverse Events Following Immunization Using the Vaccine Adverse Event Reporting System (VAERS). In: Roush SW, Baldy LM, editors. Manual for the Surveillance of Vaccine-Preventable Diseases. Atlanta: CDC; 2008.
[30] Flu Near You. United States.
[31] Flutracking. Australia.
[32] Influenzanet. Influenzanet: a network of European citizens fighting against influenza. 2018.
[33] Koppeschaar CE, Colizza V, Guerrisi C, Turbelin C, Duggan J, Edmunds WJ, et al. Influenzanet: Citizens Among 10 Countries Collaborating to Monitor Influenza in Europe. JMIR public health and surveillance. 2017;3.
[34] Eames K, Brooks-Pollock E, Paolotti D, Perosa M, Gioannini C, Edmunds W. Rapid assessment of influenza vaccine effectiveness: analysis of an internet-based cohort. Epidemiology & Infection. 2012;140:1309-15.
[35] Debin M, Colizza V, Blanchon T, Hanslik T, Turbelin C, Falchi A. Effectiveness of 2012–2013 influenza vaccine against influenza-like illness in general population: Estimation in a French web-based cohort. Human vaccines & immunotherapeutics. 2014;10:536-43.
[36] van Noort SP, Codeço CT, Koppeschaar CE, Van Ranst M, Paolotti D, Gomes MGM. Ten-year performance of Influenzanet: ILI time series, risks, vaccine effects, and care-seeking behaviour. Epidemics. 2015;13:28-36.
[37] Perrotta D, Bella A, Rizzo C, Paolotti D. Participatory online surveillance as a supplementary tool to sentinel doctors for influenza-like illness surveillance in italy. PloS one. 2017;12:e0169801.

[38] Viboud C, Charu V, Olson D, Ballesteros S, Gog J, Khan F, et al. Demonstrating the use of high-volume electronic medical claims data to monitor local and regional influenza activity in the US. PloS one. 2014;9:e102429.

[39] Aldridge RW, Hayward AC, Field N, Warren-Gash C, Smith C, Pebody R, et al. Are school absences correlated with influenza surveillance data in England? Results from Decipher my Data—a research project conducted through scientific engagement with schools. PloS one. 2016;11:e0146964.

[40] Cheng CK, Cowling BJ, Lau EH, Ho LM, Leung GM, Ip DK. Electronic school absenteeism monitoring and influenza surveillance, Hong Kong. Emerging infectious diseases. 2012;18:885.

[41] Liu TY, Sanders JL, Tsui F-C, Espino JU, Dato VM, Suyama J. Association of over-the-counter pharmaceutical sales with influenza-like-illnesses to patient volume in an urgent care setting. PloS one. 2013;8:e59273.

[42] Das D, Metzger K, Heffernan R, Balter S, Weiss D, Mostashari F. Monitoring over-the-counter medication sales for early detection of disease outbreaks—New York City. MMWR Morb Mortal Wkly Rep. 2005;54:41-6.

[43] Fluoutlook - Epidemic Forecasting Observatory.

[44] Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: a systematic review. American journal of public health. 2017;107:e1-e8.

[45] CDC. CDC Competition Encourages Use of Social Media to Predict Flu. Atlanta: CDC; 2013.

[46] Epidemic Prediction Initiative.

[47] Doctors Report Illness Tracker. Knox Spencer Associates LLC.

[48] Sickweather - Sickness Forecasting & Mapping. Baltimore: Sickweather, Inc.; 2011.

[49] Simonsen L, Gog JR, Olson D, Viboud C. Infectious disease surveillance in the big data era: towards faster and locally relevant systems. The Journal of infectious diseases. 2016;214:S380-S5.

[50] Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70:41-55.

[51] Setoguchi S, Schneeweiss S, Brookhart MA, Glynn RJ, Cook EF. Evaluating uses of data mining techniques in propensity score estimation: a simulation study. Pharmacoepidemiology and drug safety. 2008;17:546-55.

[52] Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. Journal of the American statistical Association. 1984;79:516-24.

[53] Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. American journal of epidemiology. 2006;163:1149-56.

[54] Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology (Cambridge, Mass). 2009;20:512.

[55] Mannino S, Villa M, Apolone G, Weiss NS, Groth N, Aquino I, et al. Effectiveness of adjuvanted influenza vaccination in elderly subjects in northern Italy. American journal of epidemiology. 2012;176:527-33.

[56] Cheng AC, Macartney KK, Waterer GW, Kotsimbos T, Kelly PM, Blyth CC. Repeated Vaccination Does Not Appear to Impact Upon Influenza Vaccine Effectiveness Against Hospitalization With Confirmed Influenza. Clinical Infectious Diseases. 2017;64:1564-72.

[57] Groenwold R, Hoes A, Hak E. Impact of influenza vaccination on mortality risk among the elderly. European Respiratory Journal. 2009;34:56-62.

[58] Kavanagh K, Robertson C, McMenamin J. Estimates of influenza vaccine effectiveness in primary care in Scotland vary with clinical or laboratory endpoint and method—experience across the 2010/11 season. Vaccine. 2013;31:4556-63.

[59] Zhang HT, Mcgrath LJ, Wyss R, Ellis AR, Sturmer T. Improving Control for Confounding by Frailty When Estimating Influenza Vaccine Effectiveness in Older Adults. Pharmacoepidemiology and Drug Safety. 2015;24:561.

[60] Nohynek H, Baum U, Syrjänen R, Ikonen N, Sundman J, Jokinen J. Effectiveness of the live attenuated and the inactivated influenza vaccine in two-year-olds–a nationwide cohort study Finland, influenza season 2015/16. Eurosurveillance. 2016;21.

[61] Lipsitch M, Jha A, Simonsen L. Observational studies and the difficult quest for causality: lessons from vaccine effectiveness and impact studies. International journal of epidemiology. 2016;45:2060-74.

[62] Tchetgen Tchetgen E. The control outcome calibration approach for causal inference with unobserved confounding. American journal of epidemiology. 2013;179:633-40.

[63] de Lusignan S, Crawford LJJIHI. Creating and using real-world evidence to answer questions about clinical effectiveness. 2015;22:368-73.

[64] Guarino N, Oberle D, Staab S. What is an ontology? Handbook on ontologies: Springer; 2009. p. 1-17.

[65] de Lusignan S, Shinneman S, Yonova I, van Vlymen J, Elliot AJ, Bolton F, et al. An ontology to improve transparency in case definition and increase case finding of infectious intestinal disease: database study in English general practice. 2017;5.

[66] OWL Working Group. Web Ontology Language (OWL). In: (W3C) Wwwc, editor.2012.

[67] de Lusignan SJJoiihi. In this issue: ontologies a key concept in informatics and key for open definitions of cases, exposures, and outcome measures. 2015;22.
[68] Fitzner J, Qasmieh S, Mounts AW, Alexander B, Besselaar T, Briand S, et al. Revision of clinical case definitions: influenza-like illness and severe acute respiratory infection. 2018;96:122.
[69] Tippu Z, Correa A, Liyanage H, Burleigh D, McGovern A, Van Vlymen J, et al. Ethnicity recording in primary care computerised medical record systems: an ontological approach. 2017;23:799-806.
[70] Cole NI, Liyanage H, Suckling RJ, Swift PA, Gallagher H, Byford R, et al. An ontological approach to identifying cases of chronic kidney disease from routine primary care data: a cross-sectional study. 2018;19:85.
[71] Liyanage H, Williams J, Byford R, Stergioulas LJSiht, informatics. Ontologies in Big Health Data Analytics: Application to Routine Clinical Data. 2018;255:65-9.
[72] Hassan Sadek N, Sadek AR, Tahir A, Khunti K, Desombre T, de Lusignan SJIJoCP. Evaluating tools to support a new practical classification of diabetes: excellent control may represent misdiagnosis and omission from disease registers is associated with worse control. 2012;66:874-82.
[73] Rollason W, Khunti K, de Lusignan SJJoIiHI. Variation in the recording of diabetes diagnostic data in primary care computer systems: implications for the quality of care. 2009;17:113-9.
[74] Sadek AR, van Vlymen J, Khunti K, de Lusignan SJDM. Automated identification of miscoded and misclassified cases of diabetes from computer records. 2012;29:410-4.
[75] McGovern A, Hinton W, Correa A, Munro N, Whyte M, de Lusignan SJBo. Real-world evidence studies into treatment adherence, thresholds for intervention and disparities in treatment in people with type 2 diabetes in the UK. 2016;6:e012801.