TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

COHORT STUDIES OF INFLUENZA VACCINE EFFECTIVENESS

Based on Finnish Register Data

Ulrike Baum

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1575 | MEDICA – ODONTOLOGICA | TURKU 2021

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

COHORT STUDIES OF INFLUENZA VACCINE EFFECTIVENESS

Based on Finnish Register Data

Ulrike Baum

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1575 | MEDICA – ODONTOLOGICA | TURKU 2021

University of Turku

Faculty of Medicine
Department of Clinical Research
Public Health Science
Doctoral Programme in Clinical Research

Finnish Institute for Health and Welfare
Department of Health Security, and
Department of Public Health Solutions,
Helsinki, Finland

Supervised by

Professor Kari Auranen, PhD
Department of Clinical Medicine, and
Department of Mathematics and Statistics,
University of Turku, Turku, Finland

Professor Sangita Kulathinal, PhD
Department of Mathematics and Statistics,
University of Helsinki, Helsinki, Finland

Reviewed by

Professor Nick Andrews, PhD
Statistics Modelling and Economics
Department, Public Health England,
London, United Kingdom

Professor Marjo Renko, MD, PhD
Department of Paediatrics, Institute of
Clinical Medicine, University of Eastern
Finland, Kuopio, Finland

Opponent

Associate Professor Tyra Grove Krause,
MD, PhD
Department of Infectious Disease
Epidemiology, Statens Serum Institut, and

Department of Public Health, University of
Copenhagen, Copenhagen, Denmark

ISBN 978-951-29-8576-0 (PRINT)
ISBN 978-951-29-8577-7 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)
Painosalama Oy, Turku, Finland 2021

UNIVERSITY OF TURKU
Faculty of Medicine
Department of Clinical Medicine
Public Health Science
ULRIKE BAUM: Cohort Studies of Influenza Vaccine Effectiveness Based on Finnish Register Data
Doctoral Dissertation, 91 pp.
Doctoral Programme in Clinical Research
September 2021

ABSTRACT

This thesis is motivated by the Finnish policy of estimating influenza vaccine effectiveness, *i.e.*, the vaccine-attributable reduction in influenza incidence, each season from medical and demographic register data. It presents and examines methods that enable efficient use of such routinely collected data, including the assessment and control of confounding and information bias.

Register-based cohort studies are conducted to estimate the influenza vaccine effectiveness in two-year-old children and elderly aged 65 years and older. With estimates ranging from less than 0 % to 90 %, the results concerning children suggest high variability in vaccine effectiveness across different seasons, vaccines and virus types. As the cohorts of children are fairly homogeneous, confounding is deemed negligible. By contrast, the elderly cohorts are less homogeneous and confounders are thus considered. The presence of confounding is confirmed by using off-season hospitalisation for acute respiratory infection as a negative control outcome. The confounder-adjusted analysis suggests that the influenza vaccine effectiveness in elderly was greater than 0 % but did not exceed 50 % in recent seasons.

A weighted partial likelihood approach with probabilistic deletion of false positives is proposed to correct for information bias. This novel method allows unbiased estimation of vaccine effectiveness if the sensitivity of the outcome measurement and the rate at which false positives occur are known. If these parameters are unknown, the magnitude of information bias can be assessed for a range of plausible parameter values.

This thesis demonstrates the potential of the examined cohort study design to enable timely and population-based estimation of influenza vaccine effectiveness in Finland. It also calls for validation data to ensure study validity.

KEYWORDS: Cohort Studies, Confounding Factors, Influenza, Outcome Measurement Errors, Proportional Hazards Model, Registers, Vaccine Effectiveness

TURUN YLIOPISTO
Lääketieteellinen tiedekunta
Kliininen laitos
Kansanterveystiede
ULRIKE BAUM: Influenssarokotteiden suojatehon arviointi rekisteripohjaisissa kohorttitutkimuksissa
Väitöskirja, 91 s.
Turun kliininen tohtoriohjelma
Syyskuu 2021

TIIVISTELMÄ

Tämän väitöstyön lähtökohtana on Terveyden ja hyvinvoinnin laitoksen käytäntö arvioida influenssarokotteiden suojatehoa vuosittain terveydenhuollon rekisterien ja väestötietojen perusteella. Työssä tutkitaan ja kehitetään menetelmiä, jotka mahdollistavat rekistereihin rutiininomaisesti kerättyjen tietojen tehokkaan käytön niin, että sekoittavien tekijöiden ja mittausvirheiden aiheuttamia harhoja voidaan välttää.

Rekisteritietoihin perustuvissa kohorttitutkimuksissa arvioidaan influenssarokotteiden aiheuttamaa vähenemää influenssan ilmaantuvuudessa kaksivuotiailla sekä yli 64-vuotiailla. Lapsilla suojatehon arviot vaihtelevat vuosittain alle nollasta aina 90 %:iin. Koska lapsikohortit ovat melko homogeenisia, sekoittavien tekijöiden aiheuttamaa ongelmaa voidaan pitää pienenä. Ikääntyneiden kohortit ovat sitä vastoin vähemmän homogeenisia, jolloin sekoittuminen on otettava huomioon. Sekoittavien tekijöiden merkitys osoitetaan käyttämällä influenssakauden ulkopuolisia akuutteja hengitystieinfektioita negatiivisena kontrollivasteena. Influenssarokotteiden suojatehon sekoittavilla tekijöillä vakioitu arvio on ikääntyneillä 0–50 %.

Sopivasti painotettua osittaisuskottavuutta soveltamalla rakennetaan tilastollinen malli, jonka avulla voidaan korjata influenssataudin mahdollisesti epäherkkä rekisteröinti. Uusi menetelmä sallii rokotetehon harhattoman arvioinnin edellyttäen, että vastetapahtuman havaitsemisen sensitiivisyys ja väärien positiivisten ilmaantuvuus tunnetaan. Vaikka näitä parametreja ei tunnettaisi, harhan riippuvuutta niiden arvoista voidaan kuitenkin arvioida.

Tämä väitöstyö osoittaa, kuinka rekisteripohjainen kohorttiasetelma mahdollistaa influenssarokotteiden suojatehon oikea-aikaisen ja väestöpohjaisen arvioinnin Suomessa. Tutkimusasetelman oikeellisuuden selvittäminen vaatisi kuitenkin erillisiä validointiaineistoja.

AVAINSANAT: Influenssa, Kohorttitutkimukset, Mittausvirheet, Rekisterit, Rokotteen suojateho, Sekoittavat tekijät, Suhteellisten hasardien malli

UNIVERSITÄT TURKU
Medizinische Fakultät
Abteilung für Klinische Medizin
Öffentliche Gesundheitswissenschaft
ULRIKE BAUM: Kohortenstudien zur Wirksamkeit von Influenza-Impfstoffen basierend auf finnischen Registerdaten
Doktorarbeit, 91 S.
Promotionsprogramm Klinische Forschung
September 2021

ZUSAMMENFASSUNG

Die Wirksamkeit von Influenza-Impfstoffen (Verringerung der Influenza-Inzidenz durch Impfung) wird in Finnland jährlich anhand medizinischer und demografischer Registerdaten ermittelt. In dieser Arbeit werden Methoden zur effizienten Nutzung der routinemäßig gesammelten Daten einschließlich der Kontrolle von Störfaktoren und Informationsbias vorgestellt und beurteilt.

Zur Bestimmung der Impfstoffwirksamkeit bei Zweijährigen und Senioren werden registerbasierte Kohortenstudien durchgeführt. Ergebnisse von weniger als 0 % und 90 % bei Kindern deuten auf eine hohe Variabilität der Impfstoffwirksamkeit über verschiedene Jahre hin. Der Einfluss von Störfaktoren wird aufgrund der Homogenität der Kohorten als vernachlässigbar angesehen. Bei Senioren dagegen wird dieser wegen der Inhomogenität der Kohorten berücksichtigt. Er wird bestätigt durch eine Kontrollstudie, in der eine Impfung keine Auswirkung auf die Inzidenz von akuten Atemwegsinfektionen außerhalb der Influenza-Saison haben sollte. Die um den Einfluss der Störfaktoren bereinigten Ergebnisse legen nahe, dass die Impfstoffwirksamkeit bei Senioren in den letzten Jahren zwischen 0 % und 50 % lag.

Zur Korrektur des Informationsbias, verursacht durch falsch-positive Befunde oder Nichterkennen von Influenza-Fällen, wird ein gewichteter Partial-Likelihood-Ansatz mit probabilistischer Streichung der falsch-positiven Befunde vorgeschlagen. Dieser Ansatz ermöglicht nicht nur eine unverzerrte Ermittlung der Impfstoffwirksamkeit, wenn Sensitivität und falsch-positive Fehlerrate bekannt sind. Auch bei unbekannten Werten kann innerhalb eines plausiblen Bereiches zumindest das Ausmaß der Verzerrung abgeschätzt werden.

Die Arbeit zeigt das Potenzial registerbasierter Kohortenstudien, die Impfstoffwirksamkeit zeitnah und bezogen auf Finnlands Bevölkerung zu ermitteln. Um die Studiengültigkeit sicherzustellen werden Validierungsdaten benötigt.

SCHLÜSSELWÖRTER: Ergebnismessfehler, Impfstoffwirksamkeit, Influenza, Kohortenstudien, Proportionales Hazard Modell, Register, Störfaktoren

Table of Contents

Abbreviations

ARI	Acute respiratory infection
CI	Confidence interval
DAG	Directed acyclic graph
ICD-10	International Classification of Diseases, 10th revision
QLAIV	Quadrivalent live-attenuated influenza vaccine
RNA	Ribonucleid acid
TIIV	Trivalent inactivated influenza vaccine
WHO	World Health Organization

Symbols

η	Vaccine effectiveness
θ	Hazard ratio
λ	Hazard
L	Likelihood
n	Number of cases
$\tilde{n}$	Number of cases under imperfect sensitivity
π	Sensitivity (se in Publication IV)
R	Risk set
$\tilde{R}$	Risk set under imperfect sensitivity ($\tilde{N}$ in Publication IV)
S	Survival function
$\tilde{S}$	Survival function under imperfect sensitivity
t	Time
w	Weight
x	Exposure status (v in Publication IV)

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I Baum, U., Auranen, K., Kulathinal, S., Syrjänen, R., Nohynek, H., and Jokinen, J. Cohort study design for estimating the effectiveness of seasonal influenza vaccines in real time based on register data: The Finnish example. Scandinavian Journal of Public Health, 2020; 48 (3): 316-322.

II Baum, U., Kulathinal, S., Auranen, K., and Nohynek, H. Effectiveness of 2 Influenza Vaccines in Nationwide Cohorts of Finnish 2-Year-Old Children in the Seasons 2015–2016 Through 2017–2018. Clinical Infectious Diseases, 2020; 71 (8): e255-e261.

III Baum, U., Kulathinal, S., and Auranen, K. Spotlight influenza: Estimation of influenza vaccine effectiveness in elderly people with assessment of residual confounding by negative control outcomes, Finland, 2012/13 to 2019/20. Eurosurveillance, 2021; 26 (36): 2100054.

IV Baum, U., Kulathinal, S., and Auranen, K. Mitigation of biases in estimating hazard ratios under non-sensitive and non-specific observation of outcomes – applications to influenza vaccine effectiveness. Emerging Themes in Epidemiology, 2021; 18 (1): 1-10.

1 Introduction

About a century ago, when the 1918–1919 influenza pandemic widely known as the "Spanish flu" took tens of millions of lives worldwide, the only interventions at hand to prevent infection and disease were non-pharmaceutical measures, such as isolation and personal hygiene. The first human influenza vaccines were developed more than a decade later, following the identification of the influenza virus as the aetiological agent in the early 1930s. Nowadays, influenza vaccines are considered to provide the best protection against influenza. However, the virus evolves quickly and therefore the level of protection varies across the season-specific vaccine compositions.

The evaluation of the causal effects of vaccines is important because it enables regulatory authorities as well as individuals to make evidence-informed decisions on the benefits and risks of vaccination. Typically, the efficacy and safety of a vaccine are first studied in well-controlled clinical trials and later, when the vaccine is licensed, in less-controlled observational studies. One approach to conduct such observational studies is the secondary use of medical and demographic data recorded primarily for administrative purposes. The approach's potential to provide timely and inexpensive effect estimates is especially valuable when the benefit of vaccination, such as that of seasonal influenza vaccination, must be re-evaluated frequently.

Focusing on cohort studies, this thesis investigates how to measure the direct effect of a vaccine against an acute disease in the vaccinated individual based on Finnish register data. Using the yearly estimation of influenza vaccine effectiveness in large cohorts as an example, the thesis takes on the challenges associated with the register-based cohort study design and shows under which conditions those challenges can be overcome. The following review of the literature introduces relevant concepts, such as vaccine effectiveness and bias, and summarises key facts about influenza. Subsequently, the thesis' aims, materials and methods, results, discussion, and conclusions are presented.

2 Review of the Literature

2.1 Measures of the Effects of Vaccination

2.1.1 Terminology

The aims of prophylactic vaccination, a form of active immunisation, are to protect the vaccinated individual from future infection and disease, and to reduce the transmission of the aetiological agent in the population. Vaccination may thus not only have a direct effect on the vaccinated individual but also indirect effects on all other individuals in the population (Haber, Longini, and Halloran, 1991).

The terminology with which different effects of vaccination have been referred to in the literature has evolved over the years (cf. Greenwood and Yule, 1915; Haber, Longini, and Halloran, 1991; Hanquet et al., 2013). Nowadays, *vaccine efficacy* is generally understood as the protective direct effect of vaccination under optimal conditions, while *vaccine effectiveness* measures the same effect in the field, where the storage and administration of vaccines may not be standardised and where indirect effects may apply. The term *impact* is used to describe the overall effect of a vaccination programme in a population as the weighted sum of both direct and indirect effects (Haber, Longini, and Halloran, 1991; Hanquet et al., 2013; Crowcroft and Klein, 2018).

2.1.2 Causal Model and Practical Assumptions

According to the intuitive model of causality by Rubin (1974), the direct effect of vaccination on a vaccinated individual is the difference between the observed outcome and the counterfactual outcome that would have been observed if the individual had not been vaccinated. The population average of this individual treatment effect defines the vaccine efficacy and effectiveness.

In absence of counterfactual outcomes, vaccine efficacy and effectiveness are studied by comparing the incidence of the vaccine-preventable infection or disease in vaccinated individuals with the corresponding incidence in unvaccinated individuals from the same population (Halloran and Struchiner, 1991).

However, as stated by Greenwood and Yule (1915), three conditions must be met to draw valid inference from such a study. First, the study subjects must be similar in all material aspects. Second, the study population must be randomly mixing so that vaccinated and unvaccinated subjects are equally exposed to the aetiological agent. Third, the classification into vaccinated and unvaccinated subjects must not be influenced by the outcome.

2.1.3 Vaccine Failure Models

The study of vaccine efficacy and effectiveness implies the possibility of vaccine failure, *i.e.*, the occurrence of the vaccine-preventable infection or disease in a vaccinated individual (Crowcroft and Klein, 2018). The type or reason of vaccine failure is crucial in determining the appropriateness of particular measures of vaccine efficacy and effectiveness.

Smith, Rodrigues, and Fine (1984) postulated two vaccine failure models. According to the first model, vaccines are *leaky* and provide only partial protection meaning that whenever exposed to the aetiological agent all vaccinated individuals have the same probability of infection or disease. In the second model, vaccines are thought to provide perfect protection to a large number of recipients while leaving the rest of the vaccinated without any protection, so that the direct effect of vaccination could be described as *all-or-none*.

Irrespective of the model, the occurrence of such primary vaccine failure may also depend on the level of exposure to the aetiological agent (Crowcroft and Klein, 2018), which is why the second of the three conditions set out by Greenwood and Yule (1915) is a particularly important requirement in studies of vaccine efficacy and effectiveness. In addition to primary vaccine failure, secondary vaccine failure may arise from the waning of vaccine-induced immunity over time (Crowcroft and Klein, 2018).

2.1.4 Appropriate Measures of Vaccine Effectiveness

Under the leaky vaccine failure model, vaccination reduces the instantaneous failure rate, *i.e.*, the hazard of infection or disease, in the vaccinated. Appropriate measures of vaccine efficacy and effectiveness must therefore be based on the ratio of the hazards in vaccinated and unvaccinated individuals (Smith, Rodrigues, and Fine, 1984). By contrast, under the all-or-none vaccine failure model, vaccination reduces the absolute risk, *i.e.*, the probability of infection or disease in the vaccinated. Appropriate measures must thus be based on the ratio of the absolute risks in vaccinated and unvaccinated individuals (Smith, Rodrigues, and Fine, 1984). In practice, a vaccine's primary failure type is rarely known. However, if the infection or disease is rare, hazard and risk

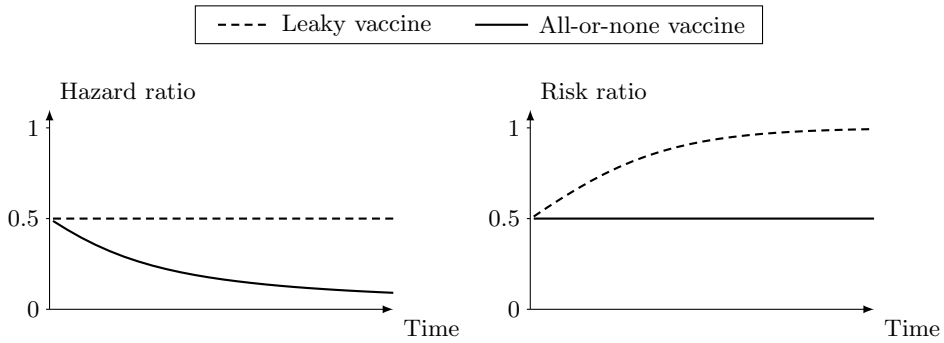

Figure 2.1: Time dependencies of measures of vaccine effectiveness. If a vaccine is leaky, the hazard ratio is time invariant but the risk ratio converges towards unity. In case of an all-or-none vaccine, the risk ratio is time invariant while the hazard ratio converges towards zero. In the present example, vaccine effectiveness is 50 %. Modified from Smith, Rodrigues, and Fine (1984) and Halloran, Longini, and Struchiner (1996).

ratios are approximately equal and can be used interchangeably.

In theory, under the assumptions of postinfection immunity and constant vaccine effectiveness, a time invariant hazard ratio indicates a leaky vaccine while a hazard ratio that decreases over time indicates an all-or-none or multimodal vaccine, the latter one being a hybrid providing vaccinated individuals with different levels of partial protection (Halloran, Longini, and Struchiner, 1996). **Figure 2.1** illustrates the time dependencies of hazard and risk ratios under the two vaccine failure models.

Nevertheless, studies often measure vaccine effectiveness based on the risk ratio irrespective of the vaccine's primary failure type and detect a decline in vaccine effectiveness over time (Tokars et al., 2020). It is important to note that this decline is not necessarily due to waning immunity. If a vaccine is leaky, the risk ratio converges towards unity because all individuals in a closed population will eventually become infected or diseased (Smith, Rodrigues, and Fine, 1984).

2.2 Threats to the Validity of Cohort Studies

2.2.1 Overview

Cohort studies are prospective observational studies in which a large number of individuals are followed over a defined period of time to study the incidence of a certain outcome. The exposure of interest may happen before or during follow-up but is, in contrast to experimental studies, not randomised (Clayton and Hills, 1993; Porta, 2008; Rothman and Greenland, 2008).

In order to draw causal inference from a study, it must be internally valid,

meaning that the study must be free from systematic errors. Given internal validity, a study might also be valid externally if the results can be generalised to populations or settings not covered in the study (Hill and Kleinbaum, 2005; Porta, 2008; Rothman, Greenland, and Lash, 2008).

Cohort studies face many threats to validity, with those pertaining to violations of internal study validity being of major concern. Therefore, the following three sections explain the systematic errors that can occur in cohort studies, *i.e.*, confounding bias, information bias, and selection bias.

2.2.2 Confounding Bias

What Is Confounding Bias and What Are Confounders?

Confounding bias is the systematic distortion of a measure of the causal relationship between an exposure and an outcome due to one or more extraneous factors (Hill and Kleinbaum, 2005; Porta, 2008; Rothman, Greenland, and Lash, 2008). These factors are commonly referred to as confounders.

It is often stated that, to qualify as a confounder, a factor must fulfill three criteria. First, it must be a causal determinant of the outcome. Second, it must be associated with the exposure. Third, it must not be a mediator in the causal relationship of interest (Greenland, Robins, and Pearl, 1999; Hill and Kleinbaum, 2005; Rothman, Greenland, and Lash, 2008). These criteria are, however, inadequate when there are multiple factors interfering with the exposure-outcome relationship because they cannot be judged separately for their potential of being confounders (Greenland, Pearl, and Robins, 1999; Rothman, Greenland, and Lash, 2008; VanderWeele and Shpitser, 2013).

As a consequence, VanderWeele and Shpitser (2013) proposed a more universal confounder definition according to which a confounder is defined as any extraneous factor that is a member of a set of factors that are together *minimally sufficient* to control for confounding. A set of factors is said to be minimally sufficient if there is no proper subset that would be sufficient by itself (Greenland, Pearl, and Robins, 1999).

Detection of Confounding and Identification of Confounders

Detection of confounding and identification of confounders should primarily be based on subject matter knowledge expressed in form of a directed acyclic graph (DAG), which is a visual representation of the underlying causal model (Greenland, Pearl, and Robins, 1999; Hernán et al., 2002). Statistical analyses alone do not warrant detection of confounding or identification of confounders since statistics alone can only reveal correlations but does not determine temporal or causal relationships (Hernán et al., 2002; VanderWeele, 2019).

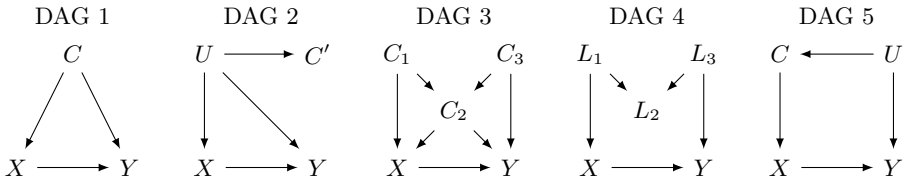

Figure 2.2: Examples of causal diagrams for the detection of confounding and identification of confounders. DAG 1: Confounding due to C. DAG 2: Confounding due to U. Controlling for C' may reduce confounding. DAG 3: Confounding due to C_1, C_2 and C_3. The sets $\{C_1,C_2\}$ and $\{C_2,C_3\}$ are both minimally sufficient. DAG 4: No confounding unless the analysis is controlled for L_2. In this case a new backdoor path opens that can only be blocked by additionally controlling for L_1 or L_3 if they are measured. DAG 5: Confounding due to C and U. Either is minimially sufficient. Restricting the definition of confounders to common causes of X and Y would miss the sufficiency of C. X, exposure; Y, outcome; C, measured confounder; U, unmeasured confounder; C', measured proxy for U; C_1, C_2 and C_3, measured confounders; L_1, L_2 and L_3, other factors. Modified from Greenland, Pearl, and Robins (1999) and VanderWeele (2019).

In a DAG, variables, *i.e.*, exposure, outcome and extraneous factors, are connected by arrows, which represent direct causal effects. Any connection from one variable to another, either direct or through other variables, is a path. A path on which two arrow heads meet in one variable is blocked. If a path follows the direction of the arrow heads, the path is directed. A variable is a cause or ancestor of another variable if there is a directed path leading from the former to the latter. In a DAG, no variable can be an ancestor of itself (Greenland, Pearl, and Robins, 1999; Hernán et al., 2002; Glymour and Greenland, 2008).

A DAG indicates the presence of confounding if, after removing all direct effects of the exposure, exposure and outcome are still connected by an unblocked backdoor path, *i.e.*, an unblocked path from the exposure through one of the exposure's direct ancestors to the outcome. If no such path exists a DAG indicates absence of confounding (Greenland, Pearl, and Robins, 1999). **Figure 2.2** illustrates four DAGs in which an uncontrolled measure of the causal relationship between exposure X and outcome Y would be confounded and one DAG in which confounding is absent.

When controlling for a factor in order to reduce or eliminate confounding, its direct effects on the exposure and outcome are removed from the DAG and all paths containing that factor are blocked. However, the stage at which the control is conducted in a study is crucial (Greenland, Pearl, and Robins, 1999). In contrast to the design-based physical block, in a cohort study only achieved by restriction, the analytical block by stratification, pooling, standardisation, or multiple regression analysis creates an association between the factor's direct ancestors potentially generating new unblocked backdoor paths

(Greenland, Pearl, and Robins, 1999; Hill and Kleinbaum, 2005). It follows that only those sets of factors that, when adequately controlled for, leave no backdoor path unblocked do sufficiently eliminate confounding (**Figure 2.2**).

In absence of profound knowledge of the underlying causal pathways required for drawing the DAG, confounders may also be identified based on other, more practical approaches, such as considering either all factors present prior to the exposure or all common causes of exposure and outcome as confounders (VanderWeele, 2019). Unfortunately, both strategies can fail in identifying factors that sufficiently control for confounding in case there are unmeasured but relevant factors as illustrated by DAG 4 and DAG 5 in **Figure 2.2**. Therefore, VanderWeele (2019) proposed another approach suggesting to control for all factors that are causes of the exposure, the outcome or both, excluding any factor not associated with the outcome except through the exposure, and for factors that are proxies for unmeasured common causes of exposure and outcome.

DAGs can also be used to express the causal relationship between an exposure, an outcome, confounders, and a negative control outcome. By definition, a negative control outcome is a factor not caused by the exposure but affected by the same factors that influence both the actual outcome and the exposure (Lipsitch, Tchetgen Tchetgen, and Cohen, 2010). Factor C', the measured proxy for an unmeasured confounder, in DAG 2 in **Figure 2.2** matches those criteria and could also depict a negative control outcome, better denoted by Y' in this case.

Negative control outcomes can be employed to empirically detect the presence of confounding bias in the estimation of an exposure-outcome (X–Y) relationship. In absence of other sources of bias, any association between the exposure and a negative control outcome indicates confounding or, after measured confounders have been controlled for, residual confounding. In practice, however, it is difficult to find a valid negative control outcome fulfilling the above criteria, as there may be unmeasured confounders that affect only either the X–Y or the X–Y' relationship but not both (Lipsitch, Tchetgen Tchetgen, and Cohen, 2010).

Direction and Magnitude of Confounding Bias

The direction and magnitude of confounding bias cannot be deduced from a DAG (Greenland, Pearl, and Robins, 1999). In general, confounding can bias the effect measure in either direction (Rothman, Greenland, and Lash, 2008). Accordingly, there are two types of confounding bias: *confounding by indication* and *healthy user bias*. Confounding by indication occurs when a causal determinant of poor health is simultaneously an indication for the

receipt of treatment, *i.e.*, the exposure. Healthy user bias occurs instead when a causal determinant of good health promotes the receipt of treatment (Porta, 2008; Remschmidt, Wichmann, and Harder, 2015).

When all confounders are measured, the magnitude of confounding bias in a study is often taken as the difference between the confounder-controlled and the crude effect measures. However, it has been shown that due to the noncollapsibility of some effect measures, such as the hazard ratio, this simple approach can be problematic because it quantifies both the effect of confounding and that of noncollapsibility. Thus, only if the latter of the two effects is small, *e.g.*, when the outcome is rare, this simple approach can be used. Otherwise advanced statistical methods, which are beyond the scope of this thesis, must be applied to disentangle the two effects (Greenland, Robins, and Pearl, 1999; Janes, Dominici, and Zeger, 2010; VanderWeele and Shpitser, 2013).

2.2.3 Information Bias

Information bias is the systematic distortion of a measure of the causal relationship between an exposure and an outcome due to measurement errors arising, *e.g.*, from intentional or unintentional inaccuracies in self-reported information or imprecise diagnostic procedures. A frequent distinction is made between exposure measurement errors and outcome measurement errors but also other study-relevant factors such as confounders may be measured inaccurately (Hill and Kleinbaum, 2005; Porta, 2008; Rothman, Greenland, and Lash, 2008).

Measurement errors in more than one factor are said to be *independent* if the measured values are statistically independent given the factors' underlying true values; otherwise those errors are said to be *dependent*. In addition, a measurement error in one factor is commonly referred to as being *nondifferential* with respect to another factor if the measurement does not depend on the latter factor's true value; otherwise the measurement error is referred to as being *differential* (Rothman, Greenland, and Lash, 2008; VanderWeele and Hernán, 2012; Tang et al., 2015).

Misclassification is a special case of measurement error. If a factor can only take one of two values, such as alive/dead, healthy/diseased or vaccinated/unvaccinated, the probabilities of misclassification are complementary to the sensitivity and specificity of the measurement procedure. The magnitude of bias due to misclassification in a study is an interplay of multiple parameters including the exposure prevalence, outcome incidence and the sensitivity and specificity of the measurement procedures (Tang et al., 2015; De Smedt et al., 2018).

The direction of information bias is generally difficult to predict, with two

exceptions. First, independent and with respect to the exposure nondifferential measurement errors in a binary outcome do not, on average, introduce bias. Second, independent and with respect to the outcome nondifferential measurement errors in a binary exposure bias the effect measure, on average, towards the null (Rothman, Greenland, and Lash, 2008; VanderWeele and Hernán, 2012; De Smedt et al., 2018). In practice, however, it can rarely be assumed that the measurement error in one factor is strictly unaffected by any other factor, even given the factor's underlying true value.

Although measurement errors can never be eliminated with certainty, their influence can be minimised already at the design stage of a study by standardising all measurement procedures, including their timing and frequency. Analytical correction of information bias, such as likelihood-based approaches or latent class modelling, requires knowledge of the causal pathways and internal or external validation data so that the error rates can be assessed. If such data are not available, simulation studies can be used to estimate the probable direction and magnitude of information bias (Hill and Kleinbaum, 2005; Tang et al., 2015; De Smedt et al., 2018).

2.2.4 Selection Bias

Selection bias is the systematic distortion of a measure of the causal relationship between an exposure and an outcome due to flaws in the process selecting the study subjects. In particular, it arises as a result of conditioning on a common effect of the exposure (or a causal determinant of the exposure) and the outcome (or a causal determinant of the outcome). Typical types of selection bias in cohort studies are differential nonresponse or loss to follow-up, and membership bias (Hernán, Hernández-Díaz, and Robins, 2004; Porta, 2008; Rothman, Greenland, and Lash, 2008).

Differential nonresponse or loss to follow-up occurs when nonresponse to a survey or noncompliance to follow-up visits is triggered by a common effect of the exposure and the outcome or their causal determinants. It leads to informative censoring and distorts the exposure-outcome distribution in the study data. If the conditional nonresponse and loss to follow-up rates are known or estimable, it may be possible to correct for this type of selection bias by measuring the association between exposure and outcome in a pseudopopulation, where uncensored observations are weighted according to the inverse probability of censoring. By constrast, membership bias can only be addressed at the design stage of a study. It arises when a factor that is a common effect of the exposure and the outcome determines the eligibility for cohort membership (Hernán, Hernández-Díaz, and Robins, 2004).

2.3 Influenza

2.3.1 Epidemiology

Influenza is a highly contagious disease caused by the influenza virus. In the human population, particularly two virus types, influenza A and influenza B, are responsible for outbreaks. Both types replicate in the epithelial cells of the human respiratory tract and spread through aerosol, droplet and contact transmission from person to person. In addition, influenza A can cross the species barrier and occasionally spreads from animals, such as pigs and birds, into the human population (Paules and Subbarao, 2017; Bresee et al., 2018).

The main targets of the human immune response are the surface proteins haemagglutinin and neuraminidase, which facilitate the host cell entry and release. These antigens continuously evolve because of the frequent occurrence of point mutations in the viral ribonucleid acid (RNA) genome and the selective pressure exerted by the host's immune response. As a result of this *antigenic drift*, immunity to influenza infection is often of temporary nature (Porta, 2008; Paules and Subbarao, 2017; Bresee et al., 2018).

Another process leading to antigenic variation is *antigenic shift*. The RNA genome of the influenza virus is segmented. If different viruses of the same type, *e.g.*, a human and an avian influenza A virus, happen to infect the same host cell, they may "exchange" genome segments and create an antigenically novel virus. Such shifts occur sporadically in influenza A viruses because of their ability to cross the species barrier. Immunity against those reassortants preexists rarely (Porta, 2008; Paules and Subbarao, 2017; Bresee et al., 2018).

Both influenza A and influenza B cause recurrent seasonal epidemics during the winter months. Influenza A additionally causes pandemics at irregular intervals as a result of antigenic shift. Based on their antigenic properties, influenza A viruses are divided into multiple subtypes of which two, A(H1N1) and A(H3N2), have recently been circulating in the human population. Influenza B viruses are subdivided into two antigenically distinct lineages: B/Victoria and B/Yamagata (Paules and Subbarao, 2017; Bresee et al., 2018).

According to a systematic review that summarised estimates of influenza reproduction numbers, one influenza case generates on average 1.3 new cases during an epidemic and up to 1.8 new cases during a pandemic (Biggerstaff et al., 2014). The cumulative risk of seasonal influenza, *i.e.*, the proportion of a population that contracts influenza during a season, can be as high as 20 % but varies greatly across seasons and age groups and depends on the case definition (Somes et al., 2018).

2.3.2 Clinical Presentation

The typical clinical picture of influenza disease is that of an acute respiratory infection (ARI) and includes various respiratory as well as systemic symptoms, such as cough, coryza and sore throat, and fever, headache and myalgia. The symptoms usually appear suddenly and persist for about a week or longer. However, influenza infection can also be asymptomatic depending on the virus' virulence and the host's immune response (Paules and Subbarao, 2017; Bresee et al., 2018).

In severe cases, influenza infection causes primary viral pneumonia or facilitates secondary bacterial pneumonia and may even lead to death. It can also exacerbate underlying chronic conditions, such as asthma and chronic obstructive pulmonary disease, or cause cardiac complications. Especially the elderly as well as young children and immunocompromised people are at high risk of severe or fatal courses of disease (Paules and Subbarao, 2017; Bresee et al., 2018).

2.3.3 Diagnosis and Treatment

Since the clinical picture of influenza is often similar to that of other ARIs, a definitive diagnosis requires laboratory confirmation. Different tests exist for the detection of either viral RNA or viral proteins in respiratory samples or in cultures based on such samples. Due to its high accuracy and relatively short turnaround time, reverse transcription polymerase chain reaction is currently considered to be the gold standard for diagnosing influenza (Paules and Subbarao, 2017; Bresee et al., 2018).

The treatment of influenza is mostly unspecific and aims at the relief of symptoms. The prophylactic and therapeutic use of influenza-specific antiviral drugs is reserved for people at high risk of severe disease to prevent the emergence of drug resistant virus populations. The antiviral drugs currently available in the European Union to treat influenza are all neuraminidase inhibitors but other agents with different mechanisms of action are under investigation (Paules and Subbarao, 2017; Bresee et al., 2018).

2.3.4 Vaccination

In analogy to natural influenza infection, vaccination with an influenza vaccine stimulates the human adaptive immune system. Seasonal influenza vaccines induce the production of antigen-specific antibodies that bind to the viral haemagglutinin and neutralise the influenza virus' replication. To cover the variety of influenza viruses, recent seasonal vaccines have been tri- or quadrivalent containing antigens of influenza A(H1N1) and influenza A(H3N2) as

well as influenza B/Victoria and/or influenza B/Yamagata viruses. The World Health Organization (WHO) regularly updates their recommended list of virus strains to be included in the next seasonal vaccine composition in reaction to the ongoing antigenic drift (Paules and Subbarao, 2017; Bresee et al., 2018).

There are two kinds of influenza vaccines: live-attenuated and inactivated influenza vaccines. The first kind is administered as a nasal spray and contains influenza virus reassortants that replicate in the nasal epithelium without causing influenza disease. The second kind is given as an injection and contains only viral components created by dissolving the viral lipid surface or using recombinant technology. In contrast to inactivated vaccines, live-attenuated vaccines also elicit a local mucosal immune response similar to that following a natural infection (Bresee et al., 2018; Luke, Lakdawala, and Subbarao, 2018).

The strength and duration of a vaccine-induced immune response depend on the vaccinee's age, history of prior infections and vaccinations, and the presence of immunocompromising conditions. Antibody levels typically peak within two to four weeks after vaccination with lower levels expected among young children and the elderly, because of immunological immaturity and immunosenescence, respectively. In addition, individuals who have not been previously exposed to influenza, either through infection or vaccination, may exhibit lower antibody levels if given a single dose only (Bresee et al., 2018).

Multiple clinical trials and observational studies have demonstrated the efficacy, effectiveness and safety of influenza vaccines, based on which the WHO recommends the annual vaccination with a seasonal influenza vaccine to health care workers and people at high risk of severe disease. Many countries including Finland have adopted this recommendation in their vaccination programmes. However, as vaccination is usually voluntary, the vaccine uptake may not be homogeneous in those populations (Mereckiene et al., 2014; Paules and Subbarao, 2017; Bresee et al., 2018).

3 Aims

The objectives of this thesis are to develop methods enabling efficient use of Finnish register data in the estimation of influenza vaccine effectiveness and to critically evaluate the methods' limitations. The specific aims are:

1. To describe the register data, study design and statistical model and to discuss potential sources of bias in the estimation of influenza vaccine effectiveness in Finland.

2. To apply the study design and statistical model in practice to estimate influenza vaccine effectiveness in two-year-old children and the elderly in Finland in recent seasons.

3. To explore which factors confound the register-based estimation of influenza vaccine effectiveness in two-year-old children and the elderly and to investigate how to assess and eliminate confounding bias.

4. To conduct a quantitative evaluation of the bias due to outcome measurement errors.

4 Materials and Methods

4.1 Data Sources and Data Linkage

All study data were retrieved exclusively from six computerised national registers maintained by the Finnish Digital and Population Data Services Agency and the Finnish Institute for Health and Welfare.

The *Population Information System* records the names, addresses, dates of birth, and dates of death (if applicable) of all Finnish citizens and of foreign citizens residing in Finland. Upon registration in the Population Information System each person is assigned a unique personal identity code. This code is widely used for administrative purposes and allows deterministic linkage of data from different national registers.

The *National Infectious Diseases Register* collects communicable disease notifications filed by doctors and laboratories on the basis of the Communicable Diseases Act. The register thus records the sample dates and virus types of all influenza-positive samples analysed in Finnish clinical microbiology laboratories as well as each patient's personal identity code.

The *Care Register for Health Care* combines patient-level administrative data on inpatient care, day surgeries, and specialised outpatient care provided in emergency rooms. Medical diagnoses are recorded using the International Classification of Diseases, 10th revision (ICD-10).

The *Register of Primary Health Care Visits* collects patient-level administrative data on outpatient care from the patient record systems of public health centres, such as well-baby clinics. The *National Vaccination Register* originates from the Register of Primary Health Care Visits and records all vaccinations given in the Finnish vaccination programme by vaccine brand and vaccination date.

The *Medical Birth Register* gathers administrative information on all births in Finland. The register includes medical as well as demographic data about each mother and child pair. The reader is referred to Publication I for further details concerning the six registers.

Table 4.1: Definition of study periods and study cohorts in Publication II.

	Study period:	Study cohort: two-year-old children	
Season	**week 40 – week 20**	Date of birth	Date of death
2015–2016	28 Sep 2015 – 22 May 2016	01 Nov 2012 – 31 Dec 2013	$\geq$ 28 Sep 2015
2016–2017	03 Oct 2016 – 21 May 2017	01 Nov 2013 – 31 Dec 2014	$\geq$ 03 Oct 2016
2017–2018	02 Oct 2017 – 20 May 2018	01 Nov 2014 – 31 Dec 2015	$\geq$ 02 Oct 2017

Table 4.2: Definition of study periods and study cohorts in Publication III.

		Study cohort: the elderly	
	Study period	Date of birth	Date of death
Season			
2012–2013	01 Oct 2012 – 31 May 2013	01 Jan 1912 – 31 Dec 1947	$\geq$ 01 Oct 2012
2013–2014	01 Oct 2013 – 31 May 2014	01 Jan 1913 – 31 Dec 1948	$\geq$ 01 Oct 2013
2014–2015	01 Oct 2014 – 31 May 2015	01 Jan 1914 – 31 Dec 1949	$\geq$ 01 Oct 2014
2015–2016	01 Oct 2015 – 31 May 2016	01 Jan 1915 – 31 Dec 1950	$\geq$ 01 Oct 2015
2016–2017	01 Oct 2016 – 31 May 2017	01 Jan 1916 – 31 Dec 1951	$\geq$ 01 Oct 2016
2017–2018	01 Oct 2017 – 31 May 2018	01 Jan 1917 – 31 Dec 1952	$\geq$ 01 Oct 2017
2018–2019	01 Oct 2018 – 31 May 2019	01 Jan 1918 – 31 Dec 1953	$\geq$ 01 Oct 2018
2019–2020	01 Oct 2019 – 31 May 2020	01 Jan 1919 – 31 Dec 1954	$\geq$ 01 Oct 2019
Off-season			
2013	01 Jul 2013 – 30 Sep 2013	01 Jan 1912 – 31 Dec 1947	$\geq$ 01 Jul 2013
2014	01 Jul 2014 – 30 Sep 2014	01 Jan 1913 – 31 Dec 1948	$\geq$ 01 Jul 2014
2015	01 Jul 2015 – 30 Sep 2015	01 Jan 1914 – 31 Dec 1949	$\geq$ 01 Jul 2015
2016	01 Jul 2016 – 30 Sep 2016	01 Jan 1915 – 31 Dec 1950	$\geq$ 01 Jul 2016
2017	01 Jul 2017 – 30 Sep 2017	01 Jan 1916 – 31 Dec 1951	$\geq$ 01 Jul 2017
2018	01 Jul 2018 – 30 Sep 2018	01 Jan 1917 – 31 Dec 1952	$\geq$ 01 Jul 2018
2019	01 Jul 2019 – 30 Sep 2019	01 Jan 1918 – 31 Dec 1953	$\geq$ 01 Jul 2019

4.2 Register-Based Cohort Study Design

4.2.1 Study Periods and Study Cohorts

The study periods were the eight influenza seasons from 2012–2013 through 2019–2020 and the seven influenza off-seasons from 2013 through 2019. The two populations chosen as the targets of inference were two-year-old children and the elderly aged 65 years and older, two age groups eligible for free seasonal influenza vaccination in Finland. **Table 4.1** and **Table 4.2** show the definitions of the study periods as well as the criteria for inclusion into the study cohorts. Ideally, the study cohorts were to comprise all individuals of the respective age registered in the Finnish Population Information System. Individuals who lived abroad or outside the National Vaccination Register's catchment area were however excluded.

4.2.2 Outcome Definitions

The primary outcome was laboratory-confirmed influenza infection, further identified by virus type as influenza A or influenza B. Respiratory samples were collected as part of routine clinical procedures of inpatient and outpatient care. Laboratory confirmation was obtained using reverse transcription polymerase chain reaction or antigen detection assays.

The secondary outcome was hospitalisation for ARI defined as any inpatient hospitalisation or emergency room visit during which at least one of the following medical conditions were diagnosed: acute upper respiratory infection (ICD-10 diagnostic codes J00–J06), pneumonia (J12–J18), acute lower respiratory infection (J20–J22), chronic obstructive pulmonary disease with acute lower respiratory infection (J44.0), cough (R05), or unspecified fever (R50.9). This unspecific outcome, similar in clinical picture to severe influenza disease, was taken as a negative control outcome during the influenza off-seasons.

4.2.3 Exposure Definitions

The exposure of interest was seasonal influenza vaccination. To account for the fact that the development of vaccine-induced immunity may take up to two weeks, three time-dependent exposure states were defined: unvaccinated, vaccinated less than or exactly two weeks ago ("transitionally or partially vaccinated"), and vaccinated more than two weeks ago ("fully vaccinated").

From 2015–2016 through 2017–2018, two-year-old children were vaccinated with quadrivalent live-attenuated influenza vaccine (QLAIV) or trivalent inactivated influenza vaccine (TIIV). In Publication II, we therefore distinguished between vaccination with QLAIV and vaccination with TIIV. In addition, we extended the transitional exposure state for those TIIV recipients who had not been previously vaccinated with at least one dose of QLAIV or two doses of TIIV, because they were recommended to be vaccinated twice within the same season. In contrast to all other children, these TIIV recipients were taken to be fully vaccinated only when the receipt of a second dose had been more than two weeks ago.

Through 2017–2018, the elderly were vaccinated exclusively with TIIV. In 2018–2019 and 2019–2020, they were instead vaccinated with quadrivalent inactivated influenza vaccine. In the off-season analyses of Publication III, the exposure of interest was vaccination in the season that preceded the off-season in question. As any of the elderly study subjects had either remained unvaccinated throughout that season or been fully vaccinated by the beginning of the off-season, we defined two constant exposure states: unvaccinated and fully vaccinated.

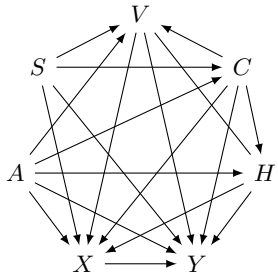

Figure 4.1: Causal diagram showing the presumed relationships between seasonal influenza vaccination (X), influenza infection or disease (Y), age (A), sex (S), history of influenza vaccination (V), presence of underlying chronic conditions (C), and number of hospitalisations in the recent past (H).

4.2.4 Potential Confounders

Demographic and health-related baseline characteristics were assessed at the beginning of each influenza season. The potential confounders considered in the analyses of influenza vaccine effectiveness in children and the elderly were age, sex, history of influenza vaccination, presence of underlying chronic conditions, and number of hospitalisations in the recent past. **Figure 4.1** shows the presumed underlying causal graph drawn based on subject matter knowledge. In Publication II, additional potential confounders specific to cohorts of children, such as gestational age at birth and number of well-baby clinic visits, were considered.

4.2.5 Follow-Up Periods

The follow-up started at the beginning of each study period and ended with the first occurrence of any of the following three events: outcome of interest, loss to follow-up (either due to death or emigration), or end of study period. In Publication II, in which we distinguished between QLAIV and TIIV, vaccination with an influenza vaccine other than the one in question was added to the list of events that ended the follow-up. During the follow-up, each study subject was regarded to be at risk of the respective outcome of interest, *i.e.*, laboratory-confirmed infection with influenza A or influenza B, or hospitalisation for ARI. The reader is referred to Publication I for illustrations of the underlying time-to-event framework.

4.2.6 Effect Measures

Adopting the leaky vaccine failure model (**Section 2.1.3**), the direct effect of influenza vaccination was defined as the hazard ratio, denoted by θ, comparing the hazard of the outcome in question in fully vaccinated study subjects with the corresponding hazard in unvaccinated study subjects. In absence of bias, the hazard ratio of a negative control outcome has the expected value of 1.

Vaccine effectiveness, *i.e.*, the *protective* direct effect of vaccination, was defined as the vaccine-attributable relative reduction in the hazard of the outcome in question. Denoting vaccine effectiveness by η, the relation between the two effect measures is

$$\eta = 1 - \theta. \tag{4.1}$$

4.3 Statistical Models

4.3.1 Cox Proportional Hazards Model

In general, a hazard $\lambda(t)$ is defined as the instantaneous failure rate at time t given survival until t. In the context of this thesis, $\lambda(t)$ describes the time-dependent rate at which the outcome of interest, *e.g.*, laboratory-confirmed influenza infection, occurs.

Denoting the hazards for unvaccinated and fully vaccinated study subjects as functions of time since the start of follow-up by $\lambda_0(t)$ and $\lambda_1(t)$, respectively, a time-dependent hazard ratio is defined as

$$\theta(t) = \frac{\lambda_1(t)}{\lambda_0(t)}. \tag{4.2}$$

Under the assumption that the hazards for unvaccinated and fully vaccinated study subjects are proportional over time, the hazard ratio is constant. The relation between the two hazards is then

$$\lambda_1(t) = \lambda_0(t) \times \theta, \tag{4.3}$$

which is widely known as the Cox proportional hazards model for two groups (Cox, 1972).

One method to check for proportional hazards is to apply the log-log transformation to the survival functions (Kalbfleisch and Prentice, 2002). If the proportionality assumption holds, then

$$\log\left[-\log\left(S_1(t)\right)\right] - \log\left[-\log\left(S_0(t)\right)\right] = \log\left(\theta\right), \tag{4.4}$$

where $S_0(t)$ and $S_1(t)$ denote the survival functions for unvaccinated and fully vaccinated study subjects.

4.3.2 Standard Partial Likelihood

In this section, the Cox proportional hazards model is used to derive the standard partial likelihood. This likelihood allows semiparametric estimation of hazard ratio θ and vaccine effectiveness η.

Let n denote the number of cases in the study cohort, $t_1 < t_2 < \ldots < t_n$ the corresponding ordered event times, and $R(t)$ the set of uncensored study subjects at risk of the outcome at time t. The indicator $x_i(t)$ marks the time-dependent exposure state of subject i and takes value 0 if i is unvaccinated at t and 1 if i is fully vaccinated at t.

The probability of the event occurring to case j at t_j, conditional on the risk set $R(t_j)$, is given by

$$L_j(\theta) = \frac{\lambda_0(t) \times \theta^{x_j(t_j)}}{\sum_{i \in R(t_j)} \lambda_0(t) \times \theta^{x_i(t_j)}} = \frac{\theta^{x_j(t_j)}}{\sum_{i \in R(t_j)} \theta^{x_i(t_j)}}, \tag{4.5}$$

in which the nuisance parameter $\lambda_0(t)$ cancels out. Under the assumption that the event times of any two study subjects are independent, the joint probability of the n ordered cases is the standard partial likelihood

$$L(\theta) = \prod_{j=1}^{n} L_j(\theta). \tag{4.6}$$

This likelihood can be considered as a function of θ and be maximised over θ or $(1 - \theta)$ to produce maximum likelihood estimates for θ or η, respectively.

We applied the above likelihood formulations in Publication II to estimate the vaccine effectiveness in a simple regression analysis without covariates. In Publication III, we used a multiple regression analysis generalising the proportional hazards model in **Equation 4.3** for more than two groups.

4.3.3 Partial Likelihood under Outcome Measurement Errors

In this section, the standard partial likelihood is modified to enable semiparametric estimation of θ and η under outcome measurement errors. The performance of this approach is evaluated in Publication IV using simulated and Finnish register data.

If the outcome measurement is imperfect, the event times are measured with error, including the possibility that some event times are not observed at all. This affects the set of study subjects observed to be at risk ($\tilde{R}(t)$) and potentially biases the observed number of cases, because correct and incorrect event times are indistinguishable by observation.

Figure 4.2 shows the eight possible paths of events for any subject in a cohort study in which the outcome is measured with error. It is assumed that

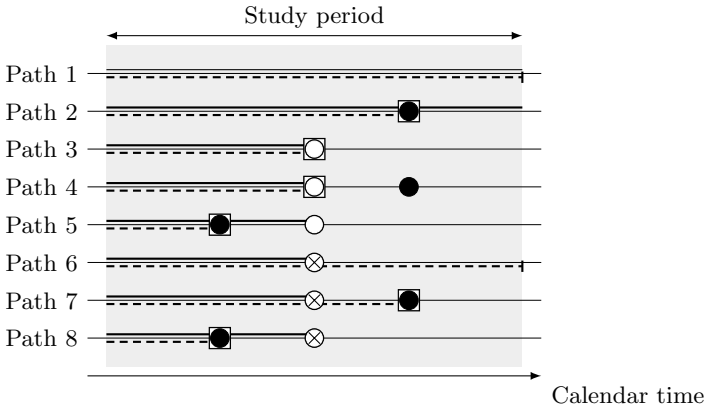

Figure 4.2: Paths of events for any subject in a cohort study in which the outcome is measured with error. The true event is depicted either by a white circle if it was observed or by a crossed circle if it was not observed. False-positive events are depicted by black circles. Although false positives may occur repeatedly, the figure shows only the first of these if any. The subject's true time at risk during the study period is marked by a solid line. In the study, the subject's follow-up (dashed line) ends at the time of the first observed event, which is highlighted by a square around the event-defining circle, or at the time of censoring (vertical bar). Although observed, the true event is not part of the data under study if there is a preceding false-positive event. The true at-risk time is then underestimated (Paths 2, 5 and 8). By contrast, Paths 6 and 7 show scenarios in which the true at-risk time is overestimated. From Publication IV.

the outcome of interest, *i.e.*, the true event, can only occur once per subject and study period. By contrast, other events that are erroneously observed to be of interest, *i.e.*, false positives, may occur repeatedly.

The problem of false-positive events can be addressed probabilistically by right censoring a random subset of the observations given the rate at which false positives occur (see Publication IV), so that only $\tilde{n}$ cases and the corresponding ordered event times $t_1 < t_2 < \ldots < t_{\tilde{n}}$ remain. Assuming the sensitivity of the outcome measurement in the unvaccinated (π_0) and the fully vaccinated (π_1) to be known and constant over time, the likelihood for the event occurring to case j at time t_j is obtained as

$$L_j(\theta) = \frac{\theta^{x_j(t_j)} \times \pi_{x_j(t_j)} \times w_{x_j(t_j)}(t_j)}{\sum_{i \in \tilde{R}(t_j)} \theta^{x_i(t_j)} \times \pi_{x_i(t_j)} \times w_{x_i(t_j)}(t_j)}, \tag{4.7}$$

where the weights $w_0(t)$ and $w_1(t)$ correct the likelihood $L_j(\theta)$ for the bias in the risk set.

The two weights are the ratios of the true survival function ($S_0(t)$ or $S_1(t)$) to the observed survival function ($\tilde{S}_0(t)$ or $\tilde{S}_1(t)$) in the unvaccinated and the fully vaccinated. Because the observed survival function is the complement probability of the true event having occurred and been observed by time t (see

Publication IV), these ratios and thus the weights are equivalent to

$$w_0(t) = \frac{1 - \frac{1-\tilde{S}_0(t)}{\pi_0}}{\tilde{S}_0(t)} \text{ and } w_1(t) = \frac{1 - \frac{1-\tilde{S}_1(t)}{\pi_1}}{\tilde{S}_1(t)}. \tag{4.8}$$

Using the nonparametric estimates of $w_0(t)$ and $w_1(t)$ that follow from the Kaplan-Meier estimator for $\tilde{S}_0(t)$ and $\tilde{S}_1(t)$ (Kaplan and Meier, 1958), the weighted partial likelihood

$$L(\theta) = \prod_{j=1}^{\tilde{n}} L_j(\theta) \tag{4.9}$$

can be maximised in analogy to the standard partial likelihood. The standard error of the resulting maximum likelihood estimate can be obtained using the Fisher information (Fisher and Russell, 1922).

5 Results

5.1 Vaccine Effectiveness in Two-Year-Old Children

5.1.1 Vaccination Coverage and Influenza Incidence

The study cohorts in the influenza seasons 2015–2016, 2016–2017 and 2017–2018 included 60 088, 60 860 and 60 345 children, respectively. By the end of follow-up, 14 %, 20 % and 22 % of the children were fully vaccinated with QLAIV while 7 %, 8 % and 9 % were fully vaccinated with TIIV. There were 309 influenza A and 79 influenza B cases in 2015–2016, 273 influenza A and 9 influenza B cases in 2016–2017, and 268 influenza A and 237 influenza B cases in 2017–2018. Less than 20 % of the cases in each season were hospitalised within one week from the day the influenza-positive specimen was sampled.

5.1.2 Confounders of Influenza Vaccine Effectiveness

To identify which factors may confound the estimation of influenza vaccine effectiveness in children, we investigated in two separate descriptive analyses which baseline characteristics were associated with both the exposure, *i.e.*, seasonal influenza vaccination, and the outcome, *i.e.*, laboratory-confirmed influenza infection. **Table 5.1** presents an extract from these analyses, taking the 2016–2017 cohort as an example. The complete analysis can be found in the Supplementary Data of Publication II.

Each season the subcohort of children vaccinated with QLAIV and the subcohort of children who remained unvaccinated throughout the season were fairly similar in their baseline characteristics. The only major difference was that the percentage of previously vaccinated children in the QLAIV subcohort was higher than the corresponding percentage in the unvaccinated subcohort. However, the baseline risk, approximated by the cumulative risk in the unvaccinated at the end of the season, did not differ between children who were previously vaccinated and children who were not previously vaccinated. We therefore deemed none of the studied baseline characteristics qualified as confounders of QLAIV effectiveness.

Table 5.1: Distribution of the baseline characteristics age (month of birth), sex, history of influenza vaccination, presence of underlying chronic conditions, and number of hospitalisations in the previous year across subcohorts, and the cumulative risk of laboratory-confirmed influenza infection in unvaccinated children in the influenza season 2016–2017.

	Relative frequency, %			Cumulative risk, ‰
Baseline characteristics	QLAIV subcohort N = 11 939	TIIV subcohort N = 5893	Unvaccinated N = 43 028	Unvaccinated N = 60 853
Age				
Nov–Dec 2013	11	20	13	4
Jan–Feb 2014	13	12	14	5
Mar–Apr 2014	13	11	15	6
May–Jun 2014	14	11	16	6
Jul–Aug 2014	16	11	16	5
Sep–Oct 2014	18	13	14	5
Nov–Dec 2014	15	22	12	4
Sex				
Female	49	47	49	5
Male	51	53	51	5
Previously vaccinated				
No	37	24	86	5
Yes	63	76	14	6
Chronically diseased				
No	88	83	89	4
Yes	12	17	11	7
Hospitalisations				
0	93	91	95	5
1	5	6	4	12
≥ 2	1	3	1	7

N, (sub)cohort size. Percentages may not total 100 because of rounding.

Likewise, we did not identify any of the studied baseline characteristics as a strong confounder of TIIV effectiveness. Children vaccinated with TIIV were fairly similar to children who remained unvaccinated, although the TIIV subcohort differed more from the unvaccinated subcohort than the QLAIV subcohort differed from the unvaccinated subcohort.

5.1.3 Estimates of Influenza Vaccine Effectiveness

Figure 5.1 shows the point and 95 % confidence interval (CI) estimates of influenza vaccine effectiveness against laboratory-confirmed influenza infection in children by vaccine, virus type and season. In 2015–2016, a season predominated by influenza A(H1N1) and influenza B/Victoria viruses, vaccine effectiveness against influenza A was estimated at 46 % (QLAIV) and 90 % (TIIV). The corresponding estimates in the subsequent two seasons, charac-

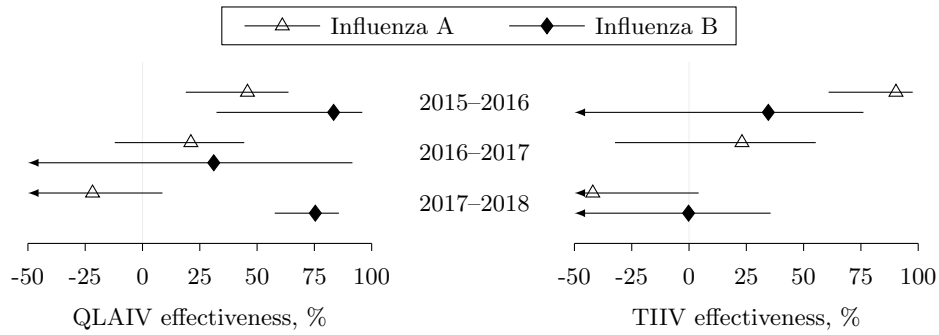

Figure 5.1: Point and 95 % CI estimates of influenza vaccine effectiveness against laboratory-confirmed influenza infection in two-year-old children by vaccine, virus type and season. From Publication II.

terised by the circulation of influenza A(H3N2) and influenza B/Yamagata viruses, were substantially lower. In 2017–2018, neither QLAIV nor TIIV provided protection against influenza A. The estimates of QLAIV effectiveness against influenza B were similar in 2015–2016 (83 %) and 2017–2018 (75 %), demonstrating that QLAIV protected equally well against the two influenza B lineages. The corresponding effectiveness of TIIV, which did not contain antigen of the respective lineage circulating in 2015–2016 and 2017–2018, was estimated at 35 % and 0 %.

5.2 Vaccine Effectiveness in the Elderly

5.2.1 Vaccination Coverage and Influenza Incidence

The study cohorts in the influenza seasons from 2012–2013 through 2019–2020 included each around one million elderly. By the end of follow-up, 37 % to 49 % of the elderly were fully vaccinated. The smallest count of laboratory-confirmed influenza cases was observed in 2012–2013, when 441 cases occurred in the unvaccinated and 185 cases in those fully vaccinated. The highest count was observed in 2017–2018, when 6650 and 5263 cases occurred in the unvaccinated and those fully vaccinated, respectively.

From 2012–2013 through 2017–2018, more than 80 % of the cases were hospitalised within one week from the day the influenza-positive specimen was sampled. The proportions of hospitalised cases were similar among unvaccinated and fully vaccinated cases. Each season the influenza A cases outnumbered the influenza B cases. The share of influenza B was particularly low in 2013–2014, 2016–2017, 2018–2019 and 2019–2020, when less than 10 % of the laboratory-confirmed influenza cases were due to influenza B.

5.2.2 Confounders of Influenza Vaccine Effectiveness

To identify which factors may confound the estimation of influenza vaccine effectiveness in the elderly, we used off-season hospitalisation for ARI as a negative control outcome and compared crude and adjusted hazard ratio estimates with the hazard ratio's expected value of 1. The main results are presented in Table 3 of Publication III.

Through all seven influenza off-seasons, fully vaccinated elderly had a higher cumulative risk of hospitalisation for ARI than unvaccinated elderly. As none of the estimated 95 % CIs for the crude hazard ratio included the value 1, we deemed that the crude estimation was confounded.

Adjusting for altogether 655 different covariate sets (see Publication III), we found multiple sets that lowered the hazard ratio estimates towards the value 1 in all seven off-seasons. One of the sets was relatively simple in terms of model complexity as it only comprised the five covariates age (categorised by year of birth into five age groups), sex, history of influenza vaccination in the previous season, presence of underlying chronic conditions in the previous year, and number of nights hospitalised in the previous five years.

After adjustment for this particular covariate set, the estimated 95 % CIs for the hazard ratio included the values 1 or 1.01 in all seven off-seasons, so that we deemed residual confounding to be absent or negligible. As a result of that, we treated the five covariates as confounders of influenza vaccine effectiveness and included them as covariates in the vaccine effectiveness analysis.

5.2.3 Estimates of Influenza Vaccine Effectiveness

Figure 5.2 shows the point and 95 % CI estimates of influenza vaccine effectiveness against laboratory-confirmed influenza infection in the elderly by virus type and season. The estimates of vaccine effectiveness against influenza A ranged from 9 % in 2017–2018 to 49 % in 2015–2016. In agreement with the findings concerning children (**Section 5.1.3**), vaccine effectiveness levels in the elderly were lower in seasons that were predominated by influenza A(H3N2) viruses. Excluding seasons with very little influenza B activity, the estimates of vaccine effectiveness against influenza B ranged from 23 % in 2017–2018 to 50 % in 2012–2013. Thus, the trivalent influenza vaccines used in those seasons provided moderate protection against influenza B despite the repeat lineage mismatch (**Figure 5.2**).

5.2.4 Estimates of Confounding Bias

Table 5.2 presents the magnitude of confounding bias quantified as the absolute difference between the crude influenza hazard ratio estimate and that

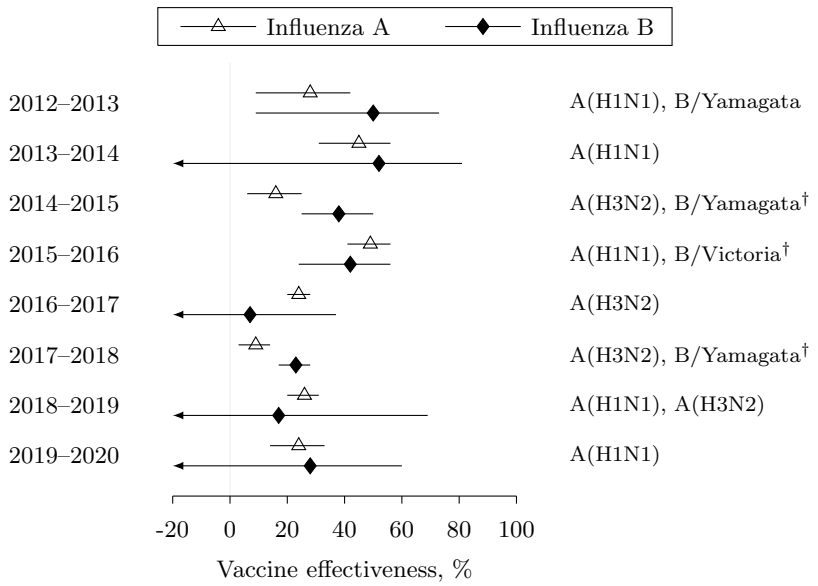

Figure 5.2: Point and 95 % CI estimates of influenza vaccine effectiveness against laboratory-confirmed influenza infection in the elderly by virus type and season adjusted for age, sex, history of influenza vaccination in the previous season, presence of underlying chronic conditions in the previous year, and number of nights hospitalised in the previous five years. The legend on the right-hand side lists the predominant virus subtypes. Subtypes marked by the † symbol were not included in the respective season's trivalent influenza vaccine composition. From Publication III.

Table 5.2: Point estimates of confounding bias in the crude estimation of influenza vaccine effectiveness against laboratory-confirmed influenza infection in the elderly by virus type and season.

	Bias, pp	
Season	Influenza A	Influenza B
2012–2013	− 5	+ 3
2013–2014	− 10	not estimated
2014–2015	− 13	− 1
2015–2016	+ 3	− 4
2016–2017	− 2	not estimated
2017–2018	− 11	+ 4
2018-2019	− 5	not estimated
2019–2020	− 7	not estimated

pp, percentage point.

adjusted for the five confounders identified in **Section 5.2.2**. Without covariate adjustment, we would have underestimated the vaccine effectiveness against influenza A in seven of the eight studied seasons by up to 13 percentage points. In one season, however, confounding introduced a positive bias that would have lead to overestimation. The effect of covariate adjustment on the estimation of vaccine effectiveness against influenza B was roughly within the same order of magnitude.

5.3 Estimates of Outcome Measurement Bias

The simulation study in Publication IV confirmed that the weighted partial likelihood derived in **Section 4.3.3** allows an unbiased estimation of hazard ratios and vaccine effectiveness under outcome measurement errors if the error rates are known. **Table 5.3** summarises the corresponding results obtained for a simulated cohort of one million individuals and 50 % vaccine effectiveness.

The magnitude of the bias in the naïve estimation, which is not corrected for outcome measurement errors, varies depending on the true cumulative risk, the proportion of false-positive events, and the sensitivity in the unvaccinated (π_0) and the fully vacinated (π_1). If the error rates in the unvaccinated and the fully vaccinated are the same, the naïve estimator is biased towards 0 % vaccine effectiveness. However, this bias amounts to only − 4 percentage points when the cumulative risk is small and false-positive events are rare (**Table 5.3**).

In addition, we showed that the adjusted estimation is robust to misspecification of the two sensitivity parameters as long as the true cumulative risk is small and the ratio π_1/π_0 is set correctly. Figure 2D of Publication IV offers a visual representation of this finding.

Table 5.3: Mean estimates of vaccine effectiveness and bias obtained from 10 000 simulated data sets under outcome measurement errors and 50 % vaccine effectiveness. The cumulative risk and proportion of false-positive events to all observed events are given for the unvaccinated at the end of the study period. The naïve estimation was conducted assuming perfect sensitivity and absence of false positives.

Setting				**Adjusted estimation**		**Naïve estimation**	
Cumulative risk, %	False-positive proportion, %	π_0, %	π_1, %	Vaccine effectiveness, %	Bias, pp	Vaccine effectiveness, %	Bias, pp
25	0	4	4	50	± 0	47	$-$ 3
25	0	5	3	50	± 0	68	$+18$
25	2	4	4	50	± 0	46	$-$ 4
25	2	5	3	50	± 0	67	$+17$
25	16	4	4	50	± 0	39	-11
81	0	4	4	50	± 0	30	-20
81	0	5	3	50	± 0	59	$+$ 9

pp, percentage point.

6 Discussion

6.1 Motivation

Following the WHO's recommendation for annual vaccination of individuals at high risk of severe influenza disease, the seasonal vaccination of the elderly population and young children with inactivated influenza vaccines was introduced in the Finnish vaccination programme in 2002 and 2007, respectively (Rapola, 2007). To enhance the vaccine uptake, the vaccination of two-year-old children with a live-attenuated influenza vaccine administered as a nasal spray was added to the programme in 2015 (Nohynek et al., 2016). With the realisation of the National Vaccination Register, real-time surveillance of the programme's implementation was rendered possible in 2012 (Baum et al., 2017). Since then, the vaccine uptake and influenza vaccine effectiveness have been monitored each season in real time to evaluate and improve the Finnish vaccination programme, which in turn benefits public health.

To ensure efficient use of the resources already available, such as the national registers, and those resources that must be invested in each study conduct, influenza vaccine effectiveness in Finland is assessed exclusively based on routinely collected medical and demographic data. Motivated by this policy, this thesis has presented and critically evaluated the register-based cohort study design currently utilised in Finland to estimate influenza vaccine effectiveness in young children and the elderly.

6.2 Main Findings

6.2.1 Register-Based Cohort Study Design

One strength of the applied register-based cohort study design is that it allows for a fast and inexpensive study conduct because it is based on the secondary use of data collected independently of the study. Another strength is the potentially high generalisability of study results, following from the fact that the study cohorts are population-based and thus highly representative. The limitation of the design is that it is difficult, if not impossible, to ensure in-

ternal study validity at the design stage. As demonstrated in this thesis, that challenge may, however, be overcome by reducing or eliminating confounding and measurement biases analytically.

6.2.2 Influenza Vaccine Effectiveness

In Publication II, we measured the vaccine effectiveness of two influenza vaccines (QLAIV and TIIV) given to two-year-old children in the influenza seasons from 2015–2016 through 2017–2018. The estimates were very heterogeneous and ranged from less than 0 % to 90 %. Neither vaccine was clearly better than the other. In Publication III, we measured the vaccine effectiveness of inactivated influenza vaccines in the elderly aged 65 years and older. The results indicate that the vaccine effectiveness did not exceed 50 % in any of the eight seasons under study, *i.e.*, from 2012–2013 through 2019–2020. Due to the small number of influenza B cases in the last two seasons, the benefit of increasing the vaccine valency could not be quantified.

These findings are broadly in line with those from other studies carried out by Buchan et al. (2018), Poehling et al. (2018), Pebody et al. (2016), Pebody et al. (2017), Rondy et al. (2018), Stein et al. (2018), Segaloff et al. (2019), Wang et al. (2018), and Wu et al. (2018), who studied influenza vaccine effectiveness in children in Canada, the United States, the United Kingdom, the European Union, Israel, and China, respectively, and Rondy, Gherasim, et al. (2017), Örtqvist et al. (2018), Rose et al. (2020), and Stuurman et al. (2020), who studied influenza vaccine effectiveness in the elderly in Europe. In addition, a worldwide meta-analysis of influenza vaccine effectiveness in the elderly by Rondy, El Omeiri, et al. (2017) agrees with our results. All but one of the above studies utilised a test-negative design, in which the odds of testing positive for influenza were compared between vaccinated and unvaccinated individuals seeking health care for ARI. The only register-based cohort study outside Finland was conducted in Sweden by Örtqvist et al. (2018).

Comparing single estimates of influenza vaccine effectiveness in order to detect differences between vaccines, seasons or study sites is often complicated. Many of the studies mentioned above as well as the cohort studies in Publication II suffer from lack of statistical precision and report wide CIs due to small numbers of cases. As a consequence, looking at differences between point estimates may not be meaningful. This problem of low precision is further exacerbated by the need for stratification, *e.g.*, by influenza type, or even subtype or lineage. Without such stratification, any differences between the estimates cannot be explained in detail because influenza vaccine effectiveness strongly depends on the match between vaccine and circulating virus strains, which is subtype- and lineage-specific and can vary between vaccines, seasons

and study sites. Furthermore, the severity of the studied outcomes must be taken into account in any comparison of vaccine effectiveness estimates because influenza vaccination reduces the severity of influenza disease (Arriola et al., 2017; Thompson et al., 2018).

The test-negative design typically includes the subtyping of all influenza-positive samples and facilitates an easy distinction between inpatient and outpatient settings and thus disease severity (Sullivan, Feng, and Cowling, 2014; Rondy, El Omeiri, et al., 2017). By contrast, the register-based design presented in this thesis cannot produce subtype- or lineage-specific estimates because clinical routine, the main source of respiratory samples, does not require subtyping. Moreover, the laboratory results collected in the National Infectious Diseases Register alone do not allow any conclusions about the severity of the primary outcome studied in this thesis. As a proxy, the risk of all-cause hospitalisation among the laboratory-confirmed influenza cases was reported. A more elaborate approach of first identifying those cases that were hospitalised specifically because of influenza or its complications, based on ICD-10 diagnostic codes, and then estimating vaccine effectiveness separately for inpatient and outpatient care could have improved the comparability with other studies.

6.2.3 Confounding Bias

In Publication II, we concluded the discussion of confounding bias with the hypothesis that two-year-old children may form quite a homogenous group in Finland, where public health care is free and covering seasonal influenza vaccination of young children. Due to methodological weaknesses, the descriptive analyses performed to identify confounders of vaccine effectiveness in two-year-old children (**Section 5.1.2**) are however a poor support for such a claim. The potential of a single factor to confound the estimation of an exposure-outcome relationship can only be reliably evaluated conditionally on all other confounders of that relationship (Greenland and Morgenstern, 2001). Therefore, a bivariate analysis as presented in this thesis may be insufficient. The same applies to the estimation of the baseline risk using only the unvaccinated. A factor does not necessarily disqualify as a confounder if it appears unassociated with the outcome among the unexposed (Greenland and Morgenstern, 2001). Nevertheless, since Nohynek et al. (2016) and Stuurman et al. (2020) did not find meaningful changes in the Finnish estimates of influenza vaccine effectiveness in young children after controlling for potential confounders, confounding bias may truly be negligible.

In Publication III, by contrast, we made a greater effort to assess the presence of confounding and to adequately control for it. Based on subject matter

knowledge, we identified five potential confounders that could be measured by the available Finnish register data, *i.e.*, age, sex, history of influenza vaccination, presence of underlying chronic conditions, and number of nights hospitalised in the recent past. Due to the novelty of the register-based design, it was however unclear whether controlling for those five would sufficiently eliminate confounding and how some of the potential confounders, such as history of influenza vaccination, should be defined given data that go back several years.

We therefore quantified the association between seasonal influenza vaccination and a negative control outcome, *i.e.*, off-season hospitalisation for ARI. We found evidence of confounding by indication but deemed residual confounding to be absent or negligible after controlling for the aforementioned five potential confounders. In addition, the negative control analysis revealed that it is sufficient to define history of influenza vaccination and presence of underlying chronic conditions using data from the previous year only, while the number of nights hospitalised in the recent past should preferably be defined based on the previous five years. This seems plausible. Individuals vaccinated in the previous season were likely also vaccinated before that. Moreover, many of the considered underlying chronic conditions require annual check-up visits. Thus, most of the previously vaccinated and chronically diseased individuals can be identified using data from the previous year only. By contrast, the number of nights hospitalised, a non-binary measure of health care utilisation and frailty, must be more informative when assessed over a longer though still recent period of time.

We assumed that off-season hospitalisation for ARI is a valid negative control outcome. As the influenza virus circulation ceases in the off-season months, the number of influenza-induced and thus vaccine-preventable ARIs should be negligible. Furthermore, the clinical picture and severity of influenza disease in the elderly are similar to those of the chosen control outcome suggesting that both outcomes share indeed the same common causal determinants apart from the aetiological agent (and vaccination). Jackson et al. (2006) expressed a legitimate concern over using a post-season outcome as a negative control, because differences between the vaccinated and unvaccinated that lead to confounding may diminish over time due to the loss of frail individuals. However, we found no strong indication for such a time-dependent effect and consequently place confidence in our negative control outcome.

Although it was not needed for the estimation of influenza vaccine effectiveness, the magnitude of confounding bias was quantified by taking the absolute difference between the crude and adjusted influenza hazard ratio estimates, disregarding the problem of noncollapsibility (**Section 2.2.2**) since laboratory-confirmed influenza infection can be considered a rare outcome.

Nevertheless, following the earlier discussed lack of precision, the robustness of those estimates, ranging from -13 to $+4$ percentage points, may be low.

6.2.4 Outcome Measurement Bias

The other major threat to the validity of cohort studies analysed in this thesis, apart from confounding bias, is the bias due to outcome measurement errors. The primary outcome in the presented register-based cohort study design is laboratory-confirmed influenza infection and the instrument used to assess the occurrence of that outcome in the study cohorts is the National Infectious Diseases Register, which records all influenza-positive findings in Finland. Hence, it is actually highly unlikely that the primary outcome is measured with error, because the accuracy with which those findings are recorded is by default very high. In a strict sense, however, this outcome definition only allows the estimation of vaccine effectiveness against laboratory confirmation, not against infection as such, which would be the ultimate goal. It must be assumed that many influenza infections are missed by that definition, because not every infected study subject seeks health care and has a respiratory sample taken for laboratory confirmation. The fundamental problem with measuring the occurrence of influenza infection based on register data is thus, first and foremost, a missing data problem. By taking the absence of a record of laboratory-confirmed influenza infection as an indication of the general absence of infection, the missing data problem is transformed into a problem of measurement error.

In Publication IV, we developed a weighted partial likelihood approach to adjust for measurement errors arising from nonsensitive observation of outcomes. Previous solutions to related research questions, such as the work by Meier, Richardson, and Hughes (2003), Tang et al. (2015), and De Smedt et al. (2018), were not applicable in our problem for two reasons. First, Meier, Richardson, and Hughes (2003) focused on chronic outcomes, which if initially missed can still be detected by later measurements. Influenza infection, however, is an acute outcome, which by standard laboratory tests can only be detected up to one week after symptom onset (Carrat et al., 2008). Second, Tang et al. (2015) and De Smedt et al. (2018) did not consider time-to-event outcomes in their analyses of biases caused by measurement errors and thus did not discuss the impact that erroneous event times have on the risk set.

To our surprise, we found that the overestimation of the risk set size due to imperfect sensitivity of the outcome measurement has a relatively small effect on the estimation of vaccine effectiveness against influenza infection. When the cumulative risk of the outcome is small, such as the risk of influenza infection, nondifferentially imperfect sensitivity biases the estimation only slightly.

However, when the cumulative risk increases, so does the bias, which on average leads to underestimation of vaccine effectiveness if the vaccine has a protective direct effect. In addition, we showed that it is mainly the ratio of the two sensitivity parameters that determines the magnitude of bias. This finding is in line with the results obtained by De Smedt et al. (2018).

Using the weighted partial likelihood (**Section 4.3.3**) and probabilistic deletion of false positives, the estimation of hazard ratios and vaccine effectiveness from erroneous time-to-event data can be corrected for outcome measurement errors when the error rates are known. In absence of other sources of bias, the adjusted estimates are unbiased. Furthermore, the method is robust to misspecification of the two sensitivity parameters as long as their ratio is set correctly and the cumulative risk is small. Unfortunately, the error rates for measuring influenza infection based on Finnish register data are not known. Shubin et al. (2014) estimated the overall sensitivity of the National Infectious Diseases Register at 4 % during the 2009–2010 influenza pandemic and detected differences related to age, region and time. It may thus be possible that the sensitivity also varies by vaccination status, *e.g.*, if physicians are more likely to take a respiratory sample from unvaccinated patients for laboratory confirmation. However, the lack of influenza-negative test results, which are not part of the register's data content, hinders the conduct of a simple validation study.

6.3 Strengths and Limitations

The strengths of the research presented in this thesis lie in the combination of theoretical and practical analyses using models, such as the leaky vaccine failure model and the causal graph in **Figure 4.1**, real world data retrieved from Finnish registers and simulated data. To compare register-based cohort study results with those from other studies, influenza vaccine effectiveness was estimated in two distinct age groups and eight very recent influenza seasons.

Moreover, instead of readily adopting common confounder selections from the literature, potential confounders were carefully selected based on subject matter knowledge, data availability and a negative control outcome analysis. In addition, the scope of existing bias considerations was broadened by taking into account that in cohort studies outcome measurement errors do not only affect the event status but also the event time.

Nonetheless, the research summarised in this thesis is also subject to several limitations. The major limitation is that the presence and magnitude of confounding bias were studied under the assumption that all other sources of bias are absent. This however is unlikely to be the case since measurement errors can never be ruled out, especially in studies that are based on

the secondary use of data. If measurement bias interferred with the negative control outcome analysis, *e.g.*, by masking the effect of residual confounding, the performed covariate adjustment might not present an adequate control of confounding.

Another limitation is that a thorough consideration of bias due to exposure measurement errors has been omitted. The National Vaccination Register's coverage of vaccinations given in the public health centres is generally very high (Baum et al., 2017). Unfortunately, seasonal influenza vaccinations may also be administered outside those centres, *e.g.*, in the private sector or in nursing homes. To date, the extent of the resulting undercoverage is unknown.

6.4 Future Perspectives

To validate the register-based cohort study design currently utilised in the yearly estimation of influenza vaccine effectiveness in Finland, future analyses should quantify the sensitivity and specificity of the Finnish register data. In addition, factors that may affect the registers' accuracy should be identified and taken into account analytically.

The accuracy of the National Vaccination Register could be estimated in surveys comparing manually validated records of seasonal influenza vaccination in representative samples of the study cohorts with those records in the register. Prior to the realisation of the National Vaccination Register, the vaccination coverage in young children was measured regularly based on surveys (Leino et al., 2007).

The optimal strategy to estimate how well influenza infections are covered by the Finnish register data would be to closely follow a representative sample of the study cohorts through an influenza season and continuously confirm each individual's infection status. Alternatively, making influenza-negative test results routinely available would provide a great opportunity for the conduct of a validation study that is, like the analyses in this thesis, solely based on the secondary use of data. Differences between the general population and the population that undergoes laboratory testing could be examined. Moreover, nested case-control studies could be performed utilising the test-negative design. This would enable a direct comparison of estimates obtained from different study designs as demonstrated by Castilla et al. (2012).

Apart from Finland, a few other countries, *e.g.*, Denmark (Thygesen et al., 2011; Grove Krause et al., 2012), have computerised national registers that facilitate the population-based study of influenza vaccinations and influenza disease. However, legal implications of deterministic data linkage based on personal identity codes might hinder the future implementation of the register-based cohort study design elsewhere.

Although this thesis focused exclusively on the estimation of influenza vaccine effectiveness in young children and the elderly, the applicability of the register-based cohort study design in Finland is much wider. The design could, *e.g.*, be utilised to also estimate influenza vaccine effectiveness in health care workers. Moreover, the design has already been used in studies assessing the safety of influenza and human papilloma virus vaccines (Baum et al., 2015; Skufca et al., 2018). Whether it is vaccine effectiveness or vaccine safety, the design can be adopted to analyse the effects of many different vaccines.

7 Conclusions

The objectives of this thesis were to develop methods enabling the efficient use of Finnish register data in the estimation of influenza vaccine effectiveness and to critically evaluate the methods' limitations. Based on the results of this thesis, the following conclusions can be drawn:

1. The secondary use of the medical and demographic data in the Finnish registers allows for a fast and inexpensive study conduct, suited for yearly estimation of influenza vaccine effectiveness.

2. It is difficult to ensure the internal validity of a register-based study, because the quality of the study data cannot be directly influenced. Validation data may be required.

3. The current data content of the Finnish registers enables a register-based conduct of cohort studies of influenza vaccine effectiveness but does not permit a register-based conduct of case-control studies.

4. A small set of only five confounders has been found sufficient to control for confounding in register-based cohort studies of influenza vaccine effectiveness in the Finnish elderly.

5. In cohort studies of influenza vaccine effectiveness, outcome measurement errors can be corrected analytically if the error rates are known. Influenza can be considered a rare disease.

6. Seasonal influenza vaccination does not provide perfect protection against influenza infection, neither in two-year-old children nor in the elderly aged 65 years and older.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisors Professor Kari Auranen and Professor Sangita Kulathinal for their dedicated support and the stimulating discussions. With incredible patience, they guided me in scientific writing, taught me statistics and helped me realise this doctoral thesis.

I also wish to thank the pre-examiners Professor Nick Andrews and Professor Marjo Renko, who devoted their time to reviewing this thesis during a worldwide emergency. In addition, I am grateful to Associate Professor Tyra Grove Krause, who accepted the invitation to be the opponent in the public examination of this thesis.

Thanks should also go to Doctor Hanna Nohynek, Adjunct Professor Jukka Jokinen, Professor Terho Heikkinen and my colleagues at the Finnish Institute for Health and Welfare. I appreciate that they encouraged me to aim for a doctorate and I am grateful for their expertise, insightful advice, and motivating Teams calls, which kept me writing during self-quarantine. I truly cherish the time we spent together in the office and the cafeteria as well as our board game and ping-pong sessions.

I also had great pleasure meeting Professor Elizabeth Halloran at the Fred Hutchinson Cancer Research Center and thank her for making this research visit possible. In addition, I would like to extend my gratitude to the Doctoral Programme in Clinical Research for financially supporting the said research visit and the writing of this thesis.

Finally, I want to warmly thank my family and Mattia for their unconditional love, and my dear friends in Helsinki, Germany and overseas for making the good times in life even better and the hard times a whole lot easier.

August 2021
Ulrike Baum

References

Arriola, C., Garg, S., Ryan, P., George, A., Zansky, S., Bennett, N., et al. (2017). Influenza Vaccination Modifies Disease Severity Among Community-dwelling Adults Hospitalized With Influenza. Clinical Infectious Diseases 65 (8), 1289–1297.

Baum, U., Sundman, J., Jääskeläinen, S., Nohynek, H., Puumalainen, T., and Jokinen, J. (2017). Establishing and maintaining the National Vaccination Register in Finland. Eurosurveillance 22 (17), 30520.

Baum, U., Leino, T., Gissler, M., Kilpi, T., and Jokinen, J. (2015). Perinatal survival and health after maternal influenza A(H1N1)pdm09 vaccination: A cohort study of pregnancies stratified by trimester of vaccination. Vaccine 33 (38), 4850–4857.

Biggerstaff, M., Cauchemez, S., Reed, C., Gambhir, M., and Finelli, L. (2014). Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: A systematic review of the literature. BMC Infectious Diseases 14, 480.

Bresee, J. S., Fry, A. M., Sambhara, S., and Cox, N. J. (2018). Inactivated Influenza Vaccines. In: Plotkin's Vaccines. Seventh Edition. Philadelphia, Pennsylvania: Elsevier, Inc.

Buchan, S. A., Booth, S., Scott, A. N., Simmonds, K. A., Svenson, L. W., Drews, S. J., et al. (2018). Effectiveness of Live Attenuated vs Inactivated Influenza Vaccines in Children During the 2012–2013 Through 2015–2016 Influenza Seasons in Alberta, Canada: A Canadian Immunization Research Network (CIRN) Study. JAMA Pediatrics 172 (9), e181514.

Carrat, F., Vergu, E., Ferguson, N. M., Lemaitre, M., Cauchemez, S., Leach, S., et al. (2008). Time lines of infection and disease in human influenza: a review of volunteer challenge studies. American Journal of Epidemiology 167 (7), 775–785.

Castilla, J., Martínez-Artola, V., Salcedo, E., Martínez-Baz, I., García Cenoz, M., Guevara, M., et al. (2012). Vaccine effectiveness in preventing influenza hospitalizations in Navarre, Spain, 2010–2011: Cohort and case-control study. Vaccine 30 (2), 195–200.

Clayton, D. and Hills, M. (1993). Statistical Models in Epidemiology. Oxford, England: Oxford University Press.

Cox, D. R. (1972). Regression Models and Life-Tables. Journal of the Royal Statistical Society, Series B 34 (2), 187–220.

Crowcroft, N. S. and Klein, N. P. (2018). A framework for research on vaccine effectiveness. Vaccine 36 (48), 7286–7293.

De Smedt, T., Merrall, E., Macina, D., Perez-Vilar, S., Andrews, N., and Bollaerts, K. (2018). Bias due to differential and non-differential disease- and exposure misclassification in studies of vaccine effectiveness. PloS One 13 (6), e0199180.

Fisher, R. A. and Russell, E. J. (1922). On the Mathematical Foundations of Theoretical Statistics. Philosophical Transactions of the Royal Society of London, Series A 222 (594-604), 309–368.

Glymour, M. M. and Greenland, S. (2008). Causal Diagrams. In: Modern Epidemiology. Third Edition. Philadelphia, Pennsylvania: Lippincott Williams and Wilkins.

Greenland, S. and Morgenstern, H. (2001). Confounding in Health Research. Annual Review of Public Health 22 (1), 189–212.

Greenland, S., Pearl, J., and Robins, J. M. (1999). Causal Diagrams for Epidemiologic Research. Epidemiology 10 (1), 37–48.

Greenland, S., Robins, J. M., and Pearl, J. (1999). Confounding and Collapsibility in Causal Inference. Statistical Science 14 (1), 29–46.

Greenwood, M. and Yule, G. U. (1915). The Statistics of Anti-Typhoid and Anti-Cholera Inoculations, and the Interpretation of Such Statistics in General. Proceedings of the Royal Society of Medicine 8, 113–194.

Grove Krause, T., Jakobsen, S., Haarh, M., and Mølbak, K. (2012). The Danish vaccination register. Eurosurveillance 17 (17), 20155.

Haber, M., Longini, I. M., and Halloran, M. E. (1991). Measures of the Effects of Vaccination in a Randomly Mixing Population. International Journal of Epidemiology 20 (1), 300–310.

Halloran, M. E., Longini, I. M., and Struchiner, C. J. (1996). Estimability and Interpretation of Vaccine Efficacy Using Frailty Mixing Models. American Journal of Epidemiology 144 (1), 83–97.

Halloran, M. E. and Struchiner, C. J. (1991). Study Designs for Dependent Happenings. Epidemiology 2 (5), 331–338.

Hanquet, G., Valenciano, M., Simondon, F., and Moren, A. (2013). Vaccine effects and impact of vaccination programmes in post-licensure studies. Vaccine 31 (48), 5634–5642.

Hernán, M. A., Hernández-Díaz, S., and Robins, J. M. (2004). A Structural Approach to Selection Bias. Epidemiology 15 (3), 615–625.

Hernán, M. A., Hernández-Díaz, S., Werler, M. M., and Mitchell, A. A. (2002). Causal Knowledge as a Prerequisite for Confounding Evaluation: An Application to Birth Defects Epidemiology. American Journal of Epidemiology 155 (2), 176–184.

Hill, H. A. and Kleinbaum, D. G. (2005). Bias in Observational Studies. In: Encyclopedia of Biostatistics. Second Edition. Chichester, England: John Wiley and Sons, Ltd.

Jackson, L. A., Jackson, M. L., Nelson, J. C., Neuzil, K. M., and Weiss, N. S. (2006). Evidence of bias in estimates of influenza vaccine effectiveness in seniors. International Journal of Epidemiology 35 (2), 337–344.

Janes, H., Dominici, F., and Zeger, S. (2010). On quantifying the magnitude of confounding. Biostatistics 11 (3), 572–582.

Kalbfleisch, J. D. and Prentice, R. L. (2002). The Statistical Analysis of Failure Time Data. Second Edition. Hoboken, New Jersey: John Wiley and Sons, Inc.

Kaplan, E. L. and Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association 53 (282), 457–481.

Leino, T., Koskenniemi, E., Saranpää, P.-R., Strömberg, N., and Kilpi, T. (2007). Rokotuskattavuus edelleen huippuluokkaa. Lääkärilehti 8 (62), 739–743.

Lipsitch, M., Tchetgen Tchetgen, E., and Cohen, T. (2010). Negative Controls A Tool for Detecting Confounding and Bias in Observational Studies. Epidemiology 21 (3), 383–388.

Luke, C. J., Lakdawala, S. S., and Subbarao, K. (2018). Influenza Vaccine–Live. In: Plotkin's Vaccines. Seventh Edition. Philadelphia, Pennsylvania: Elsevier, Inc.

Meier, A. S., Richardson, B. A., and Hughes, J. P. (2003). Discrete Proportional Hazards Models for Mismeasured Outcomes. Biometrics 59 (4), 947–954.

Mereckiene, J., Cotter, S., Nicoll, A., Lopalco, P., Noori, T., Weber, J. T., et al. (2014). Seasonal influenza immunisation in Europe. Overview of recommendations and vaccination coverage for three seasons: Pre-pandemic (2008/09), pandemic (2009/10) and post-pandemic (2010/11). Eurosurveillance 19 (16), 20780.

Nohynek, H., Baum, U., Syrjänen, R., Ikonen, N., Sundman, J., and Jokinen, J. (2016). Effectiveness of the live attenuated and the inactivated influenza vaccine in two-year-olds - A nationwide cohort study Finland, influenza season 2015/16. Eurosurveillance 21 (38), 30346.

Örtqvist, A., Brytting, M., Leval, A., and Hergens, M. P. (2018). Impact of repeated influenza vaccinations in persons over 65 years of age: A large population-based cohort study of severe influenza over six consecutive seasons, 2011/12–2016/17. Vaccine 36 (37), 5556–5564.

Paules, C. and Subbarao, K. (2017). Influenza. The Lancet 390, 697–708.

Pebody, R., Sile, B., Warburton, F., Sinnathamby, M., Tsang, C., Zhao, H., et al. (2017). Live attenuated influenza vaccine effectiveness against hospitalisation due to laboratory-confirmed influenza in children two to six years of age in England in the 2015/16 season. Eurosurveillance 22 (4), 30450.

Pebody, R., Warburton, F., Ellis, J., Andrews, N., Potts, A., Cottrell, S., et al. (2016). Effectiveness of seasonal influenza vaccine for adults and children in preventing laboratory-confirmed influenza in primary care in the United Kingdom: 2015/16 end-of-season results. Eurosurveillance 21 (38), 30348.

Poehling, K. A., Caspard, H., Peters, T. R., Belongia, E. A., Congeni, B., Gaglani, M., et al. (2018). 2015–2016 Vaccine Effectiveness of Live Attenuated and Inactivated Influenza Vaccines in Children in the United States. Clinical Infectious Diseases 66 (5), 665–672.

Porta, M. (2008). A Dictionary of Epidemiology. Fifth Edition. New York, New York: Oxford University Press, Inc.

Rapola, S. (2007). National immunization program in Finland. International Journal of Circumpolar Health 66 (5), 382–389.

Remschmidt, C., Wichmann, O., and Harder, T. (2015). Frequency and impact of confounding by indication and healthy vaccinee bias in observational studies assessing influenza vaccine effectiveness: a systematic review. BMC Infectious Diseases 15, 429.

Rondy, M., El Omeiri, N., Thompson, M. G., Leveque, A., Moren, A., and Sullivan, S. G. (2017). Effectiveness of influenza vaccines in preventing severe influenza illness among adults: A systematic review and meta-analysis of test-negative design case-control studies. Journal of Infection 75 (5), 381–394.

Rondy, M., Gherasim, A., Casado, I., Launay, O., Rizzo, C., Pitigoi, D., et al. (2017). Low 2016/17 season vaccine effectiveness against hospitalised influenza A(H3N2) among elderly: awareness warranted for 2017/18 season. Eurosurveillance 22 (41), 17–00645.

Rondy, M., Kissling, E., Emborg, H. D., Gherasim, A., Pebody, R., Trebbien, R., et al. (2018). Interim 2017/18 influenza seasonal vaccine effectiveness: combined results from five European studies. Eurosurveillance 23 (9), 18–00086.

Rose, A. M. C., Kissling, E., Gherasim, A., Casado, I., Bella, A., Launay, O., et al. (2020). Vaccine effectiveness against influenza A(H3N2) and B among laboratory-confirmed, hospitalised older adults, Europe, 2017–18: A season of B lineage mismatched to the trivalent vaccine. Influenza and Other Respiratory Viruses 14 (3), 302–310.

Rothman, K. J. and Greenland, S. (2008). Cohort Studies. In: Modern Epidemiology. Third Edition. Philadelphia, Pennsylvania: Lippincott Williams and Wilkins.

Rothman, K. J., Greenland, S., and Lash, T. L. (2008). Validity in Epidemiologic Studies. In: Modern Epidemiology. Third Edition. Philadelphia, Pennsylvania: Lippincott Williams and Wilkins.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. Journal of Educational Psychology 66 (5), 688–701.

Segaloff, H. E., Leventer-Roberts, M., Riesel, D., Malosh, R. E., Feldman, B. S., Shemer-Avni, Y., et al. (2019). Influenza Vaccine Effectiveness Against Hospitalization in Fully and Partially Vaccinated Children in Israel; 2015–16, 2016–17, and 2017–18. Clinical Infectious Diseases 69 (12), 2153–2161.

Shubin, M., Virtanen, M., Toikkanen, S., Lyytikäinen, O., and Auranen, K. (2014). Estimating the burden of A(H1N1)pdm09 influenza in Finland during two seasons. Epidemiology and Infection 142 (5), 964–974.

Skufca, J., Ollgren, J., Artama, M., Ruokokoski, E., Nohynek, H., and Palmu, A. (2018). The association of adverse events with bivalent human papilloma virus vaccination: A nationwide register-based cohort study in Finland. Vaccine 36 (39), 5926–5933.

Smith, P. G., Rodrigues, L. C., and Fine, P. E. (1984). Assessment of the Protective Efficacy of Vaccines against Common Diseases Using Case-Control and Cohort Studies. International Journal of Epidemiology 13 (1), 87–93.

Somes, M., Turner, R., Dwyer, L., and Newall, A. (2018). Estimating the annual attack rate of seasonal influenza among unvaccinated individuals: A systematic review and meta-analysis. Vaccine 36 (23), 3199–3207.

Stein, Y., Mandelboim, M., Sefty, H., Pando, R., Mendelson, E., Shohat, T., et al. (2018). Seasonal Influenza Vaccine Effectiveness in Preventing Laboratory-Confirmed Influenza in Primary Care in Israel, 2016–2017 Season: Insights Into Novel Age-Specific Analysis. Clinical Infectious Diseases 66 (9), 1383–1391.

Stuurman, A. L., Bollaerts, K., Alexandridou, M., Biccler, J., Díez Domingo, J., Nohynek, H., et al. (2020). Vaccine effectiveness against laboratory-confirmed influenza in Europe - Results from the DRIVE network during season 2018/19. Vaccine 38 (41), 6455–6463.

Sullivan, S. G., Feng, S., and Cowling, B. J. (2014). Potential of the test-negative design for measuring influenza vaccine effectiveness: a systematic review. Expert Review of Vaccines 13 (12), 1571–1591.

Tang, L., Lyles, R. H., King, C. C., Celentano, D. D., and Lo, Y. (2015). Binary Regression with Differentially misclassified Response and Exposure Variables. Statistics in Medicine 34 (9), 1605–1620.

Thompson, M., Pierse, N., Huang, S., Prasad, N., Duque, J., Newbern, E., et al. (2018). Influenza vaccine effectiveness in preventing influenza-associated intensive care admissions and attenuating severe disease among adults in New Zealand 2012–2015. Vaccine 36 (39), 5916–5925.

Thygesen, L. C., Daasnes, C., Thaulow, I., and Brønnum-Hansen, H. (2011). Introduction to Danish (nationwide) registers on health and social issues: Structure, access, legislation, and archiving. Scandinavian Journal of Public Health 39 (7), 12–16.

Tokars, J. I., Patel, M. M., Foppa, I. M., Reed, C., Fry, A. M., and Ferdinands, J. M. (2020). Waning of Measured Influenza Vaccine Effectiveness Over Time: The Potential Contribution of Leaky Vaccine Effect. Clinical Infectious Diseases 71 (10), e633–e641.

VanderWeele, T. J. (2019). Principles of confounder selection. European Journal of Epidemiology 34, 211–219.

VanderWeele, T. J. and Hernán, M. A. (2012). Results on Differential and Dependent Measurement Error of the Exposure and the Outcome Using Signed Directed Acyclic Graphs. American Journal of Epidemiology 175 (12), 1303–1310.

VanderWeele, T. J. and Shpitser, I. (2013). On the definition of a confounder. Annals of Statistics 41 (1), 196–220.

Wang, Y., Chen, L., Yu, J., Pang, Y., Zhang, J., Zhang, T., et al. (2018). The effectiveness of influenza vaccination among nursery school children in China during the 2016/17 influenza season. Vaccine 36 (18), 2456–2461.

Wu, S., Pan, Y., Zhang, X., Zhang, L., Duan, W., Ma, C., et al. (2018). Influenza vaccine effectiveness in preventing laboratory-confirmed influenza in outpatient settings: A test-negative case-control study in Beijing, China, 2016/17 season. Vaccine 36 (38), 5774–5780.

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama, Turku, Finland 2021

ISBN 978-951-29-8576-0 (PRINT)
ISBN 978-951-29-8577-7 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online))