Department of Public Health Solutions
Finnish Institute for Health and Welfare
and
Department of Public Health
University of Helsinki

EFFECT MEASURES, THEIR ESTIMATION AND INTERPRETATION

APPLICATIONS TO PNEUMOCOCCAL CONJUGATE VACCINATION

Hanna Rinta-Kokko

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Medicine of the University of Helsinki, for public examination in lecture room U3032, University main building, on 14.01.2022, at 12 noon.

Helsinki 2022

Doctoral Programme in Population Health

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

Reviewers
Professor Allen Cheng
School of Public Health and Preventive Medicine
Monash University, Melbourne, Australia

Professor Esa Läärä
Research Unit of Mathematical Sciences
University of Oulu, Oulu, Finland

Opponent
Senior researcher, Assistant Professor Albert Jan van Hoek
National Institute for Public Health and Environment, the Netherlands

Custos
Professor Juha Pekkanen
Department of Public Health
University of Helsinki, Helsinki, Finland

The Faculty of Medicine uses the Urkund system (plagiarism recognition) to examine all doctoral dissertations.

ISBN 978-951-51-7690-5 (pbk.)
ISBN 978-951-51-7691-2 (PDF)
https://ethesis.helsinki.fi

Unigrafia
Helsinki 2022

ABSTRACT

The direct and indirect effects of pneumococcal vaccination on an individual and the population are of great interest. This study focuses on the definition, estimation and interpretation of different effect measures of vaccines and vaccination against pneumococcal colonisation and disease. Vaccine efficacy, effectiveness and impact are considered as epidemiological parameters of interest which need to be estimated using observations gathered according to some study design.

In this thesis, vaccine efficacy against colonisation is defined through pneumococcal acquisition, which describes the natural process of incident occurrences of colonisation better than prevalence. Moreover, a general definition of vaccine efficacy against a multi-type pathogen is presented, with an epidemiologically meaningful interpretation as a weighted average of strain-specific efficacies. A feasible estimation method is then proposed, based on cross-sectional measurement on the current status of colonisation. It is shown that the new efficacy parameter can be estimated using an odds-ratio-based estimator by controlling for the differential time at-risk for pneumococcal acquisition. When the differences in times at-risk between vaccinated and unvaccinated individuals are taken into account, the estimation of vaccine efficacy against colonisation is shown to be less biased by within-host competition between different serotypes (strains). The estimation method is exemplified with empirical data of pneumococcal colonisation in Israeli children.

At the population level, vaccine effectiveness is the measure of vaccine-induced protection during an ongoing vaccination programme when both vaccinated and unvaccinated individuals experience the indirect effects of the vaccination programme. Vaccine impact is the population prevented fraction of the incidence of infection when exposure is the vaccination programme rather than each individual's own vaccination. Both vaccine effectiveness and impact are parameters that depend on the population dynamics of pneumococcal colonisation and disease after vaccine introduction. In this thesis, the time trends of vaccine effectiveness and impact are described with a pseudo-dynamic model that incorporates the incidences of pneumococcal carriage and disease. The model shows that the effectiveness and impact against vaccine-serotype invasive pneumococcal disease (IPD) are expected to be high and largely of the same magnitude through the post-introduction period. By contrast, the vaccine effectiveness and impact against non-vaccine-serotype IPD follow very divergent paths while the vaccine-type colonisation and disease become eliminated.

The practical estimation of vaccine effectiveness is exemplified with register data of Finnish children eligible for pneumococcal conjugate (PCV10) vaccination. Three parallel study designs, the cohort, nested case-control and indirect cohort designs, are shown to provide estimates that are broadly concordant with each other. There has been a sustained high effectiveness against invasive pneumococcal disease (IPD) caused by the PCV10 serotypes. The case numbers of non-PCV10-related IPD are small, but the time averages of the parameter estimates agreed well the expected values based on the pseudo-dynamic model.

The parameters of vaccine efficacy as proposed in this thesis can be interpreted as measures of the biological effect of the vaccine on new vaccine-type acquisitions and should therefore allow more robust comparisons across different epidemiological settings with differing levels of exposure by non-vaccine strains. Moreover, the thesis helps to interpret the time-varying parameters of vaccine impact and effectiveness during large-scale vaccinations, and their manifestation in Finnish children.

TIIVISTELMÄ

Pneumokokkirokotusten yksilöön ja koko väestöön kohdistuvat suorat ja epäsuorat vaikutukset on tärkeää tuntea. Tämä tutkimus keskittyy pneumokokkirokotteiden tehomittojen määritelmiin, estimointiin ja tulkintaan. Rokotteen teho ennen rokotusohjelman aloittamista sekä teho ja vaikuttavuus ohjelman aikana ovat kiinnostavia parametreja, jotka estimoidaan keräämällä havaintoja jonkin koeasetelman mukaisesti.

Tässä väitöstyössä tarkastellaan pneumokokkirokotteen tehoa nenänielukantajuutta vastaan kantajuuden ilmaantuvuuden kautta. Ilmaantuvuus kuvaa kantajuuden biologista luonnetta paremmin kuin sen esiintyvyys, mutta vaatii tyypillisesti pitkittäismittauksia. Työssä osoitetaan, että rokotusteho kantajuuden ilmaantuvuutta vastaan voidaan estimoida poikkileikkausaineistosta odds-suhteena. Lisäksi näytetään, että kun rokotusteho määritellään patogeenille, jolla on monta alatyyppiä kuten pneumokokille, on huomioitava eri alatyyppien keskinäinen kilpailu nenänielussa. Kilpailusta seuraa, että aika jonka rokotetut ja rokottamattomat yksilöt viettävät alttiina rokotetyypin kantajuudelle on erilainen. Kun erilaiset alttiusajat otetaan huomioon, rokotustehon estimaatti vastaa tarkemmin todellista rokotustehoa. Tätä havainnollistetaan israelilaisten päiväkotilasten kantajuusmittausten avulla.

Laajamittaisen rokotusohjelman aikana rokotusteho mittaa rokotteen yksilölle tarjoamaa suoraa suojaa tilanteessa, jossa sekä rokotetut että rokottamattomat lapset kokevat myös epäsuoria vaikutuksia (laumasuojaa ja ei-rokotetyyppien korvautumista). Rokotusohjelman vaikuttavuus mittaa kantajuuden tai taudin ilmaantuvuuden muutosta verrattuna tilanteeseen ennen rokotusohjelmaa. Sekä rokotusteho että vaikuttavuus ovat parametreja, jotka riippuvat pneumokkikantajuuden ja -taudin väestödynamiikasta. Tässä väitöstyössä rokotustehon ja vaikuttavuuden aikatrendejä kuvataan pseudodynaamisella mallilla, joka ottaa huomioon kantajuuden ja taudin ilmaantuvuuden muutokset ajassa. Mallin mukaan sekä rokotusteho että vaikuttavuus rokotetyypin vakavaa pneumokokkitautia vastaan pysyvät korkeina ja liki samansuuruisina koko rokotusohjelman ajan. Sitä vastoin rokotusteho ja vaikuttavuus ei-rokotetyypin vakavaa pneumokokkitautia vastaan ovat hyvin erisuuruiset silloin, kun rokotetyypin kantajuus on vähentynyt ja poistumassa väestöstä.

Rokotustehon estimointia havainnollistetaan käyttäen suomalaista terveysrekisteriaineistoa vakavan pneumokokkitaudin tapauksista lapsilla, jotka ovat oikeutettuja pneumokokkirokotusohjelmaan. Kolmen tutkimusasetelman eli kohortti-, pesäytetyn tapaus-verrokki- ja epäsuoran

kohorttiasetelman näytetään tarjoavan likimain samansuuruisia estimaatteja. Rokotusteho kaikkia kymmentä rokotealatyyppiä vastaan on pysynyt korkeana koko rokotusohjelman ajan. Pseudodynaamisen mallin antamat odotetut rokotustehon arvot myös ei-rokotetyypeille vastaavat hyvin toteutuneita aikakeskiarvoja, vaikka tautitapauksia on ollut vähän.

Tässä väitöstyössä esitetyt rokotustehon parametrit nenänielukantajuutta vastaan tarjoavat mahdollisuuden verrata rokotetutkimuksia erilaisissa asetelmissa, vaikka ei-rokotetyypin kantajuuden ilmaantuvuus voi vaihdella paljonkin. Lisäksi tutkimus tarjoaa keinoja tulkita ajassa muuttuvia rokotustehon ja vaikuttavuuden mittoja laajojen rokotusohjelmien aikana, erityisesti suomalaisten lasten näkökulmasta.

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I. Rinta-Kokko H, Dagan R, Givon-Lavi N, Auranen K. Estimation of vaccine efficacy against acquisition of pneumococcal carriage. Vaccine 2009;27: 3831–3937.

II. Auranen K, Rinta-Kokko H, Halloran ME. Estimating strain-specific and overall efficacy of polyvalent vaccines against recurrent pathogens from a cross-sectional study. Biometrics 2013; 69(1): 235–244.

III. Rinta-Kokko H, Auranen K, Toropainen M, Nuorti JP, Nohynek H, Siira L, Palmu AA. Effectiveness of 10-valent pneumococcal conjugate vaccine estimated with three parallel study designs among vaccine-eligible children in Finland. Vaccine 2020;38(6):1559-1564.

IV. Rinta-Kokko H, Nurhonen M, Auranen K. Impact and effectiveness of a conjugate vaccine against invasive pneumococcal disease in Finland – a modelling approach. Hum Vaccin Immunother 2021;17(6):1834-1843.

In the text the publications are referred as Studies and enumerated by their Roman numerals.

ABBREVIATIONS

AOM	Acute otitis media
ARU/ARV	Attack rate in unvaccinated/vaccinated individuals (incidence proportion is recommended instead)
CAP	Community-acquired pneumonia
CAPiTA	the Community-Acquired Pneumonia Immunization Trial in Adults
CI	Confidence interval, credible interval
COMPAS	the Clinical Otitis Media & Pneumonia Study
ECDC	European Centre for Disease Prevention and Control
FinIP	the Finnish Invasive Pneumococcal disease vaccine trial
FinOM	the Finnish Otitis Media trial
GEE	Generalized estimating equations
HIV	Human immunodeficiency virus
ICD-10	the 10th revision of the International Statistical Classification of Diseases and Related Health Problems
IPD	Invasive pneumococcal disease
IRD	Incidence rate difference
ITS	Interrupted time series
PCR	Polymerase chain reaction
PCV	Pneumococcal conjugate vaccine
NNV	Number needed to vaccinate
NVP	National vaccination programme
NVT	Non-vaccine type
PCV7	the 7-valent pneumococcal conjugate vaccine
PCV10	the 10-valent pneumococcal conjugate vaccine
PCV13	the 13-valent pneumococcal conjugate vaccine
PCR	Polymerase chain reaction
PPSV23	the 23-valent pneumococcal polysaccharide vaccine
rIRR	Ratio of incidence rate ratios
RR	Rate ratio, risk ratio
TESSy	the European Surveillance System
THL	the Finnish Institute for Health and Welfare
USA	the United States of America
UK	the United Kindom
VE	Vaccine efficacy, vaccine effectiveness
VE_{acq}	Vaccine efficacy against acquisition of pneumococcal colonisation
VE_{inv}	Vaccine efficacy against progression from carriage to disease
VE_{prev}	Vaccine efficacy against prevalence of pneumococcal colonisation
VE^{VT}	Vaccine effectiveness against vaccine-type IPD
VE^{NVT}	Vaccine effectiveness against non-vaccine-type PD
$VE^{\text{all IPD}}$	Vaccine effectiveness against all IPD

VI	Vaccine impact
VI_{tot}^{VT}	Total impact against vaccine-type IPD
VI_{tot}^{NVT}	Total impact against non-vaccine-type IPD
$VI_{tot}^{all\ IPD}$	Total impact against all IPD
VPDI	Vaccine preventable disease incidence
VT	Vaccine-type
WHO	World Health Organization

1 INTRODUCTION

Streptococcus pneumoniae is a group of bacteria with more than 90 serotypes. Bacterial colonisation of the nasopharynx is the source of transmission from one human host to another. While usually asymptomatic, pneumococcal colonisation can progress to respiratory or, on rare occasions, to systemic disease. Pneumococcal conjugate vaccines afford protection against pneumococcal colonisation and disease and are therefore able to induce both direct protection for vaccinated individuals and indirect protection (herd immunity) in the population at large through reduced transmission. Nevertheless, as the current conjugate vaccine formulations contain only a select number of pneumoccal serotypes, they can prevent only part of the disease burden caused by pneumococci. Serotypes not included in the vaccines replace those included, both in colonisation and at least partly in disease (serotype replacement). [1,2]

The direct and indirect effects of vaccination on an individual and in the population are of great interest. Vaccine efficacy is a measure of the direct protective effect of a vaccine afforded to a vaccinated individual as compared with an unvaccinated but otherwise similar individual. Vaccine effectiveness is the corresponding measure during an ongoing vaccination programme, when both vaccinated and unvaccinated individuals experience the indirect effects (herd immunity, replacement) of large-scale vaccinations. Vaccine impact is the population prevented fraction of the disease incidence when exposure is the vaccination programme rather than each individual's own vaccination. [3,4]

In this thesis, I consider vaccine efficacy, effectiveness and impact as epidemiological parameters of interest which need to be estimated using observations gathered according to some study design. Many parameters have been used and reported in literature. They describe the individual or population-level protection induced by vaccination and they can be either conditional or unconditional on the amount of exposure to infection in the vaccinated and unvaccinated. Different effect parameters also address the group or time at-risk in different manners. My interest is on choosing a parameter (estimand) of vaccine efficacy that would describe the nature of the infection and the mechanism of vaccine protection. Two questions then follow: how can the parameter be estimated in practice using a feasible study design and how can the estimates be interpreted?

The focus of my thesis is on the definition of different effect measures of vaccines and vaccination, and on the estimation and interpretation of vaccine effects against pneumococcal colonisation and disease. The two endpoints are

different in several important aspects. Episodes of colonisation are common and occur repeatedly as a continuous-time process. Only the prevalent states are directly observable at the time of active sampling [5]. By contrast, disease is rarer than colonisation and more easily observable.

In this thesis, I define the parameter (estimand) of vaccine efficacy against colonisation through acquisition, which describes the natural process of incident colonisation occurrences better than the prevalence. Moreover, I present a general definition of vaccine efficacy against a multi-type pathogen, with an epidemiologically meaningful interpretation as a weighted average of strain-specific efficacies. I focus on the problem of estimating efficacy parameters from cross-sectional data, i.e. without the need to collect repeated measurements per study subject. I propose two improved estimands that account for the differential times at-risk for acquiring the serotypes of interest between the vaccinated and unvaccinated. These parameters can be estimated from cross-sectional data and allow comparison of vaccine efficacy between different settings. I exemplify the estimation and interpretation of vaccine efficacy with empirical data of colonisation in children.

Individual and population-level effects of vaccination programmes are invariably assessed in observational settings. Large-scale vaccinations may induce strong indirect effects, which different effect measures capture differently. Vaccine effectiveness takes the unvaccinated as a simultaneous comparison group although also those are subject to the indirect effects of vaccination. By contrast, vaccine impact implicitly accounts for the indirect effects by choosing the comparison group from the completely unvaccinated population. Both effect measures depend on the population dynamics of pneumococcal colonisation and disease. In this thesis, I describe the differential time evolution of vaccine effectiveness and impact after the introduction of a pneumococcal vaccination programme until vaccine-type disease becomes eliminated in the cohort of vaccine-eligible children. I compare the estimation of vaccine effectiveness by using three parallel observational study designs and register data on invasive pneumococcal disease in Finnish children.

The aim of my thesis is to provide insight into and advice on estimating vaccine efficacy against pneumococcal colonisation. The proposed parameters take into account the underlying process of pneumococcal carriage, and the estimates they produce are therefore comparable between different settings with varying infection pressure. Moreover, the thesis helps to interpret the time-varying parameters of vaccine impact and effectiveness during large-scale vaccinations and their manifestation in Finnish children.

2 REVIEW OF THE LITERATURE

2.1 *STREPTOCOCCUS PNEUMONIAE*

Streptococcus pneumoniae (pneumococci) is a group of gram-positive bacteria. The polysaccharide capsule that covers the outer surface is the most important virulence factor of pneumococci as it protects the bacteria from phagocytosis. The capcule exhibits great diversity with varying antigenic properties and is also the basis for classifying pneumococci into a large number of serotypes [1]. Currently, 100 serotypes have been identified [6].

Pneumococci are part of the commensal flora of the human upper respiratory tract. Together with many other microorganisms such as *Haemophilus influenzae* and *Moraxella catarrhalis,* pneumococci colonise the nasopharyngeal niche. Each episode of colonisation (asymptomatic infection) starts by acquisition of a strain, followed by a period of carriage until clearance of the bacteria after a relatively short period of time. Acquisition of a new strain may occur soon after or even during an ongoing episode of carriage. Simultaneous colonisation of more than one pneumococcal strain is possible. However, pneumococci compete against each other over colonisation of the human host. [1,5,7]

Pneumococcal colonisation is an important source of horizontal spread of the pathogen in the community. Because young children have the highest susceptibility to pneumococcal acquisition and their carriage episodes are generally longer compared to older individuals, children play an important role in maintaining circulation of pneumococci in the population. [1,8]

In rare occasions, pneumococcal colonisation leads to infection of the mucosa, such as acute otitis media (AOM), sinusitis, pneumonia or a systemic infection. Apart from differing in duration of carriage, pneumococcal serotypes and strains may differ in invasiveness, i.e. in their relative ability to progress from carriage to disease per episode of colonisation (case-to-carrier ratio). The most invasive serotypes (e.g. 1 and 5) are usually the least commonly carried, and the most frequently carried (e.g. 6A and 15) are the least likely to cause invasive disease. By contrast, there are no large differences between serotypes in their propensity to cause mucosal infections. [1,2,9,10]

2.2 PNEUMOCOCCAL INFECTION OUTCOMES AND DATA SOURCES

In epidemiological studies and vaccine trials, available data and case definitions vary significantly although standardised outcome definitions have

been presented to facilitate comparison of results. In this section, pneumococcal colonisation and disease are described from the perspective of analysis, i.e. how they have been defined as study outcomes and which types of data are available for estimating their incidence and prevalence. The summary covers all relevant pneumococcal outcomes, although the main focus of the thesis is on pneumococcal colonisation and invasive disease.

Colonisation

Stable pneumococcal colonisation of the nasopharynx is common in infants, of which 40-95% carry at least one pneumococcal serotype at any given time point. The prevalence is highest in young children, peaking during the first two years of life in developed countries and persisting later in developing countries. [11–13]

Adults carry pneumococci more rarely than children, although there is large regional variation. In a study of households in the UK, the prevalence was 8% in adults (≥18 years of age) [11], but a higher prevalence of 22% was observed among study participants 5-39 years of age in Burkina Faso [12].

The prevalence, acquisition or clearance of pneumococcal colonisation have been considered as outcomes in epidemiological studies. Only the prevalent status of colonisation can be directly observed by active sampling from the nasopharynx or the oropharynx, and serotype-specific carriage prevalence as proportion of all samples has been assessed in many studies. The actual times of pneumococcal acquisition and clearance cannot be directly observed but statistical modelling can be used to impute such event times based on longitudinal data on the observed status of carriage [14,15]. When considering study outcomes, serotypes are often grouped according to some relevant property, e.g. according to whether they are included or not in a vaccine. Moreover, the density of colonisation can be used as a study outcome, measured with microbiological methods and quantitative PCR [16].

Invasive pneumococcal disease

Laboratory confirmation of invasive pneumococcal disease (IPD) from blood or cerebrospinal fluid provides a highly specific yet insensitive outcome, typically classified by serotype or serogroup. ECDC [17] defines the laboratory criteria of IPD for surveillance as at least one of the following: (i) isolation of *Streptococcus pneumoniae* from a normally sterile site, (ii) detection of *Streptococcus pneumoniae* nucleic acid from a normally sterile site, (iii) detection of *Streptococcus pneumoniae* antigen from a normally sterile site. In many countries, surveillance of IPD is organised by a hospital-based sentinel system or by enhanced population-based surveillance with a defined catchment population [18]. In Finland, all clinical microbiology laboratories are required to notify all isolations of *Streptococcus pneumoniae* from blood or

cerebrospinal fluid to the National Infectious Disease Register, a population-based, electronic laboratory surveillance system maintained by the Finnish Institute for Health and Welfare (THL) since 1995 [19].

The incidence of IPD is high in early childhood and decreases thereafter, increasing again in old age. According to the European Surveillance System (TESSy), the average incidence of IPD in Europe was 6.2 cases per 100 000 person-years in 2017 but much higher among infants and adults over 65 years of age (14.5 and 18.9 cases per 100 000 person-years, respectively) [20]. In high-risk areas, the incidences are often higher. For example, in South Africa where HIV is an important risk factor of pneumococcal diseases, the incidence of IPD after the onset of the 7-valent pneumococcal conjugate (PCV7) vaccination was 17 and 8 cases per 100 000 person-years among children <2 years and in adults 25-44 years of age, respectively [21].

Clinically suspected invasive pneumococcal disease

Not all cases that meet the clinical criteria of IPD are confirmed with a laboratory test. An isolate may not be taken, antimicrobial treatment may have started before confirmation of the causative pathogen, or blood culture practices at hospital or recording into registers may be suboptimal. Therefore, Palmu et al. [22,23] considered a more sensitive study outcome based on discharge notifications of inpatient and outpatient care with available primary or secondary diagnoses of ICD10-codes relating to IPD. An episode definition was used to prevent the same illness to be counted more than once. The incidence of the most sensitive disease outcome, 'register-based non-laboratory confirmed IPD or unspecified sepsis' was 237 cases per 100 000 person-years among vaccine-eligible children aged 3-42 months between years 2010-2013, while the incidence of laboratory confirmed IPD in the same cohort was 13 cases per 100 000 person-years [23].

Pneumonia

To assess the burden of community acquired pneumonia (CAP), hospital secondary data from administrative data sources are commonly used. As the disease definition may vary significantly by physician and national guidelines, a standardised definition based on alveolar consolidation of chest radiography was developed by a WHO working group [24].

For example, the incidences of hospital-diagnosed pneumonia in the FinIP trial were 13 and 10 cases per 1000 person-years in the control children and the children vaccinated with the ten-valent pneumococcal conjugate vaccine (PCV10), respectively, whereas the respective incidences of radiologically confirmed consolidated pneumonia (WHO definition) were 4 and 2 cases per 1000 person-years [25]. Using the same WHO definition, the incidences in the

Gambia were 41 and 26 cases per 1000 person-years among the control and vaccinated children, respectively [26].

Microbiological confirmation of pneumonia caused by pneumococci remains challenging. An immunochromatographic urinary antigen test is used in hospitals to diagnose pneumococcal pneumonia, with sensitivity of approximately 60%. This test detects the C-polysaccharide of *Streptococcus pneumoniae* that is common to all serotypes [27]. A serotype-specific urinary antigen test to identify the serotypes in the 13-valent pneumococcal conjugate vaccine (PCV13) with high sensitivity is currently available in research settings [28].

Acute otitis media (AOM)

Pneumococci, *Haemophilus influenzae* and *Moraxella catarrhalis* have been described as the most common causes of bacterial AOM [29]. In studies assessing the disease burden of AOM, the causative pathogen is typically not known. Several case definitions based on clinical confirmation, secondary data or parental questionnaires have been used to define AOM outcomes, with different scoring systems to categorise symptoms and to assess the severity of the disease [29–34]. Surrogate outcomes such as tympanostomy tube placement for recurrent, prolonged and complicated otitis, and antimicrobial purchases for uncomplicated AOM have been used as well [35].

The incidence of AOM is highest in children younger than 2 years [33]. In this age, the children have on average 1.5 attacks of acute otitis media per person per year [36]. In Icelandic children aged 1-2 years, the incidence of hospital visits related to AOM was 106 cases per 1000 population during the post-PCV10 period [29]. In Finland, the incidence of tympanostomy tube placements was 46 cases per 1000 person-years and that of antimicrobials recommended for treatment of AOM 890 cases per 1000 person-years among vaccine-eligible childen after PCV10 introduction [35].

2.3 PNEUMOCOCCAL CONJUGATE VACCINES

Pneumococcal conjugate vaccines (PCVs) contain purified capsular polysaccharides of a select number of pneumococcal serotypes, conjugated to a carrier protein. While the immune system of infants is not yet mature to mount a protective antibody response with plain polysaccharide formulations, conjugation improves the response. [5,37]

Pneumococcal conjugate vaccines have been used since 2000, when a 7-valent vaccine was licensed in the US with brand name Prevnar (in this thesis called PCV7). Current PCV formulations contain antigens of 10 or 13 serotypes

(brand names Synflorix and Prevnar 13, in this thesis called PCV10 and PCV13, respectively). The seven-valent vaccine is no longer manufactured. [13]

Another 10-valent conjugate vaccine, Pneumosil, was prequalified by the WHO in January 2020 and launched in December 2020. Pneumosil is intended to the low and middle-income countries with better matching serotype coverage in these areas as compared to the current PCVs [38]. Two investigational conjugate vaccines including 15 and 20 serotypes are currently studied in different phases of clinical evaluation. The 15-valent vaccine is intended for immunisation of both paediatric and adult populations, whereas the 20-valent vaccine has been designed primarily for prevention of the high burden of CAP in adult populations [39,40]. Moreover, two 24-valent conjugate vaccines are currently in human trials, and a 30-valent preclinical product has been announced [41].

The new wider-valent vaccines may address much of the remaining pneumococcal disease burden, but this fraction will depend on the level of serotype replacement with the non-vaccine serotypes and the emergence of non-encapsulated strains. Therefore, development of a universal, non-serotype pneumococcal vaccine has also been considered. A small number of pneumococcal protein-based vaccines as well as a whole-cell vaccine have progressed to the early stages of clinical evaluation [42–44].

2.4 VACCINE EFFECTS AND EFFECT MEASURES

Many aspects affect the choice of vaccine effect measures, their practical estimation and interpretation. Pneumococcal vaccines prevent both colonisation and disease and thus induce direct and indirect effects on the individual and population levels [45]. Moreover, the direct vaccine-induced protection can act through two different mechanisms by either reducing the rate of infection in all vaccinated individuals, or by protecting some of the vaccinated completely and leaving others without any protection. Another aspect is the choice of study design and statistical rules (estimators) to optimally estimate the effect measures of choice (estimands). Finally, the interpretation of the ensuing estimates may depend on the choice of the study outcome, estimand and estimator in complicated ways. This and the following sections cover several aspects of measuring vaccine effects that are relevant in the pneumococcal context.

In addition to the direct protective effect afforded to a vaccinated individual against pneumococcal disease, conjugate vaccines have another important ability. They protect against colonisation, which reduces the chances of transmission and further spread of the bacteria. On the population level, this leads to reduced carriage acquisition and eventually reduced numbers of disease cases. While this effect, called herd immunity, applies to those

serotypes included in the vaccine, non-vaccine serotypes may replace the vaccine types in carriage and at least partly in disease (serotype replacement). [2,5]

Vaccination affords direct protection separately against the two steps in the causal pathway from exposure to pneumococcal colonisation and further to disease. At the first step, the vaccine may protect an individual against pneumococcal carriage (including its acquisition, duration or density, i.e. the quantitative load of pneumococci in the nasopharynx). While vaccine-induced protection against acquisition and density have been observed [46], the evidence for the effect on clearance is more scarce [14,47,48]. At the second step, even if a vaccinated individual becomes a carrier of one of the serotypes included in the vaccine, vaccination may protect against carriage progressing to disease. This is the protective effect of the vaccine against case-to-carrier ratio, i.e. the conditional risk of disease given acquisition of carriage [2].

Replacement follows from different pneumococcal strains competing for colonisation of the nasopharynx. Specifically, serotype replacement is expected to occur if carriage of vaccine-type pneumococci protects the host from acquiring non-vaccine-serotypes. Because vaccination reduces vaccine-type carriage, the nonvaccine serotypes gain better ability to colonise the opened niche. This phenomenon, called within-host replacement, leads to differential times spent at-risk for acquiring the vaccine serotypes between the vaccinated and unvaccinated individuals. Following wide-spread use of conjugate vaccines, the same mechanism occurs as serotype replacement on the population-level [49]. The population-level replacement in carriage has been observed to take place readily after vaccination and be complete so that there has been little or no net change in pneumococcal carriage prevalence after vaccination [50]. However, replacement is often only partial in disease, reflecting the varying invasive potential (case-to-carrier ratio) of the serotypes involved [45,51].

If the vaccine protects some individuals perfectly while leaving others out of any protection, the effect of the vaccine is called all-or-nothing. The vaccine may also reduce the hazard of infection in all vaccinated individuals, so that vaccinated individuals may still eventually become infected. The vaccine effect is then called leaky [3,52]. A combination of all-or-nothing and leaky effects is possible as well [53]. Mehtälä et al.[14] applied a Markov transition model to longitudinal data of colonisation in Israeli toddlers and inferred that the leaky mechanism appears to be the prominent mode of vaccine action in the context of a pneumococcal conjugate vaccine.

Vaccine effect measures, i.e. efficacy, effectiveness and impact, can be considered as statistical estimands. In other words, they are parameters of interest that need to be measured using observations gathered according to some study design. Depending on the level of available information and on the

assumed mechanism of the vaccine action, estimands may be based on risk (i.e. probability) or rate (i.e. hazard). They can also be conditional or unconditional on the current state of colonisation of an individual, as well as on the amount of exposure to infection in the vaccinated and unvaccinated groups.

An estimator is a rule that is used to calculate an estimate for the parameter of interest based on observations. Different study designs allow different estimators to be used even for the same effect parameter. Different estimators also rely on different assumptions and are prone to different types and amounts of bias. Therefore, estimates are seldom perfect representations of the true vaccine effects, and their interpretation depends on the study design that was employed.

In the next few chapters, I cover vaccine efficacy, effectiveness and impact in terms of their theoretical definitions, practical estimation and interpretation.

2.5 VACCINE EFFICACY

2.5.1 DEFINITION

Vaccine efficacy is a measure of the direct protective effect of the vaccine and is defined as the relative reduction in susceptibility to infection, conditionally on a specific amount of exposure in a vaccinated person compared to an unvaccinated person [54]. Vaccine efficacy is intended to be a measure of the causal effect of vaccination on an individual as compared to a situation where the (same) individual was not vaccinated [55].

Vaccine efficacy is formally defined as $\mathrm{VE} = 1 - \mathrm{RR}$, where RR is some measure of relative risk or relative rate. In practice, the measure of risk underlying RR can be the attack rate, secondary attack rate, hazard or prevalence. In this thesis, the term rate refers to a hazard. The attack rate, however, is an exception as it is a proportion (see below).

Already in 1915, Greenwood and Yule [56] listed three conditions that need to be fulfilled in order to be able to estimate vaccine efficacy from empirical observations. First, the vaccinated and unvaccinated must be alike in all material respects, i.e. they must not differ in any such factors that may affect the liability to contract the disease. Second, the effective exposure to infection must be identical between the vaccinated and unvaccinated. Third, the confirmation of a person's vaccination status must be independent of the disease occurrence, i.e. blinded at the time of case confirmation.

2.5.2 STUDY DESIGN AND THE CHOICE OF VACCINE EFFICACY MEASURE

Greenwood and Yule [56] proposed to use attack rates to assess the protective efficacy of a vaccine. This means that vaccine efficacy is defined as the relative reduction in the attack rate, i.e. in the risk of infection in a closed population during a specific period of time, e.g. an outbreak, as follows:

$$\mathrm{VE} = 1 - \mathrm{RR} = \frac{\mathrm{ARU} - \mathrm{ARV}}{\mathrm{ARU}},$$

where ARU and ARV are the attack rates in the unvaccinated and vaccinated, respectively [57].

Rothman et al. [57] and the IEA Dictionary of Epidemiology [58] recommend using the term incidence proportion instead of attack rate. In this thesis, however, I use the term attack rate in order to agree with the established terminology in infectious disease epidemiology.

To fulfill the second condition of Greenwood and Yule [56], regarding the similarity of exposure to infection among the vaccinated and unvaccinated, it was noted already early that the attack rates should be estimated within small groups such as susceptible family members of an infectious disease case using secondary attack rates [59]. Haber et al. [60] defined formally the transmission probability as the probability that, conditionally on a contact between an infective source and a susceptible host, a successful transfer of the infectious agent will occur so that the susceptible host becomes infected. Thus, knowing the full contact history of the study participants would allow unbiased estimation of vaccine efficacy.

In general, the study objectives, feasibility and setting determine the type of data that will be available and thus largely influence what epidemiological effect measures can be estimated. Rhodes et al. [61] presented a hierarchy of vaccine efficacy effect measures, based on the amount of information available on individual-specific histories of the contacts and infection.

The hierarchy of Rhodes et al. proceeds from level I to level IV [61,62]. At level I, all contacts between infectives and susceptibles are known (who is infectious and when, and whom they contact). Vaccine efficacy can be assessed based on transmission probabilities and secondary attack rates. At level II, data on infective contacts per time period are available and the estimator can be based on hazard ratios. At level III, the observed order of events allows estimation of hazard ratios under the proportional hazards assumption. At level IV, only data on who became infected during the study period are available. This allows estimation of vaccine efficacy in terms of attack rates.

There are essentially two types of vaccine efficacy parameters: conditional parameters that aim to ensure equal exposure to intection between the vaccinated and unvaccinated groups, and unconditional parameters that may be biased due to unequal exposure. Although more information is required to estimate conditional parameters, vaccine efficacy based on e.g. secondary attack rates is biologically interpretable and robust to different transmission conditions and the indirect effects of vaccination. While moving down the levels of information hierarchy of Rhodes et al. [61] one loses the ability to condition on the actual contact between an infectious source and a susceptible person. Using these unconditional parameters in two settings with different exposure to infection, even under randomisation, could yield different efficacy estimates [63].

Another aspect that should guide the choice of the effect measure is the mode of vaccine action. A vaccine effect may be leaky or all-or-nothing. If a wrong assumption about the mode of action is made when choosing the vaccine efficacy parameter, the estimate may be biased or difficult to interpret (**Figure 1**). For an all-or-nothing vaccine, risk-based estimation is appropriate. By contrast, if the vaccinated individuals benefit from (leaky) vaccination through reduced hazard of infection, an appropriate VE parameter should take into account the person-times at risk [64].

The implications of a wrong assumption about the vaccine's action are more serious if case occurrences are common. This is the situation with pneumococcal carriage. In case of a rare outcome such as IPD, an incorrect assumption of the effect mechanism should not bias the estimation of vaccine efficacy [52].

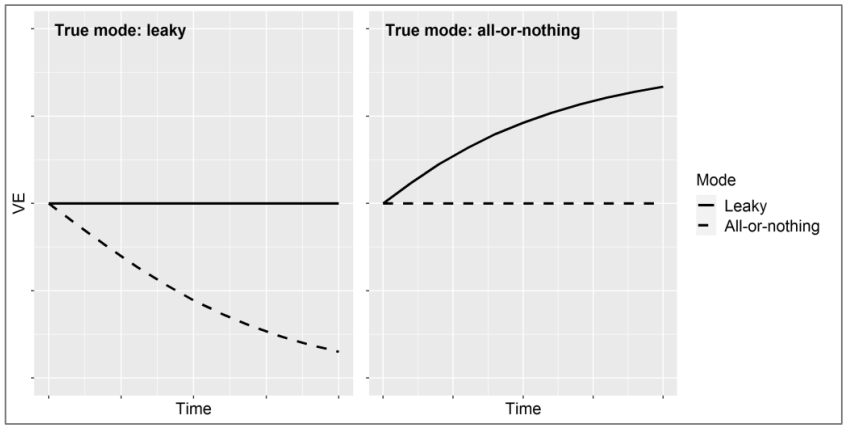

Figure 1 *Two models of the action of a vaccine. On the left panel, the true mode of action is leaky, i.e. vaccination reduces the hazard of infection in all vaccinated individuals. If an all-or-nothing mode is assumed, vaccine efficacy is underestimated. On the right panel, the true mode is all-or-nothing, i.e. the vaccine protects some individuals perfectly but leaves others out of any protection. An incorrect assumption of a leaky vaccine overestimates the efficacy.* [52]

2.5.3 VACCINE EFFICACY AGAINST PNEUMOCOCCAL INFECTION IN TRIALS

The efficacies of pneumococcal conjugate vaccines (PCVs) against pneumococcal colonisation and disease have been estimated in a number of individually randomised trials. In these trials, several parameters have been used to quantify vaccine efficacy.

Rate or hazard ratios have been estimated to assess vaccine efficacy against invasive pneumococcal disease (IPD), pneumonia and acute otitis media (AOM) [26,30,34,65–68]. Some studies have also taken into account the possibility of recurrent episodes in the same child. Dagan et al. [47] used odds ratios as estimators of the relative hazard of the first carriage acquisition of a child in the recipients of the 9-valent conjugate vaccine compared with control subjects.

Risk ratios have been used to assess vaccine efficacy against pneumococcal carriage [68,69] and disease [70,71]. Finally, in some studies the difference in prevalence of both carriage and disease has been tested between the vaccinated and unvaccinated [65,72].

In 1993, the safety and immunogenicity of a 5-valent PCV was investigated in infants in a study conducted in the Gambia. After three doses of PCV and

revaccination with a polysaccharide vaccine, the odds of vaccine-type colonisation was reduced by 89% compared to the control group [73].

Two trials assessed the efficacy of a 9-valent PCV against pneumococcal colonisation in Israel and Soweto, South Africa [47,72]. A 50% reduction in the rate of new vaccine-type acquisition was observed among Israeli toddlers during the two-year follow-up. In this analysis, a new acquition was specifically defined as the first observation of a serotype per child, and rate reduction was defined through odds ratios. Also in Soweto, a clear reduction of 18% was observed in the proportion of vaccine-type colonisation at 9 months of age.

The efficacy of two 7-valent PCVs against AOM was assessed in the FinOM trial conducted in the Tampere area, Finland [30,66]. Recurrent episodes of acute otitis media in the same child were evaluated to estimate hazard ratios. Both vaccines reached practically identical point estimates of 57% and 56% against vaccine-type AOM, and clearly negative point estimates of -33% and -27% against non-vaccine-type AOM. In a trial in Czech Republic and Slovakia, almost similar point estimate of 53% against the first episode of vaccine-type AOM was reported using similar method to estimate hazard ratio [33]. A clearly positive vaccine efficacy of 7% against all-cause otitis media was estimated with PCV7 in the Kaiser Permanente trial in Northern California [65].

Several trials have assessed vaccine efficacy against pneumonia, resulting in broadly similar estimates with different vaccines. In the Gambia and Soweto, South Africa, radiologically confirmed pneumonia was reduced with a 9-valent PCV by 37% and by 20% among children without HIV infection, respectively [26,71]. In Bohol, the Philippines, an 11-valent PCV reached to a vaccine efficacy of 23% (-1…41%) against community-acquired pneumonia (CAP) [67]. In the COMPAS trial in Argentina, Panama and Colombia, the efficacy of PCV10 against consolidated CAP was 26% [34]. In South Africa, vaccine efficacy was defined through risk ratios, and in all other studies through hazard ratios.

The trials of a 9-valent PCV in South Africa and PCV7 in California estimated high efficacies of 83% and 97% against vaccine-type IPD, respectively [65,71]. Somewhat lower efficacy of the 9-valent PCV was estimated in the Gambia, 77% [26]. In Latin American children, the efficacy of PCV10 against vaccine-type IPD was as high as 100% [34].

In the CAPiTA trial, the efficacy of PCV13 was assessed among adults ≥65 years of age in the Netherlands [70]. Vaccine efficacy against vaccine-type pneumococcal CAP was 46% and against vaccine-type IPD 75%.

2.6 VACCINE EFFECTIVENESS

2.6.1 DEFINITION

Vaccine effectiveness quantifies the relative reduction in the hazard or risk of infection that vaccination affords to a vaccinated individual during an ongoing vaccination programme and can be defined as VE = 1-RR. It is an effect measure which corresponds to vaccine efficacy in the context of routine use of a vaccine and has also been called field efficacy [60]. In this thesis, both vaccine efficacy and vaccine effectiveness are denoted by VE.

Vaccine effectiveness is invariably estimated in observational settings. The vaccinated and unvaccinated are compared in parallel, although also the unvaccinated are subject to the indirect effects of vaccination. The effectiveness parameter is not explicitly conditioned on replacement in infection. However, control individuals are often sampled from those at risk for infection at the time of the case occurrence, i.e. matching on the current level of the force of infection.

2.6.2 STUDY DESIGNS IN ESTIMATING VACCINE EFFECTIVENESS

The broad types of study designs that have been used to estimate vaccine effectiveness are the cohort and case-control designs, with several modifications depending on the amount and type of available data.

In a cohort study, the underlying population consists of individuals who have an equal possibility to be identified as cases if they contract the infection in question during the follow-up period [74]. In contrast to a randomised experiment, in which the exposure status of each study participant is assigned at random, the individual itself (or his/her parent) decides whether or not to be vaccinated. Moreover, the properties of the vaccinated and unvaccinated that affect the liability to contract the disease may not be randomly assigned, violating the first condition of Greenwood and Yule [56].

The source population is formally defined and enumerated, the vaccination statuses of the cohort members are identified and these are followed over time until the study outcome or end of follow-up. Although such prospective data collection is laborious if data are not available from health registers, it provides a temporal framework and offers the potential to infer causal effects of vaccination. Study subjects can also be selected at the present time while their disease and vaccination statuses are assessed retrospectively [75]. This is a typical approach when using register data whose collection process had not been specifically designed for the study at hand. Nevertheless, data are more immediately and less costly available than in a prospective study.

With the cohort design, hazards can be directly estimated. Two other often-used measures of disease occurrence are the incidence proportion (average risk), i.e. the proportion of subjects who experience the outcome at any time during the follow-up, and incidence odds, i.e. the ratio of the number of subjects who experience the outcome versus the number of those that do not experience it. These measures share the same numerator, the number of incident cases. The difference is in denominators that involve person-time at risk, persons at risk, or survivors. The corresponding effect measures are hazard ratio, risk ratio and odds ratio. [76]

A cohort study may be very inefficient because large amounts of data are collected for those not experiencing the infection, although the same precision in the case-control comparison would be achieved with much fewer controls. This is one motivation for the case-control design. While the underlying population is perceived similarly to the cohort design, cases are identified first and controls are only then selected from the same source population. Data regarding the exposure status are collected retrospectively for both groups. [77]

In a case-control study, cases and controls should meet the same inclusion criteria except for the outcome. Efforts should be focused on accurately ascertaining the disease status and vaccination history in both groups. However, as data of a relatively small number of study subjects (compared to cohort studies) are typically collected, the method can be resource-efficient and particularly useful for diseases or outcomes that are uncommon. [75,76]

The apparent relative risk measure in case-control studies is the odds ratio. The actual estimand which the odds ratio will approximate, however, depends on how the controls are selected. Correspondingly to the cohort design, controls can be selected randomly from the survivors at the end of the follow-up, from the source population at risk at the beginning of the follow-up, or from the risk set, i.e. longitudinally throughout the course of the study. Odds ratios obtained will then estimate three different estimands, namely odds, risk, or hazard ratios in the source population, respectively. [64,76,78]

The screening method, in which the vaccination status of cases is compared to the population-level vaccination coverage, is another approach for assessing vaccine effectiveness [79]. This method is used in situations where data on control individuals are not available or reliable, and where the validity of the external estimate of vaccination coverage is considered reliable. In contrast to cohort and case-control designs, the proportion of vaccinated in the underlying cohort or among the controls is not estimated, but instead the true (assumed) level of the population coverage is used. The relative risk of disease is estimated through the odds of vaccination in cases compared to that in population. The main shortcomings of the screening method are that the accuracy of the vaccination coverage cannot usually be tested, and that a

detailed stratified analysis of risk factors may not be possible due to unavailability of stratified estimates of vaccination coverage.

The indirect cohort design is an attractive approach where data on vaccination coverage are not needed and only surveillance data are required in order to estimate the vaccine effectiveness. The method was initially suggested for assessing the effectiveness of a pneumococcal polysaccharide vaccine in immunologically impaired individuals in a phase IV setting [80]. Intrinsically, the indirect cohort approach is a modified case-control method in which the odds of vaccination in vaccine-type IPD cases is compared to that in non-vaccine-type IPD cases. The usefulness of the method has been questioned due to the apparently non-valid assumption that vaccination has no effect on non-vaccine type carriage [81]. This assumption is needed to ensure that the odds of vaccination among the controls equals to that of the catchment population. The size of the bias has been studied and concluded small (2-5%), especially if adjusted for the calendar time-period. An advantage of the method is that the non-vaccine cases serve as well-matched controls in terms of risk factors and use of health care [82,83].

The cohort design allows the use of almost all available information on the source population over the follow-up period. Therefore, the effect measures on the levels of the hierarchy of Rhodes et al. [61] are estimable, except (usually) the contacts between individuals and the relative transmission probabilities of level I. The case-control design, if appropriately matched, also provides approximations of hazard and risk ratios at the information level II. The screening and indirect cohort methods are case-control designs as well but require strong assumptions to be valid.

2.6.3 VACCINE EFFECTIVENESS AGAINST PNEUMOCOCCAL INFECTION IN OBSERVATIONAL STUDIES

The most common approaches that have been used to estimate vaccine effectiveness against pneumcoccal disease and colonisation are the matched case-control and the indirect cohort designs [82,84,85]. The most common effect parameter has been hazard ratio. The screening method has been used e.g. in Italy [86]. The cohort design is rarely feasible, but it was employed in New South Wales and Western Australia in the entire child population to estimate the effectiveness of PCV13 against IPD [87]. The two Finnish studies that have employed cohort design are presented in this thesis [88,89].

The estimated effectiveness of PCV7 against vaccine-type IPD was usually high, but there has been some variation. For example, the effectiveness was 88% in Germany [85] and 94% among non-Aboriginal children in Australia [87], but somewhat lower in the UK (56%, 95% confidence interval -7…82%) [82]. For the same outcome, the effectiveness of PCV10 was estimated at 73% and 84%

in two studies in Brazil [90,91] and that of PCV13 at 75% in the UK [92] and 97% in the US [93].

Matched case-control studies were conducted to estimate the effectiveness of PCV13 against radiologically confirmed and presumed bacterial pneumonia in the Gambia and South Africa. The point estimates were 43% and 39%, respectively [94,95].

2.7 VACCINE IMPACT

2.7.1 DEFINITION AND ESTIMATION

Vaccine impact is defined as the relative reduction in the incidence of infection in a population that experiences the vaccination programme compared to the incidence in a completely unvaccinated population [4]. Vaccine impact is expressed as 1-RR, where RR is the ratio of the incidences (vaccinated vs. unvaccinated). Here, incidence is typically interpreted as the rate of infection on the population level.

Vaccine impact is usually estimated in observational settings, where some part of the population is vaccinated. Vaccinated individuals experience both direct protection through the vaccine-induced immune response and indirect protection through the reduced exposure to infection, and the total impact quantifies this net benefit. Indirect impact (herd protection) concerns the unvaccinated part of the population, which benefits from reduced transmission during a large-scale vaccination programme. The overall impact is the weighted average of the total impact in the vaccinated and the indirect impact in the unvaccinated. The overall impact is the parameter that is usually measured in observational studies, as the vaccination status of individuals is often unknown.

A usual approach to assess vaccine impact is the so called before-after design, in which the incidences of infection before and after vaccine introduction are compared. Another approach is to choose the unvaccinated comparison populations from geographically distinct areas. In a cluster-randomised trial, all eligible, consenting individuals in a cluster are administered the same vaccine depending on the treatment arm of the cluster. [96]

In the before-after design, some issues arise from the use of a historical control. To accurately estimate the baseline incidence, several years of data before introduction are preferable, although it is often not clear how long a pre-vaccination period is needed or when to start the post-introduction period. Sometimes a transition period following immediately the vaccine introduction, when the vaccination coverage is still rising, is left out of the analysis. [97]

The incidence of infection may vary also due to factors not related to vaccination. There may be secular trends and cyclical variations in serotype distribution, or changes in case detection, surveillance methods and case definitions as well as changes in population characteristics [4]. Interrupted time series (ITS) analysis is a quasi-experimental design that can be used to account for the time trend in the incidence of infection before vaccine introduction. The post-intervention time trend in the hypothetical absence of the intervention (the counterfactual) is then compared with the observed post-intervention time trend [98].

The key assumption of the ITS method is that the pre-intervention trend continues unchanged into the post-intervention period and there are no external factors systematically affecting the trend. To help the interpretation, Thorrington et al. [99] proposed to compare the observed post-introduction incidence rates with control conditions that are not likely affected by vaccination or other public health interventions. They calculated age-specific ratios of incidence rate ratios (rIRR) for a pneumococcal disease outcome over each of the five control conditions separately, as well as over the composite of the conditions. Ratios greater than one imply an increase in the incidence of pneumococcal infection in the post-introduction period greater than seen in the control conditions. By contrast, ratios less than one imply a decrease in pneumococcal infection that is greater than expected based on the control incidence.

Another approach for choosing the control conditions was presented in a study that assessed the impact of PCV10 and PCV13 on pneumonia hospitalisations in five countries in the Americas [100]. The authors derived, separately for each country and age-group, a composite control by first aggregating pre-vaccination data across 17 ICD-10 chapters and 20 additional conditions and then assigning weights for each of them to generate a synthetic composite control whose pre-vaccination trend best matched the pre-vaccination pneumonia trend. The post-vaccination data from the weighted synthetic controls were then used as a counterfactual against which the impact of PCV10 and PCV13 was assessed.

2.7.2 VACCINE IMPACT AGAINST PNEUMOCOCCAL INFECTION IN OBSERVATIONAL STUDIES

The impact of PCVs has been assessed against pneumococcal colonisation and several disease outcomes including invasive pneumococcal disease (IPD, meningitis and/or non-meningitis), pneumonia and (acute) otitis media in paediatric and adult populations. The interest has been in the impact against the disease burden due to all serotypes as well as against serotype-specific incidence. The type of available data has often driven the choice of the outcome and study design.

Colonisation

The impact of PCVs on colonisation has been investigated in cross-sectional studies as well as using prospective surveillance e.g. in Massachusetts and Turkey [101,102]. PCVs have led to substantial decreases in vaccine-type colonisation in children eligible for vaccination programmes both in high- and low-income settings. In Iceland, the total impact of PCV10 was as high as 94% [103]. Similar reductions of 63% and 64% in vaccine-type colonisation were observed in vaccine-eligible children in Fiji and Kenya, respectively, three years after the introduction of PCV10 [104,105]. Among American Indian children <5 years of age, the decline in PCV13-specific colonisation was 60% two years after the introduction of PCV13 compared to the PCV7-era [106].

Vaccinations have induced also some indirect impact against vaccine-type colonisation in unvaccinated age-groups both in high and low-income settings [103,107,108]. Non-vaccine serotypes have offset the decrease in vaccine serotypes after PCV introduction in many settings [103,108–110].

Invasive pneumococcal disease

Prior to the introduction of PCVs, 6–11 most common serotypes accounted for over 70% of all IPD in children worldwide [18]. These serotypes are covered by the current PCVs, which have led to strong decreases in vaccine-type IPD in pediatric populations e.g. in Finland [19,111]. PCVs have also induced indirect impact on unvaccinated older populations. However, although children continue to benefit from vaccination programmes, increases in non-vaccine-type IPD compromise the benefits in adult age groups in many countries [112–115]. Somewhat controversial results have been reported from the US, where no clear increase in the non-vaccine-type IPD incidence has been detected in adult age groups [116].

In Brazil, PCV10 reduced the incidence of all IPD by 44% in children 2-23 months 3 years after introduction. No impact was observed in unvaccinated age groups, but an increase in the incidence of all IPD was detected in adults ≥18 years [117]. In France, PCV13 implementation led to a major reduction of 44% in the incidence of all IPD across all age groups, but a rebound in the incidence in both children and adults occurred five years after introduction due to the emergence of several non-PCV13 serotypes [118].

Clinically suspected IPD

In addition to the specific outcome of laboratory confirmed IPD, the impact on the more sensitive clinically suspected IPD was evaluated in Finland in order to better assess the overall disease burden of pneumococcal disease among vaccine-eligible children. The point estimate of the relative reduction was smaller than that of the laboratory confirmed IPD (34% vs. 80% [19]), but the

absolute reduction in incidence was more than twofold (122 vs. 50 per 100 0000 person-years). [23]

Pneumonia

Similarly to IPD, PCVs have clearly decreased the incidence of pneumonia in children. In Kenya, the impact of the PCV10 programme decreased hospital admissions of clinically defined and radiologically confirmed pneumonia by 27% and 48%, respectively, 13 years after introduction among children aged 2-143 months [119]. In Finland, the incidence of pneumonia hospitalisations decreased by 23% and 18% in vaccine-eligible and older unvaccinated children after PCV10 introduction, respectively [120]. In the Netherlands, the net impact of PCV7 and PCV10 on community-acquired pneumonia was clear in children up to 2 years of age, and some decrease in the incidence was observed also in older age groups [121].

With regard to the elderly, declines were observed in all-cause pneumonia hospitalisations in Finland [122] and in pneumococcal pneumonia hospitalisations in Portugal and UK [99,123]. By contrast, increasing trends in pneumonia hospitalisations were observed in Brazil [124].

Acute otitis media

A number of post-licensure studies have reported varying impact estimates of PCVs on AOM-related endpoints. In a Finnish study of children 3-54 months of age, the relative rate reductions were 18% and 15% in antimicrobial purchases and tympanostomy tube placements, respectively, used as surrogate endpoints of AOM and recurrent, prolonged and complicated otitis [35]. In Iceland, a 24% reduction was observed in hospital admissions and visits due to AOM among children aged 3-23 months [29]. In the UK, a 22% reduction in otitis media diagnoses was observed after PCV7 introduction, and a subsequent decrease of 19% after PCV13 introduction in children <10 years of age based on the national primary care database of general practitioners [125].

2.7.3 VACCINE IMPACT AGAINST PNEUMOCOCCAL INFECTION IN TRIALS

In Navajo and White Mountain Apache Indian reservations in USA, 8292 children were enrolled in a study between April 1997 and May 2000. There were 38 randomisation units, 36 units for Navajo and two for Apache, that were defined by geography and population size to group communities with significant social interactions of adults and children into the same randomisation units [126]. Within each unit all enrolled children received the same vaccine, i.e. PCV7 or meningococcal type C conjugate vaccine as control [127].

The per-protocol total impact against the primary outcome, vaccine-type IPD was 77% (95% CI -9…95%) and that against all IPD 54% [127]. The total impact against vaccine- and non-vaccine-type colonisation were 60% and -33%, respectively. The indirect impact among unvaccinated children against vaccine-type colonisation was 73% (-7…93%). [48]

The FinIP vaccine trial was a nation-wide cluster-randomised, double-blind phase 3-4 field trial that was conducted in 72 geographical areas in Finland. The enrolment and vaccinations took place in the well-baby clinics of the participating public health care centres serving altogether nearly 80% of the Finnish population. The enrolment of children aged 6 weeks to 18 months started in February 2009 and ended, as planned, when PCV10 was introduced into the national vaccination programme in September 2010. The blinded follow-up lasted until the end of January 2012. Enrolled children were vaccinated with PCV10 or a control vaccine (hepatitis A or B) according to either 3+1 or 2 +1 schedule. [128]

The data on disease outcomes were collected from established administrative national health registries: invasive pneumococcal disease from infectious disease register, clinically suspected, non-laboratory confirmed IPD and pneumonia hospitalisations from hospital discharge register, and two surrogate outcomes for otitis media, tympanostomy tube placements and antimicrobial purchases, from benefits and hospital discharge registers. Episode definitions were used to ensure that repeated hospital visits and admissions due to the same illness were not counted more than once in the analyses. [22,23,25,128,129]

The total impact of the combined 2+1 and 3+1 schedules against the primary outcome, vaccine-type IPD was estimated at 100% and that against all IPD at 93% [128]. The total impact against clinically suspected IPD was 50% [22] and that against hospital-diagnosed pneumonia and consolidated pneumonia 27% and 45%, respectively [25]. The point estimates for tympanostomy tube placements and antimicrobial purchases were 13% (95% CI -2…26%) and 7%, respectively [129,130]. No major differences were observed between the 3+1 and 2+1 schedules with any of the trial outcomes. In a post-hoc subgroup analysis, no clear differences in impact were observed by sex, gestational age or birth weight [131].

In the FinIP, the total impact was defined as 1 minus the rate of the disease outcome in question in PCV10 clusters compared to that in control clusters. It was estimated using negative binomial regression allowing for possible overdispersion due to the cluster design, i.e. variability beyond what is expected if the incidence rate is assumed to follow the same Poisson distribution across all clusters of the same treatment group [132]. In the published articles and elsewhere [133], the effect measure was called effectiveness in order to make the difference with individually randomised

trials that aim to estimate vaccine efficacy, as well as to account for the parallel design. O'Brien et al. [127] called the effect measure in their trial vaccine efficacy, but discussed that it is not what is usually estimated in individually-randomised trials but instead a combination of the direct and indirect effects of vaccination. In this thesis, the effect measure of cluster randomised design is called impact, highlighting the fact that the control children are not exposed to the indirect effects in contrast to the PCV vaccinated study participants.

There are some limitations in the cluster randomised design, such as the increased difficulty to maintain masking, reduction of the effective sample size, and the potential mixing of the intervention and control populations [127]. The results are less generalisable with other settings compared to the vaccine efficacy from individually randomised trials, as the effect size depends strongly on vaccination coverage [62].

Jaffar et al. [134], when planning a trial in the Gambia, listed the potential limitations of a cluster randomised trial that could be encountered in their setting. Blinding may be difficult to sustain, especially if the vaccine would have a strong effect. The herd effects might have remained low in the Gambia where mixing of the intervention and control populations was intense, which would have led to lower power compared to an individually randomised trial of the same size. Therefore, the authors decided to choose the individually randomised trial design.

2.7.4 MODELLING VACCINE IMPACT AGAINST PNEUMOCOCCAL INFECTION

Longitudinal data with repeated measurements on colonisation have been used to study the process of pneumococcal carriage to improve understanding the pneumococcal diversity in unvaccinated populations [135–138]. To estimate hazards of acquisition and clearance, as well as the strength of competitive interactions between serotypes, Markov transition models have often been used. The model describes transitions between the uncolonised as well as singly or multiply colonised states.

Population-level compartmental models of pneumococcal transmission have also been used [139–141]. Compartmentalisation typically reflects health states relevant for transmission (susceptible, infectious, recovered) or population (age classes, social mixing groups) and the aim is to track changes in compartments without specifying which individuals are involved [142].

Nurhonen et al. [143] employed an individual-based simulation model considering also the age-specific contact structure of the population to explore the extent of indirect effects after large-scale vaccinations with PCVs. Similar

models have been used to explore the potential impact of catch-up campaigns in Vietnam and in Kenya, suggesting that additional reductions in IPD of any serotype can be obtained soon after PCV implementation [144,145]. In these types of models, vaccine efficacy parameters are used as inputs. It is therefore essential for coherent predictions to model the vaccine response correctly at the individual level. For example, population-level models have invariably been based on the assumption of leaky mode of vaccine action, but usually without empirical evidence. However, Mehtälä et al. [14] used Markov transition models to study the correct estimation of vaccine efficacy from longitudinal data on colonisation and concluded that the plausible mode of action in the context of pneumococcal conjugate vaccines is leaky protection.

Because any model soon becomes difficult to define or computationally intractable if the number of parameters to be estimated increases, pneumococcal serotypes have usually been treated as groups (e.g. vaccine- and non-vaccine types). However, some studies have payed special attention to the differences between serotypes. For example van Effelterre et al. [146] included 18 serotypes responsible for most of the IPD in the US children as well as serotype-specific susceptibility to antibiotics into their model predicting the impact of PCV7.

Another approach that has been used to predict post-vaccination disease patterns are so called pseudo-dynamic models. They utilise serotype-specific carriage data and estimates of invasiveness as the case-to-carrier ratios (probabilities of disease per carriage episode) without explixitly modelling pneumococcal transmission. Weinberger et al. [45] evaluated which serotypes are the most likely to increase in disease following vaccination, and Flasche et al. [147] predicted the total impact of PCV7 on IPD in children, using data of 13 sites in 8 countries. A similar method by Nurhonen et al. [51] assumed partial or complete elimination of vaccine-type colonisation, complete replacement by non-vaccine-type colonisation and stable case-to-carrier ratios. The authors presented a sequential algorithm for the identification of the most optimal additional serotypes to current or prospective vaccine formulations.

2.8 VACCINE PREVENTABLE DISEASE INCIDENCE

In this review, all effect measures of vaccination as presented thus far have been relative. This means that they have been based on the ratios of the risks or hazards of infection, comparing the vaccinated and unvaccinated. The difference in the incidence of infection between vaccinated and unvaccinated subjects (incidence rate difference, IRD), however, is also important as it represents the vaccine preventable disease incidence (VPDI). The relative reduction in the incidence due to vaccination may be small but if the baseline incidence is high, the vaccine-preventable disease burden may still be

considerable. Therefore, VPDI is a useful measure of the public health importance of vaccination. [148,149]

If the measurement of the disease outcome is not specific so that false positive cases are possible (e.g. in case of pneumonia), the absolute rate difference may provide a more accurate measure of the true effect of the vaccine as the false positive incidence cancels out in calculating differences. Denote by a the incidence of the actual infection of interest (e.g. pneumococcal pneumonia) and by b the incidence of false positive infection (pneumonia due to other causes). Estimating the absolute rate difference and thus VPDI allows cancelling the false positive incidence: $\text{VPDI} = (a^U + b) - (a^V + b) = a^U - a^V$, where superscripts V and U denote the vaccinated and unvaccinated, respectively. The rate ratio would not lead to such cancelling of rate b.

Separate estimates of vaccine-preventable disease burden can be calculated for different disease manifestations as well as against health-system endpoints such as health-care visits or drug use. In the FinIP trial, the full disease burden caused by pneumococcus was estimated by assessing the VPDI of IPD, clinically suspected but non-laboratory confirmed IPD, hospital-diagnosed pneumonia, tympanostomy tube placements and antimicrobial purchases [150]. It was shown that over 95% of the reduction in the total number of disease episodes in vaccinated children were due to mild upper respiratory infections.

In the CAPiTA trial, VPDI among adults ≥65 years of age in the Netherlands was assessed for a number of pneumococcal disease outcomes, such as (vaccine-type) IPD, community acquired and pneumococcal pneumonia and death [151]. Similarly to the FinIP trial, the relative reductions of the more sensitive, clinically defined outcomes were smaller than those of the more specific etiologically confirmed outcomes. For example, the relative reductions in clinical community-acquired pneumonia (CAP) and vaccine-type IPD were 8% and 72%, respectively. By contrast, VPDI of CAP (121 cases per 100 000 person-years) was much higher that that of vaccine-type IPD (6 cases per 100 000 person-years).

In both trials, the number needed to vaccinate (NNV) was calculated as the inverse of VPDI (1/VPDI). NNV with three or four doses of PCV10 was 671 to prevent one episode of laboratory-confirmed IPD, but only 4 to prevent one outcome episode (any of those) during the two-year follow-up [150]. In the CAPiTA trial, NNV with one dose of PCV13 was 3411 to prevent one case of vaccine-type IPD, but 165 to prevent one case of CAP [151].

3 AIMS OF THE STUDY

The purpose of this study is to define the effect measures of vaccines and vaccination in the context of pneumococcal colonisation and disease, and to estimate and interprete different vaccine effects. The detailed objectives are as follows:

1. To define an estimand of type-specific and all-type vaccine efficacy against acquisition of pneumococcal colonisation and an appropriate but practically feasible estimator that allows its estimation from cross-sectional data.

2. To improve the above estimand by conditioning on time at-risk to acquiring the target strains and thereby to better describe the biological efficacy of a vaccine against pneumococcal acquisition.

3. To compare three observational study designs in the estimation of vaccine effectiveness against invasive pneumococcal disease in Finnish children eight years into the vaccination programme.

4. To build a pseudo-dynamic model to describe and compare the expected behaviour of the vaccine impact and effectiveness against invasive pneunomococcal disease during the post-introduction period.

4 METHODS FOR ESTIMATING VACCINE EFFICACY, EFFECTIVENESS AND IMPACT AGAINST PNEUMOCOCCAL INFECTION

This chapter covers the methods developed in this thesis, as well as the corresponding statistical models. Chapter 5 covers the corresponding empirical results.

4.1 VACCINE EFFICACY AGAINST PNEUMOCOCCAL COLONISATION (I, II)

4.1.1 DEFINITION OF THE ESTIMANDS AND ESTIMATORS OF VACCINE EFFICACY

In contrast to the well-standardised methods of measuring vaccine efficacy against IPD, there is wide variation in reporting vaccine effects on colonisation. In many studies the vaccine efficacy against colonisation has been estimated based on prevalence ratios and prevalence odds ratios [47,48,72,73,152]. Nevertheless, at least ideally, the definition and measurement of vaccine efficacy should be based on the actual mechanism of the vaccine's protective action [62]. The most rational way to approach vaccine efficacy against pneumococcal colonisation appears to be through hazards (i.e. per capita rates) of acquisition.

In order to estimate vaccine efficacy against acquisition of pneumococcal colonisation, longitudinal samples, meaning repeated observations of the current status of colonisation in the study subjects, are ideally needed. Moreover, as the exact times of acquisition cannot be observed, statistical modelling would be needed to capture the dynamics of multiple acquisition and clearance events. However, collecting frequently sampled data is laborious, expensive, and unpleasant for the study participants. Therefore, our interest was to define a more feasible *estimator* that allows using cross-sectional data.

We defined the *estimand* of vaccine efficacy against acquisition based on acquisitions of a target strain *i* as

$$\mathrm{VE}_{\mathrm{acq,i}} = 1 - \frac{q_i^{(T)}}{q_i^{(C)}}, \qquad (1)$$

where $q_i^{(T)}$ and $q_i^{(C)}$ are the (constant) hazards of acquiring strain *i* colonisation in the vaccinated and unvaccinated, respectively. The target strain can be a vaccine type, non-vaccine type or a group of them. If one assumes that the serotype distribution of colonisation is stable (stationary) over time and the duration of strain i carriage episode is not affected by vaccination, the following epidemiological relationship between the odds of colonisation, hazard and mean duration D_i^j holds separately for the vaccinated (T) and unvaccinated (C) [57]:

$$\frac{p_i^{(j)}}{1-p_i^{(j)}} = q_i^j D_i^j,$$

where $p_i^{(j)}$ is the prevalence of serotype *i* and *j*= *T, C*. The estimand of vaccine efficacy $\mathrm{VE}_{\mathrm{acq,i}}$ can now be obtained as a ratio of the odds of vaccination in those colonised with strain *i* versus those not colonised (see Study I equation (2)):

$$\mathrm{VE}_{\mathrm{acq,i}}^{(\mathrm{odds})} = 1 - \frac{D_i q_i^{(T)}}{D_i q_i^{(C)}} = 1 - \frac{p_i^{(T)}/(1-p_i^{(T)})}{p_i^{(C)}/(1-p_i^{(C)})}. \qquad (2)$$

The prevalences p_i are directly estimable from cross-sectional data and the equation (2) therefore provides an *estimator* of $\mathrm{VE}_{\mathrm{acq,i}}$. Estimation of vaccine efficacy against acquisition of colonisation is thus possible using standard methods for estimating odds ratios and without long follow-up of study participants. Note that the expression (2) does not depend on the duration of colonisation.

The approach of cross-sectional measurement of pneumococcal acquisition has a close connection to risk-set sampling in nested case-control studies, where controls are selected from those at risk at the time of case occurrence [153]. In this thesis, the method is extended from traditional nested case-control studies into two directions as subjects may be at risk for acquisition of several competing serotypes and experience recurrent colonisation of the same serotype. Then, the at-risk set corresponds to all states in a multi-state model from which there are transitions to the target states [154].

Estimand (2) does not take into account the within-host competition between serotypes that occurs because carriers may be additionally protected against acquisition of another strain [7,15]. In order to adjust for the different times that the vaccinated and unvaccinated spend at risk for acquiring a new episode of the target-type colonisation, we followed the ideas of at-risk set and multi-state model in Study II. **Figure 2** presents two models of colonisation with two strains and three or four states of colonisation. Both models can be generalised to any number of states. We applied the models separately for vaccinated and unvaccinated children. Strains 1 and 2 represent vaccine and

non-vaccine strains, respectively. The first model (A) allows transitions to double colonisation whereas the second model (B) does not.

In Study II, we defined a modified *estimand* of vaccine efficacy, again defined through a hazard ratio, but now conditioning this parameter on those carriage states against which the vaccine has no direct biological effect (i.e. the states of empty nasopharynx or colonisations with any of the non-vaccine serotypes). In the simplest form, the estimand can be specified through conditioning on the uncolonised state only:

$$\mathrm{VE}_{\mathrm{acq,i}} = 1 - q_{0,i}^{T}/q_{0,i}^{C}. \qquad (3)$$

Here, the sub-indices express the transitions from the empty nasopharynx (0) to the target strain i. The expression (3) corresponds to the relative reduction in the hazard of acquiring the target strain i when comparing an uncolonised vaccinated individual to an uncolonised unvaccinated individual.

The estimand of vaccine efficacy against all target strains can be defined by generalising the estimand (3). We denote the predefined target set of vaccine serotypes as W and write the estimand as follows (see Study II, equation (1)):

$$\mathrm{VE}_{\mathrm{W|0}} = 1 - \frac{\sum_{i\in W} q_{0,i}^{T}}{\sum_{i\in W} q_{0,i}^{C}} \qquad (4)$$

Correspondingly, vaccine efficacy can be defined by including in the target set W any group of strains (e.g. the non-vaccine strains).

A more comprehensive estimand can be defined by including in the at-risk set all those states against which the vaccine does not confer direct biological protection (denoted as $\bar{V}_0$):

$$\mathrm{VE}_{\mathrm{W|\bar{V}_0}} = 1 - \frac{\sum_{i\in W} q_{\bar{V}_0,i}^{T}}{\sum_{i\in W} q_{\bar{V}_0,i}^{C}}. \qquad (5)$$

Here, the sub-indices express transitions from the larger risk-set ($\bar{V}_0$) to the target strain $i \in W$. However, if the interest is in the estimation of vaccine efficacy against the non-vaccine types, the reference set should comprise only the uncolonised state.

These conditional estimands (4) and (5) have a clear interpretation as measures of the direct protective effect of the vaccine on an individual.

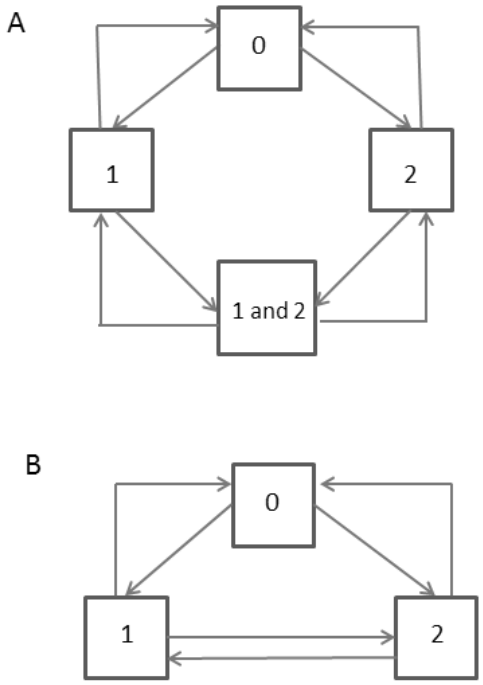

Figure 2 *A: Model A of colonisation. The model has two strains and four states of colonisation. The states are uncolonised (0), colonised with strain 1, colonised with strain 2 and colonised with both strains (1 and 2). The model is governed by eight transition hazards. B: Model B of colonisation. The model has two strains and three states of colonisation. The states are uncolonised (0), colonised with strain 1 and colonised with strain 2. The model is governed by six transition hazards. Both models are considered separately for the vaccinated and unvaccinated.*

After applying the properties of reversibility and stationarity to the underlying processes of colonisation in the vaccinees and controls (Study II, supplementary material), the estimand (4) can be expressed in terms of the stationary probabilities of colonisation as follows (Study II, equation (4)):

$$\mathrm{VE}_{\mathrm{W}|0} = 1 - \frac{p_W^{(T)}/p_0^{(T)}}{p_W^{(C)}/p_0^{(C)}}, \tag{6}$$

These stationary propabilities are directly estimable from cross-sectional data when the serotype distribution of colonisation is stationary. Therefore, the equation (6) also provides an *estimator* $\widehat{\mathrm{VE}}_{\mathrm{W}|0}$ of vaccine efficacy against the target state(s).

We define a more comprehensive estimator by including all states in $\bar{V}_0$ in the denominator (Study II, equation (6)):

$$\widehat{\mathrm{VE}}_{\mathrm{W}|\bar{\mathrm{V}}_0} = 1 - \frac{p_W^{(T)}/p_{\bar{V}_0}^{(T)}}{p_W^{(C)}/p_{\bar{V}_0}^{(C)}}, \quad (7)$$

The use of the estimator (7) may increase the precision in the estimation of vaccine efficacy compared to (6).

The conditional estimands account for changes in susceptibility to acquisition within an individual (within-host replacement). However, some conditions need to be fulfilled. The serotype distribution should be stationary in the sense that the statistical properties of the carriage processes in the vaccinated and unvaccinated do not change over time. The between-strain competition needs to be symmetrical in terms of relative hazards and the mean durations of colonisation in the vaccinated and unvaccinated should be equal. The last two assumptions will be discussed in more detail in Section 5.1.1.

4.1.2 APPLICATION TO EMPIRICAL DATA AND THE SIMULATION STUDY

Longitudinal data on pneumococcal carriage in Israeli toddlers was used to validate the estimators (2), (6) and (7) of vaccine efficacy. The children were vaccinated with a 9-valent pneumococcal conjugate vaccine (N=132) or a meningococcus group C conjugate vaccine as controls (N=130) and followed by monthly nasopharyngeal samples for a year. The children were aged 12-43 months at enrolment and attended eight day-care centres. Serotype-specific episodes of carriage were defined as series of isolates of a homologous (same) serotype at consecutive visits. Altogether 772 carriage episodes of the most common serotypes/groups were identified and defined as acquisition. The data have been described in more detail in Dagan et al. [47] and Study I.

Data from all 12 cross-sections were used to gain enough samples to obtain serotype-specific estimates of vaccine efficacy. Logistic regression was used to estimate $\mathrm{VE}_{\mathrm{acq}}^{(\mathrm{odds})}$, $\widehat{\mathrm{VE}}_{\mathrm{W}|0}$ and $\widehat{\mathrm{VE}}_{\mathrm{W}|\bar{\mathrm{V}}_0}$. When comparing the performance of different cross-sectional estimators, two other estimators, $\mathrm{VE}_{\mathrm{prev}}$ and $\mathrm{VE}_{\mathrm{acq}}^{(\mathrm{rate})}$, were also used. $\mathrm{VE}_{\mathrm{prev}}$ is the vaccine efficacy based on the ratio of serotype i prevalences, $\mathrm{VE}_{\mathrm{prev}} = 1 - p_i^{(T)}/p_i^{(C)}$ (see Study I, Section 2.3). $\mathrm{VE}_{\mathrm{acq}}^{(\mathrm{rate})}$ was estimated with Cox proportional hazards model using time since vaccination as the time-variable. To account for the dependence between a child's samples, generalized estimating equations (GEE) were used in logistic regression models, and including an individual frailty into the Cox regression model. R-functions gee() with an exchangeable correlation structure and coxph() with a cluster term to allow between-subject heterogeneity were used [155]. The

analyses were adjusted for age-group and the presence of the target serotype in the day care centre at the time of the sample.

We constructed a simulation study with four vaccine and five non-vaccine strains to study the sensitivity of the cross-sectional estimates to departures from the model assumptions. We chose the hazards of colonisation and clearance so that the stationary strain distribution mimics that typical of serotypes of *S. pneumoniae*, with prevalence ranging from approximately 20% for the most common type to <2% for rare types. Repetitions of samples in a cohort of 1000 vaccinated and 1000 unvaccinated children were taken at the stationary phase of the process. In these simulations, odds ratios for vaccine efficacy estimates were estimated through logistic regression and using R-function glm().

4.2 VACCINE EFFECTIVENESS AND IMPACT AGAINST INVASIVE PNEUMOCOCCAL DISEASE (III, IV)

4.2.1 MODEL (IV)

The vaccine impact (VI) and vaccine effectiveness (VE) against pneumococcal carriage and disease may vary over time because of the changing population dynamics of pneumococcal carriage. In order to explicitly describe the expected behaviour of VI and VE during the post-introduction period, we constructed a simple pseudo-dynamic model to follow the incidences of pneumococcal carriage and disease in cohorts of children until vaccine-type (VT) disease becomes eliminated and a new steady-state is reached.

In the model, we described the indirect effects of vaccination by assuming that the proportion *f(t)* of episodes of vaccine-type carriage of all carriage episodes decreases over time *t* due to reduced VT transmission and will immediately be replaced by a corresponding hazard of non-vaccine type (NVT) acquisition. This was achieved by assuming that the per capita hazard of carriage acquisition (force of infection) C is constant over time irrespective of the child's age and vaccination status. The forces of VT carriage acquisition at time t are then $f(t)C$ and $f(t)C(1-\mathrm{VE}_{\mathrm{acq}})$ in the unvaccinated and vaccinated children, respectively, where $\mathrm{VE}_{\mathrm{acq}}$ is the leaky vaccine efficacy against carriage acquisition. (Note that in Study IV $\mathrm{VE}_{\mathrm{acq}}$ is denoted as $\mathrm{VE}_{\mathrm{col}}$. Note also that C denotes here the force of infection while it was used to denote "control" elsewhere in this summary.)

The hazards of IPD follow by multiplying the hazards of carriage acquisition with the corresponding case-to-carrier ratios, assumed to remain constant during the study period. The hazards of IPD depend on the time-dependent proportion of VT carriage (*f(t)*), the pre-vaccination hazards of VT and NVT IPD and the vaccine efficacies against acquisition and progression from

carriage to disease (VE_{acq}, VE_{inv}), respectively. Once the hazards of IPD had been defined for the two serotype categories (VT and NVT) and vaccinated and unvaccinated individuals, we derived the expressions of vaccine impact against VT, NVT and all IPD by comparing the hazards in the post-introduction period to those in the pre-vaccination period. Correspondingly, we derived the expressions of vaccine effectiveness for the three serotype groups by comparing hazards in the vaccinated to those in the unvaccinated in the post-introduction period. The detailed derivations of the expressions of vaccine impact and effectiveness are presented in Study IV.

In this thesis, vaccine effectiveness and the total vaccine impact, i.e. the parameters that describe vaccine effects on vaccinated individuals, are compared. Vaccine effectiveness is a measure of the benefit of becoming vaccinated during an ongoing vaccination programme. The total vaccine impact accounts for both direct vaccine protection and the indirect effects of the vaccination programme. **Table 1** presents the total vaccine impact and vaccine effectiveness against VT, NVT and all IPD, i.e. VI_{tot}^{VT} versus VE^{VT}, VI_{tot}^{NVT} versus VE^{NVT} and $VI_{tot}^{all\ IPD}$ versus $VE^{all\ IPD}$. The corresponding parameters of indirect and overall impact are presented in Tables 3 and 4 of Study IV.

Note that in Study IV on page 3 (left column), there is a typo in the equation describing VE against VT. The equation should read as follows: $VE^{VT} = 1 - (1 - VE_{col})(1 - VE_{inv})$.

Table 1 *The total impact of a vaccination programme against VT, NVT and all IPD in a before-after study setting and vaccine effectiveness against VT, NVT and all IPD in a parallel study setting. Modified from Tables 3 and 4 of Study IV.*

Serotype group	Total impact	Vaccine effectiveness
VT	$\mathrm{VI}_{\mathrm{tot}}^{\mathrm{VT}} = 1 - \frac{f(t)}{f(0)}(1 - \mathrm{VE}_{\mathrm{col}})(1 - \mathrm{VE}_{\mathrm{inv}})$	$\mathrm{VE}^{\mathrm{VT}} = 1 - (1 - \mathrm{VE}_{\mathrm{col}})(1 - \mathrm{VE}_{\mathrm{inv}})$
NVT	$\mathrm{VI}_{\mathrm{tot}}^{\mathrm{NVT}} = 1 - \frac{1 - f(t)(1 - \mathrm{VE}_{\mathrm{col}})}{1 - f(0)}$	$\mathrm{VE}^{\mathrm{NVT}} = 1 - \frac{1 - f(t)(1 - \mathrm{VE}_{\mathrm{col}})}{1 - f(t)}$
All IPD	$\mathrm{VI}_{\mathrm{tot}}^{\mathrm{all\ IPD}} = 1 - \frac{(1 - \mathrm{VI}_{\mathrm{tot}}^{\mathrm{VT}})\mathrm{IPD}_0^{\mathrm{VT}} + (1 - \mathrm{VI}_{\mathrm{tot}}^{\mathrm{NVT}})\mathrm{IPD}_0^{\mathrm{NVT}}}{\mathrm{IPD}_0^{\mathrm{VT}} + \mathrm{IPD}_0^{\mathrm{NVT}}}$	$\mathrm{VE}^{\mathrm{all\ IPD}} = 1 - \frac{\frac{f(t)}{f(0)}(1 - \mathrm{VE}^{\mathrm{VT}})\mathrm{IPD}_0^{\mathrm{VT}} + \frac{1 - f(t)}{1 - f(0)}(1 - \mathrm{VE}^{\mathrm{NVT}})\mathrm{IPD}_0^{\mathrm{NVT}}}{\frac{f(t)}{f(0)}\mathrm{IPD}_0^{\mathrm{VT}} + \frac{1 - f(t)}{1 - f(0)}(1 - VE^{NVT})\mathrm{IPD}_0^{\mathrm{NVT}}}$

4.2.2 APPLICATION TO EMPIRICAL DATA

Setting

In Finland, PCV10 was selected on the basis of a public tender and was introduced into the national vaccination program in September 2010. Children born on June 2010 or later have been eligible to be immunised with a 3-dose schedule at 3, 5, and 12 months (2+1). No catch-up vaccinations were offered for older children. There was no previous use of pneumococcal conjugate vaccines in the Finnish national vaccination program except for rare risk groups, and private use of PCV7 before the national vaccination program was estimated to be <2% on the basis of national sales data. After introduction to the national vaccination program, the uptake of the first dose of PCV10 rose quickly over 90% and was estimated at 96% in the 2017 birth cohort [156].

Use of PCV13 and the 23-valent pneumococcal polysaccharide vaccine (PPSV23) in adults at risk and the elderly is currently recommended. However, there is no national pneumococcal adult vaccination programme, and the cumulative adult coverage for PPSV23 and PCV13 in 2018 was 3% and 9%, respectively, based on the National Vaccination Registry [156].

To demonstrate the estimation of PCV10 effectiveness with cohort, nested case-control and indirect cohort designs (Study III), we defined the study cohort of vaccine-eligible children based on the Finnish Population Information System. All children were born in or after 6/2010 and followed from 1/2011 until 12/2018 (aged 6–102 months).

In Study IV, we used cohort design to estimate vaccine effectiveness in the vaccine-eligible children until the end of 2016 (aged 6–78 months). For the impact analysis, a reference cohort was selected from the time-period predating PCV10 introduction, including children born between 6/2002 and 6/2008 and followed from 1/2003 through 12/2008. In order to have cohorts of equal size, the target cohort was smaller than in Study III.

Register data

We retrieved the vaccination status of each individual child in the vaccine-eligible cohorts from the National Vaccination Register. The child was defined as vaccinated if at least one dose of PCV10 was registered. Cases of IPD were identified from the National Infectious Diseases Register, serotyped at the THL reference laboratory, and categorised according to the causative serotype into three mutually exclusive groups: PCV10 serotypes (1, 4, 5, 6B, 7F, 9V, 14, 18C, 19F, 23F), PCV10-related serotypes that belong to the same serogroups

as the PCV10 serotypes, and non-PCV10-related serotypes. All register-based information was linked by using the unique national personal identity code.

Empirical data and statistical methods

In Study III, VE was estimated using three different study designs. In the cohort design, all children in the study cohort were included and followed until IPD, death or end of study period. VE was estimated with Poisson regression adjusted by age group (6–23, 24–47, 48–102 months), sex and calendar year. In the nested case-control design, five controls for each IPD case were selected from the case's risk set, matching with age, sex and calendar year. The risk set included all children in the study cohort who were at risk of IPD at the time of case occurrence. Conditional logistic regression was used to estimate VE. In the indirect cohort design, VE was estimated as the odds of vaccination in IPD cases of the target serotype or serotype group compared to the odds of vaccination in non-PCV10-related cases. VE was estimated with logistic regression adjusted by age group, sex and calendar year. To assess any time trends in VE, the study period was split into two parts: years 2011–2014 (early period) and 2015–2018 (latter period).

In Study IV, we estimated VE using the cohort design. Poisson regression was used to compare the incidence rates between the vaccinated and unvaccinated children, and the person-times of the cohorts were used as offsets. The total, indirect and overall impact of the vaccination programme were estimated by comparing IPD rates between the target and reference cohorts.

It seemed obvious that the small numbers of cases in the data precluded appropriate use of asymptotic frequentist methods. Therefore, statistical inferences were performed within the Bayesian framework in both Studies III and IV. Uninformative prior distributions were used, i.e. a normal distribution with mean 0 and variance 10^6 for the effect parameters and a Gamma distribution with mean 1 and intensity 10^5 for precision. Results are presented as point estimates (posterior means) and 95% posterior probability (credible) intervals (CI). All analyses were carried out with R-library INLA [155,157]. The methodology of Bayesian inference using integrated Laplace approximations is described in Rue et al. [158].

In both Studies III and IV, we estimated VE as a time-average of the time-varying effectiveness measure over the defined time-period (years 2011-2014 or 2015-2018 in Study III, and years 2011-2016 in Study IV). In addition, we estimated VI both as a time-average and separately for each year, comparing to the whole reference cohort (Study IV). In **Figure 3**, the estimation of VI and VE in vaccine-eligible children, as well as vaccine efficacy in a hypothetical individually randomised vaccine trial are presented in the Finnish setting.

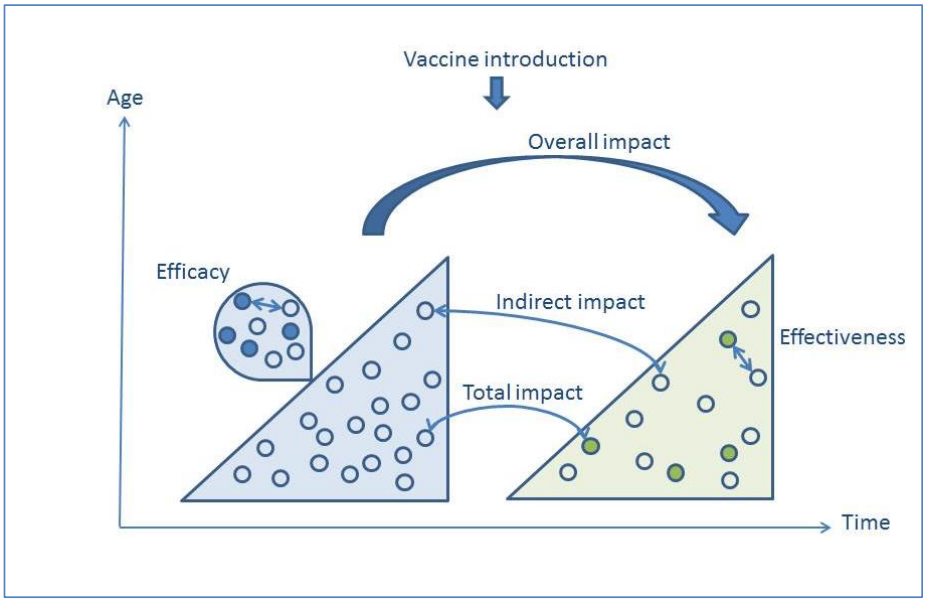

Figure 3 *Vaccine efficacy, effectiveness and impact. The cohort of vaccine-eligible children is presented as a light green triangle and the unvaccinated reference cohort as a light blue triangle. Vaccinated and unvaccinated IPD cases are presented with filled and unfilled circles, respectively. Vaccine impact (total, indirect and overall) is estimated by comparing the target and reference cohorts. Vaccine effectiveness is estimated within the target cohort. A (hypothetical) trial to estimate vaccine efficacy is presented in terms of a sample of the unvaccinated reference cohort, in which half of the participants would be vaccinated.*

5 RESULTS

5.1 VACCINE EFFICACY AGAINST PNEUMOCOCCAL COLONISATION (I, II)

5.1.1 PERFORMANCE OF THE ESTIMATORS OF VACCINE EFFICACY AND THEIR APPLICATION TO EMPIRICAL DATA (I, II)

When the three cross-sectional estimators of vaccine efficacy ($\mathrm{VE}_{\mathrm{acq}}^{(\mathrm{odds})}$, $\widehat{\mathrm{VE}}_{\mathrm{W}|0}$ and $\widehat{\mathrm{VE}}_{\mathrm{W}|\bar{\mathrm{v}}_0}$ were applied to data on pneumococcal carriage in Israeli toddlers, the point estimates corresponded to each other and to $VE_{acq}^{(rate)}$ relatively well (**Table 2**). The estimators were also compared to the vaccine efficacy based on prevalence ratio, $\mathrm{VE}_{\mathrm{prev}}$, which often provided the lowest estimates. The 9-valent pneumococcal conjugate vaccine was found to be efficacious against acquisition of serotypes 6B, 9V and 23F. The efficacy against all vaccine serotypes based on the odds-ratio based estimators varied from 0.39 to 0.43 (95% CI 0.17...0.55). The estimated efficacy against the non-vaccine types was clearly different from zero only when using $\mathrm{VE}_{\mathrm{acq}}^{(\mathrm{odds})}$ (-0.28; 95% CI -0.59... -0.03). This was expected because using $\mathrm{VE}_{\mathrm{acq}}^{(\mathrm{odds})}$ does not account for the apparent within-host replacement by the NVT strains. The precision of the estimation was higher with $\widehat{\mathrm{VE}}_{\mathrm{W}|\bar{\mathrm{v}}_0}$ than with $\widehat{\mathrm{VE}}_{\mathrm{W}|0}$ because the former estimator incorporates more data in the estimation.

The performance of the cross-sectional estimators of vaccine efficacy was considered also in terms of departures from the model assumptions (Study II, Section 6). When vaccine efficacy is defined for a single strain, two assumptions need to be fulfilled. The between-strain competition needs to be symmetrical in terms of relative hazards, and the hazards of clearance of all strains need to be equal. However, if the estimation of vaccine efficacy is considered for the aggregate state of all vaccine strains, the assumption of symmetrical competition is not needed. In this case, the only condition required for cross-sectional estimation is the similarity of clearance hazards of the aggregate states in the vaccinees and controls. Moreover, the aggregate conditional estimands are coherent in the sense that they are the weighted averages of the serotype-specific estimands and therefore fully agree with any strain-specific efficacy.

A simulation study with four vaccine and five non-vaccine strains was performed to study the assumption of equal durations (Study II). If the vaccine affects the hazard of clearance of vaccine strains, the mean of the cross-sectional estimator $\widehat{\mathrm{VE}}_{\mathrm{W}|\bar{\mathrm{v}}_0}$ was close to the combined efficacy $1-(1-\widehat{\mathrm{VE}}_{\mathrm{W}|\bar{\mathrm{v}}_0})(1-\mathrm{VE}_{\mathrm{D}})$ where VE_{D} is one minus the ratio of the mean duration of colonisation in

vaccinees to that in controls (data not shown). The estimate based on $\widehat{\mathrm{VE}}_{W|\bar{v}_0}$ can thus be interpreted more generally as a summary measure of the effect of being vaccinated on susceptibility to acquisition and on duration of colonisation. If individuals are infectious throughout the period of colonisation, this quantity is the relative reduction in the transmission potential of the strain in question, afforded by the vaccine.

Table 2 *Estimates of serotype-specific vaccine efficacy against carriage acquisition in Israeli toddlers, based on prevalence ratio (VE_{prev},), odds-ratio based estimators ($VE_{acq}^{(odds)}$, $\widehat{VE}_{W|0}$ and $\widehat{VE}_{W|\bar{V}_0}$) and rate ratio ($VE_{acq}^{(rate)}$), with respective 95% confidence intervals. Estimates of $VE_{acq}^{(odds)}$ and $VE_{acq}^{(rate)}$ are presented in Table 2 of Study I.*

Serotype/ group	$\mathbf{VE_{prev}}$	$\mathbf{VE_{acq}^{(odds)}}$	$\widehat{\mathbf{VE}}_{\mathbf{W|0}}$	$\widehat{\mathbf{VE}}_{\mathbf{W|}\bar{V}_0}$	$\boldsymbol{VE_{acq}^{(rate)}}$
6B	0.68	0.76 (0.38...0.91)	0.82 (0.52...0.93)	0.78 (0.47...0.90)	0.85 (0.45...0.96)
9V	0.59	0.60 (0.18...0.8)	0.68 (0.26...0.87)	0.67 (0.28...0.84)	0.60 (0.13...0.82)
14	0.23	0.30 (-0.30...0.62)	0.10 (-0.75...0.54)	0.37 (-0.12...0.64)	0.18 (-0.66...0.60)
19F	0.1	0.08 (-0.47...0.43)	-0.01 (-0.63...0.38)	0.06 (-0.43...0.37)	-0.00 (-0.92...0.48)
23F	0.34	0.39 (0.06...0.60)	0.46 (0.15...0.66)	0.48 (0.23...0.65)	0.54 (0.30...0.69)
6A	0.23	0.21 (-0.15...0.45)	0.08 (-0.36...0.38)	0.08 (-0.36...0.38)	0.25 (-0.11...0.49)
15	-0.53	-0.63 (-1.37...-0.12)	-0.27 (-0.97...0.18)	-0.27 (-0.97...0.18)	-0.32 (-0.86...0.06)
19A	0.01	-0.04 (-1.03...0.47)	0.01 (-0.96...0.50)	0.01 (-0.96...0.50)	-0.00 (-0.92...0.48)
VT	0.35	0.42 (0.24...0.56)	0.39 (0.17...0.55)	0.43 (0.28...0.55)	0.46 (0.30...0.58)
NVT	-0.14	-0.28 (-0.59...-0.03)	-0.10 (-0.38...0.13)	-0.10 (-0.38...0.13)	-0.14 (-0.37...0.04)

5.2 VACCINE EFFECTIVENESS AND IMPACT AGAINST INVASIVE PNEUMOCOCCAL DISEASE (III, IV)

5.2.1 COMPARISON OF TIME TRENDS IN VACCINE EFFECTIVENESS AND VACCINE IMPACT (IV)

We explored the time-related behavior of the parameters of vaccine effectiveness (VE) and vaccine impact (VI) under a scenario that corresponds to vaccination of children. This means a high vaccination coverage of 90% (suitable to Finland), moderate efficacy against carriage acquisition (VE_{acq}=50%) and high efficacy against progression to disease (VE_{inv}=90%). We assumed that the proportion of vaccine type carriage *f(t)* decreases from 60% to zero in six years. The initial proportion of NVT cases out of all IPD (20%) was set similar to the situation in Finnish children in the pre-vaccination era. Here, we summarise the results of the above scenario. Another scenario with moderate vaccination coverage 50% and efficacy against disease progression VE_{inv}=50% is reported in Study IV.

The time evolution of the parameters of vaccine effectiveness and impact are graphically displayed in **Figure 4** according to the expressions presented in **Table 1**. The comparison of VE and VI is made for the vaccinated individuals, i.e. with VE (as a measure of the benefit of becoming vaccinated during an ongoing vaccination programme) and VI_{tot} (as a measure of both the direct vaccine-induced protection and the indirect effects of the vaccination programme). Other comparisons for the unvaccinated and all individuals are presented in Study IV. All trends described in this section are expected based on the model.

As exposure to VT infection decreases eventually to zero, the impact against VT IPD (VI_{tot}^{VT}) increases to 100%. VI_{tot}^{VT} is affected by the direct and indirect effects and quantifies the net benefit of the vaccination programme to a vaccinated individual. By contrast, because exposure to NVT infection increases over time, VI_{tot}^{NVT} is always negative and decreases over time.

Although the incidence of VT IPD decreases in vaccinated children due to the direct and indirect protection (herd effects), the ratio of the incidence rate in the vaccinated to that in the unvaccinated, and thus VE^{VT}, is expected to remain stable at any time since the programme onset. This follows from the assumption that both VE_{acq} and VE_{inv} are constant irrespective of the dynamics of the indirect effects. The assumption of constant VE_{acq} may be a simplification because it discards within-host competition between serotypes. As the vaccination programme enhances NVT acquisition among the vaccinated, VE^{VT} could increase over time (cf. Study II). This effect, however,

should be small, especially when VT incidence is very low. The time-independence of VE^{VT} can also be seen in **Table 1**.

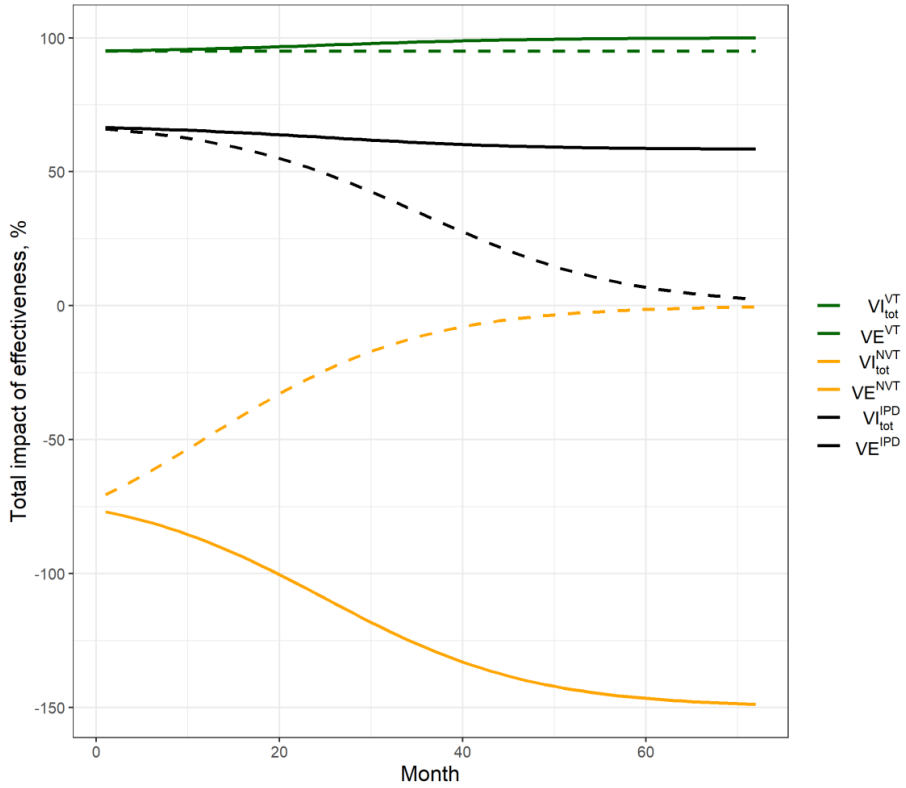

Figure 4 *Vaccine effectiveness and total impact as functions of time since vaccination onset (Study IV). Parameter values:* VE_{col}*=50%,* VE_{inv}*=90%, f(0)=60% and the initial fraction of NVT IPD out of all IPD 20%. VT carriage decreases to 0% in about 70 months. The lines represent the total impact (solid line) and effectiveness (dashed line) against VT (green), NVT (orange) and all IPD (black). The expressions of vaccine effectiveness and impact are presented in Table 1.*

The total impact of the programme on vaccinated individuals, VI_{tot}^{VT}, is always higher than VE^{VT}, as it accounts also for the indirect effects of vaccination. However, if VE_{acq} and VE_{inv} are high, the difference between VI_{tot}^{VT} and VE^{VT} is not notable.

The vaccine effectiveness against NVT IPD (VE^{NVT}) describes the within-host competition between the vaccinated and unvaccinated. Based on the model, VE^{NVT} increases from a negative level and approaches zero when vaccine-type carriage has been eliminated and the vaccinated and unvaccinated children

carry non-vaccine types equally. Eventually, there is no within-host competition between vaccine and non-vaccine serotypes. Also $VE^{all\ IPD}$ should approach zero as eventually all disease is caused by the non-vaccine types.

5.2.2 APPLICATION TO EMPIRICAL DATA

We estimated vaccine effectiveness (VE) within the target cohort of vaccine-eligible children in Finland in Studies III and IV. The study periods were from 2010 to 2016 (Study III) and until the end of 2018 (Study IV). Time trends in VE estimates were assessed in Study III. In Study IV, we estimated the total, indirect and overall vaccine impact (VI) both as time average and separately for each year to assess any time trends. A schematic representation of the estimation of vaccine effectiveness and impact among vaccine-eligible children in a setting that corresponds to Finland is presented in **Figure 3**.

In both studies III and IV, the estimates of vaccine effectiveness and impact against the vaccine types broadly corresponded to those of the theoretical model of Study IV. While the incidence of PCV10 serotypes decreased substantially during the follow-up, VI_{tot}^{VT} was higher than VE^{VT} and increased over years reaching 100% in six years. VE^{VT} remained consistently high until 2018 (Study III).

The number of non-PCV10-related IPD cases was very small and thus it was not possible to assess the time trend in VI_{tot}^{NVT}. Nevertheless, the time-average was negative (-78%; 95% CI -214… -2) as expected due to replacement in carriage. VE^{NVT} was close to zero (-17%; 95% CI -622…65).

In neither study not all serotypes fell into the two categories of vaccine and non-vaccine types. Specifically, the non-vaccine serotype 19A exhibited properties of both categories. The incidence of 19A IPD increased especially among unvaccinated children and resulted in a negative vaccine impact. However, vaccine effectiveness against 19A IPD was positive and that against all IPD was higher than expected (Study IV). We separated vaccine-related serotypes from non-vaccine-type IPD in the analysis. The point estimates of both the vaccine effectiveness and total impact against the vaccine-related serotypes were positive, whereas those against the non-vaccine-related serotypes were negative.

We assessed the time trends of vaccine impact estimates by splitting the study period into yearly periods and were able to show an increasing trend in VI_{tot}^{VT}. However, the small number of cases prevented meaningful estimation of the time trend in VI_{tot}^{NVT}. In Study III, VE^{VT} was shown to stay constant over the study period, but a decreasing trend in $VE^{all\ IPD}$ was observed, as expected according to the model.

In estimating vaccine effectiveness, cohort, nested case-control and indirect cohort designs provided estimates that were broadly concordant with each other, but those with the cohort design were usually the most precise (Study III).

6 DISCUSSION

6.1 SUMMARY OF THE MAIN FINDINGS

6.1.1 VACCINE EFFICACY (I, II)

Ideally, the definition and measurement of vaccine efficacy should be based on the actual mechanism of the vaccine's protective action. We defined the estimand of vaccine efficacy against colonisation through acquisition (VE_{acq}) because it better than prevalence describes the natural process of incident colonisation occurrences, and because the effect of conjugate vaccines is most likely on acquisition rather than clearance [47,48].

Longitudinal data are typically used to estimate hazards of acquisition and clearance of multitype pneumococcal colonisation, and would thereby allow the estimation of vaccine efficacy against colonisation. However, collecting frequently sampled data on colonisation is laborious, expensive and unpleasant for the study participants. In Studies I and II we focused on the problem of estimating vaccine efficacy from cross-sectional data, i.e. without the need to collect repeated measurements per study subject. In Study I, we defined an estimator of VE_{acq}, $VE_{acq}^{(odds)}$, through the odds ratio of the target-serotype colonisation in vaccinees versus controls using the quantitative relationship between the prevalence and incidence of colonisation. The method is simple to use and allows, under stationary conditions, estimation of VE_{acq} from one observation of the current status of colonisation per study subject.

The estimand (VE_{acq}), as well as those that had previously been presented in literature, is unconditional in the sense that it does not take into account potential within-host changes in the pneumococcal flora occurring after vaccination. Therefore, in Study II, we considered the estimation of vaccine efficacy by choosing the controls from those individuals who are considered susceptible to acquisition of the select set of target strains. We proposed two improved estimands, and the corresponding estimators were again expressed as odds ratios, making the vaccine efficacy parameters estimable from one measurement per study subject. When these conditional estimands are derived for aggregate sets of strains, e.g. all vaccine or non-vaccine strains, they are coherent in the sense that they are the weighted averages of serotype-specific estimands.

We studied the performance of the odds-ratio-based estimators of vaccine efficacy against acquisition by reanalysing longitudinal data from a trial of a 9-valent pneumococcal conjugate vaccine in Israel [47]. The efficacy against non-vaccine-serotype (NVT) colonisation was estimated practically zero with all estimators except the unconditional $VE_{acq}^{(odds)}$, because the conditional estimators better account for within-host replacement by non-vaccine strains. The null efficacy is what we aimed for, as the vaccine has no direct biological effect on these strains. Overall, we were able to show improvement in terms of less biased estimators and increased precision through choosing the controls from those individuals that are uncolonised or from all those susceptible to acquiring the strains of interest.

Moreover, we studied the performance of the cross-sectional estimators in terms of departures from the model assumptions. If the multitype model is partitioned into three aggregate states so that transitions occur between the target set (W), the reference set $\overline{V_0}$ (i.e. the colonisation states including state 0 against which the vaccine has no direct biological protection), and the "rest" (R), the assumption of symmetrical competition between strains is not needed. The only condition required for cross-sectional estimation is the similarity of clearance hazards of the aggregate states W in the vaccinees and controls. We then studied the implications of this assumption. If the vaccine enhances clearance of vaccine strains, the parameter of vaccine efficacy can be interpreted more generally as a summary measure of the effect of being vaccinated on susceptibility to acquisition and on duration of colonisation. If individuals are infectious throughout colonisation, this quantity is related to the transmission potential of the strain in question.

Importantly, the interpretation of the vaccine efficacy estimates of Study II as measures of the biological effect of the vaccine on new vaccine-type acquisitions is what we aimed for. They should provide more comparable estimation of vaccine efficacy against colonisation than the unconditional VE_{acq} and especially the prevalence-based parameters across different epidemiological settings with differing levels of exposure by non-vaccine strains. However, also the unconditional VE_{acq} is of use as it may be more informative about the extent of within-host replacement. Moreover, as explained earlier, a more general interpretation as the vaccine efficacy against transmission potential may be given.

6.1.2 VACCINE EFFECTIVENESS AND IMPACT (III, IV)

In Studies III and IV, we described the time-dependency of the vaccine impact (VI) and effectiveness (VE) against invasive pneumococcal disease (IPD) among vaccine-eligible children using a pseudo-dynamic model that incorporates the incidences of pneumococcal carriage and disease during the

post-introduction period. Based on the model, we described explicitly the expected time trends in vaccine impact and effectiveness and showed that these two measures may behave very differently over time.

The effectiveness and total impact against non-vaccine-type IPD (VE^{NVT} and VI_{tot}^{NVT}, respectively) are expected to reach clearly different values as the NVT IPD incidence increases due to replacement in carriage. While VI_{tot}^{NVT} decreases, VE^{NVT} increases to 0% once vaccine-type infection has been eliminated and exposure to NVT colonisation is equal in the vaccinated and unvaccinated parts of the population.

Vaccine effectiveness against vaccine-type IPD (VE^{VT}), by contrast, stays constant although the incidence of VT IPD decreases, as both vaccinated and unvaccinated children benefit from herd protection in a similar manner. The total impact against vaccine-type IPD (VI_{tot}^{VT}) increases to 100% when the vaccine serotypes have been eliminated.

The argument that VE^{VT} remains constant over time may be a simplification because it discards within-host competition between serotypes in the nasopharynx. Since vaccination reduces vaccine-type carriage, the non-vaccine serotypes may gain better ability to colonise the opened niche [49]. If this was taken into account in the estimation, VE^{VT} could increase over time. Although the effect should be small especially in the situation where vaccine-type incidence is very low, it may be relevant in the context of vaccine efficacy estimation (cf. Studies I and II).

We estimated the vaccine impact and effectiveness using register data of vaccine-eligible children in Finland. While the incidence of PCV10 serotypes decreased substantially during the study perios, VI_{tot}^{VT} was higher than VE^{VT} and increased over years reaching 100% in six years. VE^{VT} remained consistently high. These results corresponded to those of the pseudo-dynamic model.

The number of non-PCV10-related IPD cases was very small and it was thus not possible to assess the time trend in VI_{tot}^{NVT}. Nevertheless, the time-average was negative due to replacement in carriage. VE^{NVT} was less negative and close to the expected value of zero.

Another factor in the Finnish setting was that not all serotypes fall into the two broad categories of vaccine and non-vaccine types. The non-vaccine serotype 19A has been the main replacing serotype in Finnish population including children. However, in vaccine-eligible children, 19A seems to have occurred more often in unvaccinated as compared to vaccinated children. As there have been changes in the distribution of circulating 19A clones [159], the assumption about constant case-to-carrier ratios may have not been fully accurate. Therefore, we separated vaccine-related types from non-vaccine-type IPD

(Studies III and IV). The point estimates of both the vaccine effectiveness and total impact against the vaccine-related serotypes were positive, whereas those against the non-vaccine-related serotypes were negative.

In Study III, the cohort, nested case-control and indirect cohort designs resulted in estimates of vaccine effectiveness that were broadly concordant with each other, although those with the cohort design were usually the most precise. PCV10 offered a sustained and high effectiveness against PCV10-serotype-IPD to vaccinated children during the first decade after introduction into the programme.

6.2 COMPARISON TO OTHER STUDIES

6.2.1 VACCINE EFFICACY (I, II)

Vaccine efficacy against pneumococcal colonisation has been assessed in several studies based on testing the hypothesis of equal vaccine-type prevalence in vaccinees and controls [72,160] or estimating the prevalence (or prevalence odds ratio) as the effect measure of the relative risk of vaccine-type colonisation [33,73,126]. The interpretation of vaccine efficacy across different studies may then be confounded by the variability in the choice of the effect parameter and factors related to the setting, e.g. specific colonisation endpoint (acquisition, duration or density), force of acquisition, control vaccine, timing of colonisation measurements and sample size [161].

To harmonise the interpretation and to obtain better external validity of vaccine efficacy against acquisition, we proposed to define vaccine efficacy against acquisition through a hazard ratio and an estimator through the odds ratio of the target serotype colonisation in vaccinees versus controls from cross-sectional measurements. In some studies, similar ideas have been applied. In an Israeli trial, odds ratios were used as estimators of the hazard of acquisition in PCV9 vaccine recipients compared with control subjects [47]. In a PCV7 trial conducted among American Indians [48], serotype-specific odds ratios were determined by restricting denominators to those not carrying pneumococci, i.e. defining the effect measure similarly to our vaccine efficacy estimator that is conditioned on uncolonised states (Study II). More recently, the concept of hazard-based estimation of vaccine efficacy has been applied in several observational studies by choosing the odds-ratio-based estimators [101–103,109,110,162]. Moreover, in a very recent study, vaccine efficacy against transmission of SARS-CoV-2 was estimated from cross-sectional data [163]. Referring to Study I, the authors defined the combined efficacy of a vaccine against acquisition and duration of infection as the odds ratio of vaccination in those that were tested positive and negative. They interpreted the parameter

as efficacy against transmission potential of an individual similary as we have suggested in Studies I and II.

Methodologically, the estimators of vaccine efficacy against acquisition, as specified in Studies I and II, are related to the nested case control design in a setting with multiple possible endpoints and sampling based on prevalent cases. Related statistical methods to estimate vaccine efficacy against colonisation or invasive pneumococcal disease in a setting with multiple serotypes include sieve analysis [164] and the indirect cohort method [80]. Our approach generalises the indirect cohort method to the analysis of transient, common and recurrent (colonisation) events with appropriate adjustment for replacement carriage within the host.

Sieve analysis is a method to assess variation in vaccine efficacy across different serotypes, based on cross-sectional data [164]. Sieve is the vaccine's strain-specific immune barrier to infection, and the aim of the analysis is to examine the ratio of two serotype-specific vaccine efficacies, defined through risk ratios. No absolute vaccine efficacy estimates will be directly available with the method, except if the efficacy against the comparator serotype is known to be zero. The main difference between our approach and the sieve analysis is that the outcomes in the latter method are non-transient (see also Auranen et al. [161]).

Our concept of conditional estimators (Study II) has been used to advice the aggregation of type-specific hazards to simplify a multi-type model and to adjust for the time each individual spends being at risk for the respective transitions [14,165]. The need for correct estimation of vaccine efficacy against colonisation and disease is apparent in the modelling studies that aim at predicting the impact of large-scale vaccinations and advising the optimal serotype compositions of new vaccines. [139–141,143]

6.2.2 VACCINE EFFECTIVENESS AND IMPACT (III, IV)

To our knowledge, Study III is the first to compare three parallel study designs in the estimation of PCV effectiveness against IPD in the same setting. Moreover, although vaccine effectiveness and impact have been compared in other settings, our follow-up was long and we paid special attention to the expected time trends in the effect parameters.

Two previous studies have compared the indirect cohort and matched case-control designs in the same setting. In Brazil, the effectiveness against PCV10-serotype IPD was estimated somewhat higher with the matched case-control design than with the indirect cohort design (84% and 73%, respectively) [90,91]. In the US, the estimates of vaccine effectiveness against PCV13 serotype IPD were almost equal with the two designs (96% and 97%, respectively) [93,166].

Only few studies have compared vaccine effectiveness and impact simultaneously. In the Gambia and Navarra, Spain, cases in children were compared both in parallel using a (nested) case-control design and with the cases from time period before vaccinations to estimate the effectiveness and impact of PCV13 against radiologically confirmed pneumonia and IPD, respectively over years 2011-2014 in both studies [94,167]. In Navarra, the total and indirect impacts against IPD were high and similar to each other (76% and 78%, respectively). Vaccine effectiveness was estimated at 95%. In the Gambia, the point estimates were lower due to the unspecific outcome. The overall impact and effectiveness were estimated at 29% and 43% against radiologically confirmed pneumonia, respectively.

6.3 STRENGTHS AND LIMITATIONS

6.3.1 VACCINE EFFICACY (I, II)

In order to use an odds-ratio-based estimator for a hazard-ratio-based estimand of vaccine efficacy, several conditions need to be fulfilled. First, the turnover of pneumococcal serotypes in the study cohort must be at least approximately stationary so that cross-sectional samples of colonisation at any time point are statistically similar. In such a situation, the odds ratio estimates the hazard ratio [57,168,169]. Instead, the prevalence ratio may not be valid, because in case of a common disease outcome such as pneumococcal colonisation and when a vaccine is efficacious, it overestimates the prevalence odds ratio [170]. With regard to the stationary serotype distribution in carriage, such plateau is usually obtained within a few months after birth, in particular when exposure to pneumococci is high [1,171–173]. According to a simulation study, at least twice the mean duration of a carriage episode needs to be passed after vaccination to reach a steady state in pneumococcal carriage [174].

The second assumption in our multitype model is that the strains compete with each other symmetrically. It means that for any two strains, the ratio of their two hazards of acquisition to the doubly colonised state equals the ratio of the respective hazards of colonisation from empty nasopharynx. The assumption can be relaxed if vaccine efficacy is estimated against a group of strains, e.g. against all vaccine strains. Then, the transitions of the Markov model occur between the target set, the reference set and the “rest”. The group “rest” includes states with two vaccine strains and can be disregarded from the actual vaccine efficacy estimation.

The third condition required for cross-sectional estimation is the similarity of clearance hazards of the single strains or the aggregate target group in the vaccinees and controls, i.e. that vaccination does not affect the duration of

carriage. Two studies that specifically assessed vaccine efficacy against duration, did not find any effect [47,48]. Importantly, if the assumption does not hold, the vaccine efficacy estimand can still be interpreted more generally as a summary measure of the effect of being vaccinated on susceptibility to acquisition and duration of colonisation.

6.3.2 VACCINE EFFECTIVENESS AND IMPACT (III, IV)

In Study III, we used three parallel designs, i.e. the cohort, case-control and indirect cohort designs to estimate vaccine effectiveness. All three designs are unconditional in the sense that they do not adjust for the differential within-host replament between the vaccinated and unvaccinated. By contrast, vaccine impact implicitly adjusts for replacement by comparing the hazard of infection in the vaccinated with a completely unvaccinated population not experiencing any indirect effects. Therefore, under strong indirect effects induced by the vaccination programme, vaccine effectiveness and impact may behave very differently over time.

On the levels of information hierarchy of Rhodes et al. [61], the cohort design allows direct estimation of hazard ratios. Also the case-control design allows estimation of hazard ratios through odds ratios by sampling controls from those who were at risk at the time of case occurrence [76]. The indirect cohort design is on the same level of information as the case-control design with regard to the contact structure of study subjects, being essentially a case-control design. However, it requires only disease surveillance data, and no data on vaccination status of the population or vaccination coverage are needed [80]. None of the three designs has information on all contacts that would allow estimation of vaccine efficacy through secondary attack rates. However, in our setting with nation-wide registers and a homogeneous population, such bias of inequal exposure to infection would be averaged over.

VE estimates based on the indirect cohort design may be biased due to disproportionate replacement by non-vaccine serotypes among vaccinated and unvaccinated IPD cases [82]. According to our study (III), VE estimates were higher when using the indirect cohort design than with the other designs but the difference was not large. Moreover, our theoretical study (Study IV) showed that the bias diminishes as the proportion of VT carriage decreases and exposure to NVT colonisation becomes more equal in the vaccinated and unvaccinated parts of the population.

When using the three parallel study designs in VE estimation, we adjusted or matched all analyses in a similar manner with age group, sex and calendar year to reach the best possible comparability. However, the different designs could be used in more optimal ways than what we did in our comparison study, depending on the available data. Generally, the cohort design is optimal in

utilising the entire follow-up of the underlying cohorts, whereas the indirect cohort and nested case-control designs may allow for better adjustment than the cohort design as data are often more easily available for small subsets of the population. In our study, the cohort design reached the best precision in the sense that the estimated credible intervals were narrower than with the other designs.

The estimation of vaccine effectiveness is often based on numbers of vaccinated and unvaccinated cases per person-time, accumulated until certain time point. This means that the effectiveness becomes a time average, although the measure itself could be time-varying. As long as VE against the vaccine types is assumed – and observed as in our study – to remain constant over time, this is not a problem. However, as VE against the non-vaccine serotypes is expected to increase from a negative level towards zero, a time-average may hide the trend if follow-up does not accrue constantly over time. In our study (III), the number of non-PCV10-related IPD cases was very small and thus it was not possible to assess the time trend in VE^{NVT}. Nevertheless, the time-average was close to the expected value of zero.

Vaccine effectiveness studies are subject to ecological and sampling biases. Vaccine impact is a dynamic measure that depends on the level of vaccination coverage and distribution in the population. Changes in coding practices, access to care, diagnosis or treatment guidelines or in other interventions will affect the results. A long pre-vaccination period and analysis at the serotype group level (e.g. VT/NVT) instead of serotype-specific level help in balancing out secular trends. In our study, the pre-vaccination period was relatively long and there were no major changes in coding practices or diagnostics of invasive pneumococcal disease during the study period. Moreover, no notable pre-vaccine trends were observed and the method of choice was thus the before-after design instead of interrupted time-series analysis.

A strength of our application with empirical data are the nation-wide health registers that can be linked with unique personal identity codes, allowing the estimation of VE in the whole vaccine-eligible cohort, and VI with a long pre-vaccination period. Small case numbers made the estimation based on short time bands, evaluation of possible time-trends in VE, and comparison with the theoretical model difficult. However, the observed time trends in VI estimates corresponded to the model predictions. Despite the small data and the simplified structure of the model, our study should serve to exemplify how the unobservable carriage process modifies disease dynamics and how VI and VE behave in the long-term post-vaccination.

6.4 CONCLUSIONS

Vaccine effect measures need to be chosen based on the natural history and epidemiology of the infection in question and on the assumed mechanism of vaccine action. The setting and feasibility of data collection then guide the choice of the appropriate estimators.

The most rational way to approach vaccine efficacy against susceptibility to pneumococcal colonisation appears to be through hazard of acquisition, which describes the natural process of incident occurrences of colonisation better than the prevalence. Moreover, vaccine efficacy against a multi-type pathogen can be interpreted as a weighted average of strain-specific efficacies. Importantly, cross-sectional measurements of the current status of pneumococcal colonisation can be used to estimate vaccine efficacy through odds ratios without long follow-up of study participants. If the estimand is defined controlling for the differential time at-risk for acquiring the serotypes of interest, it better describes the direct biological effect of the vaccine on transmission potential or susceptibility to acquisition. The unconditional estimand can also be used to assess the extent of within-host replacement.

Individual and population level effects of vaccination programmes are invariably assessed in observational settings. Vaccine effectiveness and impact are used to describe the benefits for a vaccinated individual and for a whole population of a vaccination programme. Both parameters change over time and may behave very differently due to the changing population dynamics of pneumococcal carriage and disease. Both parameters also depend on other factors such as secular trends and vaccination coverage, and they are not directly comparable between settings.

The conditional estimands of vaccine efficacy as proposed in this thesis can be interpreted as measures of the biological effect of the vaccine on new vaccine-type acquisitions and should therefore allow more robust comparisons across different epidemiological settings with differing levels of exposure by non-vaccine strains. The thesis also helps to interpret the time-varying parameters of vaccine impact and effectiveness during large-scale vaccinations, and their manifestation in Finnish children.

ACKNOWLEDGEMENTS

This study was conducted at the Finnish Institute for Health and Welfare (THL) in the Department of Vaccination and Immune Protection, later called Department of Public Health Solutions.

First and foremost, I want to thank my supervisor, Professor Kari Auranen for guiding me in pneumococcal epidemiology, statistics and scientific writing, and for support and encouragement in finalising this work.

I sincerely thank the reviewers of my thesis, Professors Allen Cheng and Esa Läärä for investing their time and effort. The insightful comments significantly improved the thesis. I also thank senior researcher, Assistant Professor Albert Jan van Hoek who has kindly agreed to be my opponent in the public examination of the thesis.

I warmly thank Docent Arto Palmu for supervision in register-based studies and for encouragement in writing the thesis. I wish to thank the multiprofessional team for the several years of collaboration on assessing the impact of the pneumococcal vaccination programme: Docent Jukka Jokinen, Professor Pekka Nuorti, Hanna Nohynek, MD, PhD, Lotta Siira, PhD, and Maija Toropainen, PhD. I appreciate their support and contribution as my co-authors. I sincerely thank my other co-authors Professor Ron Dagan, Doctor Noga Givon-Lavi and Professor M. Elizabeth Halloran.

I wish to thank my fellow team members in the PneumoCarr project, Panu Erästö, PhD, Fabian Hoti, PhD, Juha Mehtälä, PhD as well as Markku Nurhonen, MSc, with whom I had the opportunity to work also in other projects. I appreciate his expertise in modelling and contribution as co-author.

I am grateful to Professor Terhi Kilpi for providing excellent research conditions and for expertise in pneumococcal studies. Thanks should also go to Esa Ruokokoski and Jonas Sundman for technical support in IT, and to Sanna Tossavainen for help in practical issues and for the endless efforts in building team spirit and happiness into our daily work. I warmly thank all my colleagues for the enjoyable chats both at the office and remotely.

Finally, I wish to thank my dear friends and family for bringing joy and perspective to my life. I warmly thank my parents and parents-in-law for help and care and Pekka for continuous support.

Espoo, January 2022
Hanna Rinta-Kokko

REFERENCES

1. Bogaert D, de Groot R, Hermans P. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. *Lancet Infect Dis*. 2004;4:144-154.

2. Simell B, Auranen K, Käyhty H, Godlblatt D, Dagan R, O'Brien K. The fundamental link between pneumococcal carriage and disease. *Expert Rev Vaccines*. 2012;11(7):841-855.

3. Halloran M, Haber M, Longini IM. Interpretation and estimation of vaccine efficacy under heterogeneity. *Am J Epid*. 1992;136(3):328-343.

4. Hanquet G, Valenciano M, Simondon F, et al. Vaccine effects and impact of vaccination programmes in post-licensure studies. *Vaccine*. 2013;31:5634-5642.

5. Käyhty H, Auranen K, Nohynek H, Dagan R, Mäkelä H. Nasopharyngeal colonization: a target for pneumococcal vaccination. *Expert Rev Vaccines*. 2006;5(5):651-667.

6. Ganaie F, Saad J, McGee L, et al. A new pneumococcal capsule type, 10D, is the 100th serotype and has a large cps fragment from an oral *Streptococcus*. *mBio*. 2020;11(3):e00937-20.

7. Lipsitch M, Dykes J, Johnson S, et al. Competition among *Streptococcus pneumoniae* for intranasal colonization in a mouse model. *Vaccine*. 2000;18:2895-2901.

8. Leino T, Hoti F, Syrjänen R, Tanskanen A, Auranen K. Clustering of serotypes in a longitudinal study of *Streptococcus pneumoniae* carriage in three day care centres. *BMC Infectious Diseases*. 2008;8:173-10.

9. Brueggeman A, Griffiths D, Meats E, Peto T, Crook D, Spratt B. Clonal relationships between invasive and carriage *Streptococcus pneumoniae* and serotype- and clone-specific differences in invasive disease potential. *J Infect Dis*. 2003;187:1424-1432.

10. Hanage W, Auranen K, Syrjänen R, et al. Ability of pneumococcal serotypes and clones to cause acute otitis media: implications for the prevention of otitis media by conjugate vaccines. *Infect Immun*. 2004;72(1):76-81.

11. Goldblatt D, Hussain M, Andrews N, et al. Antibody responses to nasopharyngeal carriage of *Streptococcus pneumoniae* in adults: a longitudinal household study. *J Infect Dis*. 2005;192(3):387-393.

12. Mueller J, Yaro S, Ouedraogo M, et al. Pneumococci in the African meningitis belt: meningitis incidence and carriage prevalence in children and adults. *PLoS ONE*. 2012;7(12):e52464.

13. WHO. Pneumococcal conjugate vaccines in infants and children under 5 years of age: WHO position paper - February 2019. *Wkly Epidemiol Rec*. 2019;94(8).

14. Mehtälä J, Dagan R, Auranen K. Estimation and interpretation of heterogeneous vaccine efficacy against recurrent infections. *Biometrics*. 2016;72:976-985.

15. Melegaro A, Choi Y, Pebody R, Gay N. Pneumococcal carriage in United Kingdom families: estimating serotype-specific transmission parameters from longitudinal data. *Am J Epidemiol*. 2007;166:228-235.

16. Satzke C, Turner P, Virolainen-Julkunen A, et al. Standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*: Updated recommendations from the World Health Organization Pneumococcal Carriage Working Group. *Vaccine*. 2014;32:165-179.

17. ECDC. 2008/426/EC: Commission Decision of 28 April 2008 amending. Decision 2002/253/EC laying down case definitions for reporting communicable diseases to the Community network under Decision No 2119/98/EC of the European Parliament and of the Council (notified under document number C(2008) 1589). Published online 2008. Accessed January 17, 2021. https://op.europa.eu/en/publication-detail/-/publication/3e53de24-26d6-4645-b9ab-3931f3874c9e/language-en

18. WHO. WHO Vaccine-Preventable Diseases Surveillance Standards. Pneumococcus. Last updated September 5, 2018. Published online 2018. Accessed January 9, 2021. https://www.who.int/immunization/monitoring_surveillance/burden/vpd/standards/en/

19. Jokinen J, Rinta-Kokko H, Siira L, et al. Impact of ten-valent pneumococcal conjugate vaccination on invasive pneumococcal disease in Finnish Children – a population-based study. *PlosOne*. 2015;10(3):e0120290.

20. ECDC. *Invasive Pneumococcal Disease*. ECDC; 2019. Accessed August 8, 2020. https://www.ecdc.europa.eu/en/publications-data/invasive-pneumococcal-disease-annual-epidemiological-report-2017

21. von Gottberg A, de Gouveia L, Tempia S, et al. Effects of vaccination on invasive pneumococcal disease in South Africa. *N Engl J Med*. 2014;371:1889-1899.

22. Palmu A, Jokinen J, Nieminen H, et al. Vaccine effectiveness of the pneumococcal *Haemophilus influenzae* protein D conjuate vaccine (PHiD-CV10) against clinically suspected invasive pneumococcal disease: a cluster randomized trial. *Lancet Respir Med*. 2014;2:717-727.

23. Palmu A, Kilpi T, Rinta-Kokko H, et al. Pneumococcal conjugate vaccine and clinically suspected invasive disease. *Pediatrics*. 2015;136(1):e22-7.

24. Cherian T, Mullholland E, Carlin J, et al. Standardized interpretation of paediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies. *Bull World Health Organ*. 2005;83:353-359.

25. Kilpi T, Jokinen J, Puumalainen T, et al. Effectiveness of pneumococcal *Haemophilus influenzae* protein D conjugate vaccine against pneumonia in children: a cluster-randomised trial. *Vaccine*. 2018;36:5891-5901.

26. Cutts F, Zaman S, Enwere G, et al. Cutts FT, Zaman SM, Enwere G, et al. Efficacy of nine-valent pneumococcal conjugate vaccine against pneumonia

and invasive pneumococcal disease in The Gambia: randomised, double-blind, placebo-controlled trial. *Lancet*. 2005;365:1139-1146.

27. Molinos L, Zalacain R, Menendez R, et al. Sensitivity, specificity, and positivity predictors of the pneumococcal urinary antigen test in community-acquired pneumonia. *Ann Am Thorac Soc*. 2015;12(10):1482-1489.

28. Wunderink R, Self W, Anderson E, et al. Pneumococcal community-acquired pneumonia detected by serotype-specific urinary antigen detection assays. *Clin Infect Dis*. 2018;66(10):1504-1510.

29. Sigurdsson S, Kristinsson K, Erlendsdóttir H, Hrafnkelsson B, Haraldsson A. Decreased incidence of respiratory infections in children after vaccination with ten-valent pneumococcal vaccine. *Pediatr Infect Dis J*. 2015;34(12):1385-1390.

30. Eskola J, Kilpi T, Palmu A, et al. Efficacy of a pneumococcal conjugate vaccine against acute otitis media. *N Engl J Med*. 2001;344:403-409.

31. Magnus M, Vestrheim D, Nystad W, et al. Decline in early childhood respiratory tract infections in the Norwegian mother and child cohort study after introduction of pneumococcal conjugate vaccination. *Pediatr Infect Dis J*. 2012;31(9):951-955.

32. McCormick D, Chonmaitree T, Pittman C, et al. Nonsevere acute otitis media: a clinical trial comparing outcomes of watchful waiting versus immediate antibiotic treatment. 2005;115(6):1455-1465.

33. Prymula R, Peeters P, Chrobok V, et al. Pneumococcal capsular polysaccharides conjugated to protein D provide protection against otitis media caused by both *Streptococcus pneumoniae* and nontypable *Haemophilus influenzae*: a randomized double blind efficacy study. *Lancet*. 2006;367:740-748.

34. Trenaghi M, Sáez-Llorens X, López P, et al. Efficacy of pneumococcal nontypable *Haemophilus influenzae* protein D conjugate vaccine (PHiD-CV) in young Latin American children: A double-blind randomized controlled trial. *PLoS Med*. 2014;11(6):e1001657.

35. Palmu A, Rinta-Kokko H, Nohynek H, Nuorti J, Jokinen J. Impact of national ten-valent pneumococcal conjugate vaccine program on reducing antimicrobial use and tympanostomy tube placements in Finland. *Pediatr Infect Dis J*. 2018;37(1):97-102.

36. Niemelä M, Uhari M, Möttönen M, Pokka T. Costs arising from otitis media. *Acta Paediatr*. 1999;88(5):553-556.

37. Goldblatt D. Conjugate vaccines. Editorial Review. *Clin Exp Immunol*. 2000;119:1-3.

38. SII. Serum Institute of India launches India's first fully indigenously developed pneumococcal vaccine, "PNEUMOSIL". Press Release. Published online December 28, 2020. Accessed January 9, 2021. https://www.seruminstitute.com/news_pneumosil_281220.php

39. Merck. Merck Submits Applications for Licensure of V114, the Company's Investigational 15-valent Pneumococcal Conjugate Vaccine, for Use in Adults to the U.S. FDA and European Medicines Agency. Press release. Published online June 22, 2020. Accessed January 9, 2021. https://www.merck.com/news/

40. Pfizer. U.S. FDA accepts for priority review the biologics license application for Pfizer's investigational 20-valent pneumococcal conjugate vaccine for adults 18 years of age and older. Press release. Published online August 12, 2020. Accessed January 9, 2021. https://www.pfizer.com/news/

41. Klugman K, Rodgers G. Time for a third-generation pneumococcal conjugate vaccine. *Lancet Infect Dis*. Published online July 20, 2020. doi:10.1016/S1473-3099(20)30513-2

42. Hammitt L, Campbell J, Borys D, et al. Efficacy, safety and immunogenicity of a pneumococcal protein-based vaccine co-administered with 13-valent pneumococcal conjugate vaccine against acute otitis media in young children: a phase IIb randomized study. *Vaccine*. 2019;37(51):7482-7492.

43. Keech C, Morrison R, Anderson P, et al. A Phase 1 randomized, Placebo-controlled, observer-blinded trial to evaluate the safety and immunogenicity of inactivated *Streptococcus pneumoniae* whole-cell vaccine in adults. *Pediatr Infect Dis J*. 2020;39(4):345-351.

44. Odutola A, Ota A, Antonio M, et al. Immunogenicity of pneumococcal conjugate vaccine formulations containing pneumococcal proteins, and immunogenicity and reactogenicity of co-administered routine vaccines - a phase II, randomised, observer-blind study in Gambian infants. *Vaccine*. 2019;37(19):2586-2599.

45. Weinberger D, Harboe Z, Flasche S, Scott A, Lipsitch M. Prediction of serotypes causing invasive pneumococcal disease in unvaccinated and vaccinated populations. *Epidemiology*. 2011;22:199-207.

46. Baggett H, Watson N, Deloria Knoll M, et al. Density of upper respiratory colonization with *Streptococcus pneumoniae* and its role in the diagnosis of pneumococcal pneumonia among children aged <5 Years in the PERCH study. *Clin Infect Dis*. 2017;64(S3):S317-27.

47. Dagan R, Givon-Lavi N, Zamir O, et al. Reduction of nasopharyngeal carriage of *Streptococcus pneumoniae* after administration of a 9-valent pneumococcal conjugate vaccine to toddlers attending day care centers. *J Infect Dis*. 2002;185:927-936.

48. O'Brien K, Millar E, Zell E, et al. Effect of pneumococcal conjugate vaccine on nasopharyngeal colonization among immunized and unimmunized children in a community-randomized trial. *J Infect Dis*. 2007;196:1211-1220.

49. Lipsitch M. Vaccination against colonizing bacteria with multiple serotypes. *Proc Natl Acad Sci*. 1997;94:6571-6576.

50. Hanage W, Finkelstein J, Huang S, et al. Evidence that pneumococcal serotype replacement in Massachusetts following conjugate vaccination is now complete. *Epidemics*. 2010;2(2):80-84.

51. Nurhonen M, Auranen K. Optimal serotype compositions for pneumococcal conjugate vaccination under serotype replacement. *PLoS Comput Biol*. 2014;10(2):e1003477.

52. Smith P, Rodriques L, Fine P. Assessment of the protective efficacy of vaccines against common diseases using case-control and cohort studies. *Int J Epidemiol*. 1984;13(1):87-93.

53. Halloran M, Longini I, Strichiner C. Estimability and interpretation of vaccine efficacy using frailty mixing models. *Am J Epid*. 1996;144(1):83-97.

54. Halloran M, Haber M, Longini I, Struchiner C. Direct and indirect effects in vaccine efficacy and effectiveness. *Am J Epid*. 1991;133(4):323-331.

55. Rubin D. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psych*. 1974;68(5):688-701.

56. Greenwood M, Yule U. The statistics of anti-typhoid and anti-cholera inoculations, and the interpretation of such statistics in general. *Proc R Soc Med*. 1915;8:113-194.

57. Rothman K, Greenland S, Lash T, (ed.). *Modern Epidemiology*. 3rd ed. Lippincott Williams & Wilkins; 2008.

58. Porta M. *IEA Dictionary of Epidemiology*. 6th ed. Oxford University Press; 2014.

59. Top F. Measles in Detroit, 1935. *Am J Publ Health*. 1938;28:935-943.

60. Haber M, Longini I, Halloran M. Measures of effects of vaccination in a randomly mixing population. *Int J Epid*. 1991;20(1):300-310.

61. Rhodes P, Halloran M, Longini I. Counting process models for infectious disease data: distinguishing exposure to infection from susceptility. *J R Statist Soc B*. 1996;58(4):751-762.

62. Halloran M, Struchiner C, Longini I. Study designs for evaluating different efficacy and effectiveness aspects of vaccines. *Am J Epid*. 1997;146(10):789-803.

63. Halloran M, Longini I, Struchiner C. *Design and Analysis of Vaccine Studies*. Springer; 2010.

64. Greenland S. Interpretation and choice of effect measures in epidemiological analyses. *Am J Epid*. 1987;125(5):761-768.

65. Black S, Shinefield H, Fireman B, et al. Efficacy, safety and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. *Pediatr Infect Dis J*. 2000;19:187-195.

66. Kilpi T, Åhman H, Jokinen J, et al. Protective efficacy of a second pneumococcal conjugate vaccine against pneumococcal acute otitis media in infants and children: randomized, controlled trial of a 7-valent pneumococcal polysaccharide-meningococcal outer membrane protein complex conjugate vaccine in 1666 children. *Clin Infect Dis*. 2003;37(9):1155-1164.

67. Lucero M, Nohynek H, Williams G, Tallo V, Simoes E, et al. Efficacy of an 11-valent pneumococcal conjugate vaccine against radiologically confirmed pneumonia among children less than 2 years of age in the Philippines - A randomized, double-blind, placebo-controlled trial. *Pediatr Infect Dis J*. 2009;28:455-462.

68. Prymula R, Kriz P, Kaliskova E, Pascal T, Poolman J, Schuerman L. Effect of vaccination with pneumococcal capsular polysaccharides conjugated to *Haemophilus influenzae*-derived protein D on nasopharyngeal carriage of *Streptococcus pneumoniae* and H. influenzae in children under 2 years of age. *Vaccine*. 2010;28:71-78.

69. Cheung YB, Zaman S, Nsekpong E, et al. Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian children who participated in a 9-valent pneumococcal conjugate vaccine trial and in their younger siblings. *Pediatr Infect Dis J*. 2009;28:990-995.

70. Bonten M, Huijts S, Bolkenbaas M, et al. Polysaccharide Conjugate Vaccine against Pneumococcal Pneumonia in Adults. *N Eng J Med*. 2015;372:1114-1125.

71. Klugman K, Madhi S, Huebner R, et al. A trial of a 9-valent pneumococcal conjugate vaccine in children with and those without HIV infection. *N Engl J Med*. 2003;349:1341-1348.

72. Mbelle N, Huebner R, Wasas A, Kimura A, Chang I, Klugman K. Immunogenicity and impact of nasopharyngeal carriage of a nonavalent pneumococcal conjugate vaccine. *J Infect Dis*. 1999;180:1171-1176.

73. Obaro S, Adegbola R, Banya W, Greenwood B. Carriage of pneumococci after pneumococcal vaccination. *Lancet*. 1996;348:271-272.

74. Breslow N. In: Ahrens W, Pigeot I (eds.) Handbook of epidemiology, Section 7 Case-Control Studies. In: 2nd edition. Springer Science+Business Media; 2014.

75. Song J, Chung K. Observational studies: Cohort and case-control studies. *Plast Reconstr Surg*. 2010;126:2234-2242.

76. Pearce N. Classification of epidemiological study designs. *Int J Epid*. 2012;41:393-397.

77. Keogh R, Cox D. *Case-Control Studies*. Cambridge University Press; 2014.

78. Rodrigues L, Kirkwood B. Case-control designs in the study of common diseases: updates on the demise of the rare disease assumption and the choice of sampling scheme for controls. *Int J Epid*. 1990;19(1):205-213.

79. Farrington C. Estimation of vaccine effectiveness using the screening method. *Int J Epid*. 1993;22(4):742-746.

80. Broome C, Facklam R, Fraser D. Pneumococcal disease after pneumococcal vaccination. An alternative method to estimate the efficacy of pneumococcal vaccine. *N Eng J Med*. 1980;303(10):549-552.

81. Moberley S, Andrews R. An evaluation of the indirect cohort method to estimate the effectiveness of the pneumococcal polysaccharide vaccine. *J Vaccines & Immunization*. 2014;2(1):4-6.

82. Andrews N, Waight P, Borrow R, et al. Using the indirect cohort design to estimate the effectiveness of the seven valent pneumococcal conjugate vaccine in England and Wales. *PlosOne*. 2011;6(12):e28435.

83. Omori R, Cowling B, Nishiura H. How is vaccine effectiveness scaled by the transmission dynamics of interacting pathogen strains with cross protective immunity? *PLoS ONE*. 2012;7(11):e50751.

84. Dominguez A, Ciruela P, Hernandez S, et al. Effectiveness of the 13-valent pneumococcal conjugate vaccine in preventing invasive pneumococcal disease in children aged 7–59 months. A matched case-control study. *PLoS ONE*. 2017;12(8):e0183191.

85. Rückinger S, van der Linden M, Reinert R, von Kries R. Efficacy of 7-valent pneumococcal conjugate vaccination in Germany: An analysis using the indirect cohort method. *Vaccine*. 2010;28:5012-5016.

86. Martinelli D, Pedalino B, Cappelli M, et al. Towards the 13-valent pneumococcal conjugate universal vaccination. Effectiveness in the transition era between PCV7 and PCV13 in Italy, 2010– 2013. *Hum Vacci Immunother*. 2014;10(1):33-39.

87. Gidding H, McCallum L, Fathima P, et al. Effectiveness of a 3 + 0 pneumococcal conjugate vaccine schedule against invasive pneumococcal disease among a birth cohort of 1.4 million children in Australia. *Vaccine*. 2018;36:2650-2656.

88. Rinta-Kokko H, Auranen K, Toropainen M, et al. Effectiveness of 10-valent pneumococcal conjugate vaccine estimated with three parallel study designs among vaccine-eligible children in Finland. *Vaccine*. 2020;38(6):1559-1564.

89. Rinta-Kokko H, Nurhonen M, Auranen K. Impact and effectiveness of a conjugate vaccine against invasive pneumococcal disease in Finland — a modelling approach. *Hum Vaccin Immunother*. 2021;17(6):1834-1843.

90. Domingues C, Verani J, Renoiner E, et al. Effectiveness of ten-valent pneumococcal conjugate vaccine against invasive pneumococcal disease in Brazil: A matched case-control study. *Lancet Respir Med*. 2014;2(6):464-471.

91. Verani J, Domingues C, de Moraes J. Indirect cohort analysis of 10-valent pneumococcal conjugate vaccine effectiveness against vaccine-type and vaccine-related invasive pneumococcal disease. *Vaccine*. 2015;33:6145-6148.

92. Andrews N, Waight P, Burbidge P, et al. Serotype-specific effectiveness and correlates of protection for the 13-valent pneumococcal conjugate vaccine: a postlicensure indirect cohort study. *Lancet Infect Dis*. 2014;14(9):839-846.

93. De Serres G, Pilishvili T, Link-Gelles R, et al. Use of surveillance data to estimate the effectiveness of the 7-valent conjugate pneumococcal vaccine in children less than 5 years of age over a 9 year period. *Vaccine*. 2012;30:4067-4072.

94. Mackenzie G, Hill P, Sahito S, et al. Impact of the introduction of pneumococcal conjugate vaccination on pneumonia in The Gambia: population-based surveillance and case-control studies. *Lancet Infect Dis*. 2017;17(9):965-973.

95. Madhi S, Groome M, Zar H, et al. Effectiveness of pneumococcal conjugate vaccine against presumed bacterial pneumonia hospitalisation in HIV-uninfected South African children: a case–control study. *Thorax*. 2015;70:1149-1155.

96. Moulton L, O'Brien K, Kohberger R, et al. Design of a Group-Randomized *Streptococcus pneumoniae* Vaccine Trial. Controlled Clinical Trials. 2001;22:438-452.

97. Pirçon JY, Talarico C, Bollaerts K, Hausdorff W, Clake C. The choice of analytical methodology can alter conclusions regarding herd effects of paediatric pneumococcal vaccination programmes. *Vaccine*. 2018;36(46):6933-6943.

98. Lopez Bernal J, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *Int J Epid*. 2017;46(1):348-355.

99. Thorrington D, Andrews N, Stowe J, Miller E, van Hoek A. Elucidating the impact of the pneumococcal conjugate vaccine programme on pneumonia, sepsis and otitis media hospital admissions in England using a composite control. *BMC Medicine*. 2018;16:13.

100. Bruhn C, Hetterich S, Schuck-Paim C, et al. Estimating the population-level impact of vaccines using synthetic controls. *PNAS*. 2017;114(7):1524-1529.

101. Laughlin A, Hsu K, Silverio A, Marchant C, Pelton S. Direct and indirect effects of PCV13 on nasopharyngeal carriage of PCV13 unique pneumococcal serotypes in Massachusetts' children. *Pediatr Infect Dis J*. 2014;33:504-510.

102. Soysal A, Karabağ-Yılmaz E, Kepenekli E, et al. The impact of a pneumococcal conjugate vaccination program on the nasopharyngeal carriage, serotype distribution and antimicrobial resistance of *Streptococcus pneumoniae* among healthy children in Turkey. *Vaccine*. 2016;4(33):3894-3900.

103. Sigurdsson S, Erlendsdóttir H, Quirck S, et al. Pneumococcal vaccination: Direct and indirect effect on carriage of vaccine types and antibiotic resistance in Icelandic children. *Vaccine*. 2017;35::5242-8.

104. Dunne E, Satzke C, Ratu F, et al. Gould K, Hinds J, Tikoduadua L, Kado J, Rafai E, Kama M, Mulholland K, Russell F. Effect of ten-valent pnemococcal conjugate vaccine introduction on pneumococcal carriage in Fiji: results from four annual cross-sectional carriage surveysLancet Glob Health. *Lancet Glob Health*. 2018;6:e1375-85.

105. Hammitt L, Akech D, Morpeth S, et al. Population effect of 10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of *Streptococcus pneumoniae* and non-typeable *Haemophilus influenzae* in Kilifi, Kenya: findings from cross-sectional carriage studies. *Lancet Glob Health*. 2014;2:e397-405.

106. Grant L, Hammitt L, O'Brien S, et al. Impact of the 13-valent pneumococcal conjugate vaccine on pneumococcal carriage among American Indians. *Pediatr Infect Dis J*. 2016;35(8):907-914.

107. Heinsbroek E, Tafatatha T, Phiri A, et al. Pneumococcal carriage in households in Karonga District, Malawi, before and after introduction of 13-valent pneumococcal conjugate vaccination. *Vaccine*. 2018;36(48):7369-7376.

108. Usuf E, Bottomley P, Bojang R, et al. Persistence of nasopharyngeal pneumococcal vaccine serotypes and increase of nonvaccine serotypes among vaccinated infants and their mothers 5 years after introduction of pneumococcal conjugate vaccine 13 in the Gambia. *Clinical Infectious Diseases*. 2019;68(9):1512-1521.

109. Ricketson L, Wood M, Vanderkooi O, et al. Trends in asymptomatic nasopharyngeal colonization with *Streptococcus pneumoniae* after introduction of the 13-valent pneumococcal conjugate vaccine in Calgary, Canada. *Pediatr Infect Dis J*. 2014;33(7):724-730.

110. Southern J, Andrews N, Sandu P, et al. Pneumococcal carriage in children and their household contacts six years after introduction of the 13-valent pneumococcal conjugate vaccine in England. *PLoS One*. 2018;13(5):e0195799.

111. Rinta-Kokko H, Palmu A, Auranen K, et al. Long-term impact of 10-valent pneumococcal conjugate vaccination on invasive pneumococcal disease among children in Finland. *Vaccine*. 2018;36:1934-1940.

112. De Wals P, Lefebre B, Deceuninck G, Longtin J. Incidence of invasive pneumococcal disease before and during an era of use of three different pneumococcal conjugate vaccines in Quebec. *Vaccine*. 2018;6(3):421-426.

113. Ladhani S, Collins S, Djennad A, et al. Rapid increase in non-vaccine serotypes causing invasive pneumococcal disease in England and Wales, 2000–17: a prospective national observational cohort study. *Lancet Infect Dis*. 2018;18(4):441-451.

114. Naucler P, Galanis I, Morfeld E, Darenberg J, Örtqvist Å, Henriques-Normark B. Comparison of the impact of pneumococcal conjugate vaccine 10 or pneumococcal conjugate vaccine 13 on invasive pneumococcal disease in equivalent populations. *Clin Infect Dis*. 2017;65(11):1780-1789.

115. van der Linden M, Falkenhorst G, Perniciaro S, Imöhl M. Effects of infant pneumococcal conjugate vaccination on serotype distribution in invasive pneumococcal disease among children and adults in Germany. *PLoS ONE*. 2015;10(7):e0131494.

116. Matanock A, Lee G, Gierke R, Kobayashi M, Leidner A, Pilishvili T. Use of 13-valent pneumococcal conjugate vaccine and 23-valent pneumococcal polysaccharide vaccine among adults aged ≥65 years: updated recommendations of the advisory committee on immunization practices. *MMWR Morb Mortal Wkly Rep*. 2019;68:1069-1075.

117. Andrade A, Minamisava R, Policena G, et al. Evaluating the impact of PCV-10 on invasive pneumococcal disease in Brazil: A time-series analysis. *Hum Vaccin Immunother*. 2016;12(2):285-292.

118. Ouldali N, Varon E, Levy C, et al. Invasive pneumococcal disease incidence in children and adults in France during the pneumococcal conjugate vaccine era: an interrupted time-series analysis of data from a 17-year national prospective surveillance study. *Lancet Infect Dis*. Published online 2020:Published online.

119. Silaba M, Ooko M, Bottomley C, et al. Effect of 10-valent pneumococcal conjugate vaccine on the incidence of radiologically-confirmed pneumonia and clinically-defined pneumonia in Kenyan children: an interrupted time-series analysis. *Lancet Glob Health*. 2019;7:e337-46.

120. Palmu A, Rinta-Kokko H, Nohynek H, Nuorti J, Kilpi T, Jokinen J. Impact of ten-valent pneumococcal conjugate vaccine on pneumonia in Finnish children in a nation-wide population-based study. *PLoS ONE*. 2017;12(3):e0172690.

121. van Deursen A, Schurink-van't Klooster T, Man W, et al. Impact of infant pneumococcal conjugate vaccination on community acquired pneumonia hospitalization in all ages in the Netherlands. *Vaccine*. 2017;35(51):7107-7113.

122. Okasha O, Rinta-Kokko H, Palmu A, Ruokokoski E, Jokinen J, Nuorti J. Population-level impact of infant 10-valent pneumococcal conjugate vaccination on adult pneumonia hospitalisations, Finland. *Thorax*. 2018;73:262-269.

123. Kislaya I, Rodrigues A, Sousa-Uva M, et al. Indirect effect of 7-valent and 13-valent pneumococcal conjugated vaccines on pneumococcal pneumonia hospitalizations in elderly. *PLoS ONE*. 2019;14(1):e0209428.

124. Andrade A, Afonso E, Minamisava R, et al. Direct and indirect impact of 10-valent pneumococcal conjugate vaccine introduction on pneumonia hospitalizations and economic burden in all age-groups in Brazil: A time-series analysis. *PLoS ONE*. 2017;12(9):e0184204.

125. Lau W, Murray M, El-Turki A, et al. Impact of pneumococcal conjugate vaccines on childhood otitis media in the United Kingdom. *Vaccine*. 2015;33:5072-5079.

126. Millar E, O'Brien K, Watt J, et al. Effect of community-wide conjugate pneumococcal vaccine use in infancy on nasopharyngeal carriage through 3 years of age: a cross-sectional study in a high-risk population. *Clin Infect Dis*. 2006;43(1):8-15.

127. O'Brien K, Moulton L, Reid R, et al. Efficacy and safety of seven-valent conjugate pneumococcal vaccine in American Indian children: group randomised trial. *Lancet*. 2003;362:355-361.

128. Palmu A, Jokinen J, Borys D, et al. Effectiveness of the ten-valent pneumococcal *Haemophilus influenzae* protein D conjugate vaccine (PHiDCV10) against invasive pneumococcal disease: a cluster randomised trial. *Lancet*. 2013;381:214-222.

129. Palmu A, Jokinen J, Nieminen H, et al. Effect of pneumococcal *Haemophilus influenzae* protein D conjugate vaccine (PHiD-CV10) on outpatient antimicrobial purchases: a double-blind, cluster randomised phase 3-4 trial. *Lancet Infect Dis*. 2014;14:205-212.

130. Palmu A, Jokinen J, Nieminen H, et al. Effectiveness of the ten-valent pneumococcal conjugate vaccine against tympanostomy tube placements in a cluster-randomized trial. *The Pediatr Infect Dis J*. 2015;34(11):1230-1235.

131. Nieminen H, Rinta-Kokko H, Jokinen J, et al. Effectiveness of the 10-valent pneumococcal conjugate vaccine among girls, boys, preterm and low-birth-weight infants - results from a randomized, double-blind vaccine trial. *Vaccine*. 2019;37:3715-3721.

132. Donner A, Klar N. *Design and Analysis of Cluster Randomization Trials in Health Research*. Arnold; 2000.

133. Hayes R, Alexander N, Bennett S, Cousens S. Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Stat Methods Med Res*. 2000;9(2):95-116.

134. Jaffar S, Leach A, Halla A, et al. Preparation for a pneumococcal vaccine trial in The Gambia: individual or community randomisation? *Vaccine*. 2000;18:633-640.

135. Auranen K, Mehtälä J, Tanskanen A, Kaltoft M. Between-strain competition in acquisition and clearance of pneumococcal carriage - epidemiologic evidence from a longitudinal study of day-care children. *Am J Epidemiol*. 2010;171:169-176.

136. Hoti F, Erästö P, Leino T, Auranen K. Outbreaks of *Streptococcus pneumoniae* carriage in day care cohorts in Finland - implications for elimination of transmission. *BMC Infect Dis*. 2009;9:102.

137. Lipsitch M, Abdullahi O, D'Amour A, et al. Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal serotypes in Kenya with a Markov transition model. *Epidemiology*. 2012;23(4):510-519.

138. Melegaro A, Gay N, Medley G. Estimating transmission parameters of pneumococcal carriage in household. *Epidemiol Infect*. 2004;132:433-441.

139. Choi Y, Jit M, Flasche S, Gay N, Miller E. Mathematical modelling long-term effects of replacing Prevnar7 with Prevnar13 on invasive pneumococcal diseases in England and Wales. *PLoS ONE*. 2012;7(7):e39927.

140. Melegaro A, Choi Y, George R, Edmunds W, Millar E, Gay N. Dynamic models of pneumococcal carriage and the impact of the heptavalent pneumococcal conjugate vaccine on invasive pneumococcal disease. *BMC Infect Dis*. 2010;10:90.

141. Ojal J, Flasche S, Hammitt L, et al. Sustained reduction in vaccine-type invasive pneumococcal disease despite waning effects of a catch-up campaign in Kilifi, Kenya: A mathematical model based on pre-vaccination data. *Vaccine*. 2017;35:4561-4568.

142. Willem L, Verelst F, Bilcke J, Hens N, Beutels P. Lessons from a decade of individual-based models for infectious disease transmission: a systematic review (2006-2015). *BMC Infect Dis*. 2017;17:612.

143. Nurhonen M, Cheng A, Auranen K. Pneumococcal transmission and disease in silico: a microsimulation model of the indirect effects of vaccination. *PLoS ONE*. 2013;8(2):e56079.

144. Flasche S, Ojal J, Le Polain de Waroux O, et al. Assessing the efficiency of catch-up campaigns for the introduction of pneumococcal conjugate vaccine: a modelling study based on data from PCV10 introduction in Kilifi, Kenya. *BMC Med*. 2017;15(1):113.

145. Le Polain de Waroux O, Edmunds W, Takahashi K, et al. Predicting the impact of pneumococcal conjugate vaccine programme options in Vietnam. *Hum Vaccin Immunother*. 2018;14(8):1939-1947.

146. Van Effelterre T, Moore M, Fierens F, et al. A dynamic model of pneumococcal infection in the United States: implications for prevention through vaccination. *Vaccine*. 2010;28(21):3650-3660.

147. Flasche S, Le Polain de Waroux O, O'Brien K, Edmunds W. The serotype distribution among healthy carriers before vaccination is essential for predicting the impact of pneumococcal conjugate vaccine on invasive disease. *PLoS Comput Biol*. 2015;11(4):e1004173.

148. Feikin D, Scott J, Gessner B. Use of vaccines as probes to define disease burden. *Lancet*. 2014;383:1762-1770.

149. Gessner B, Feikin D. Vaccine preventable disease incidence as a complement to vaccine efficacy for setting vaccine policy. *Vaccine*. 2014;32:3133-3138.

150. Palmu A, Jokinen J, Nieminen H, et al. Vaccine-preventable disease incidence of pneumococcal conjugate vaccine in FinIP trial. *Vaccine*. 2018;36(14):1816-1822.

151. Gessner B, Jiang Q, Van Werkhoven C, et al. A public health evaluation of 13-valent pneumococcal conjugate vaccine impact on adult disease outcomes from a randomized clinical trial in the Netherlands. *Vaccine*. 2019;37(38):5777-5787.

152. Dagan R, Muallem M, Melamed R, Leroy O, Yagupsky P. Reduction of pneumococcal nasopharyngeal carriage in early infancy after immunization with tetravalent pneumococcal vaccines conjugated to either tetanus toxoid or diphtheria toxoid. *Pediatr Infect Dis*. 1997;16(11):1060-1064.

153. Wacholder S, McLaughlin J, Silverman D, Mandel J. Selection of controls in case-control studies. I Principles. *Am J Epid*. 1992;135(9):1019-1028.

154. Lubin J. Extension of analytic methods for nested and population-based incident case-control studies. *J Chron Dis*. 1986;39(5):379-388.

155. R Core Team (2018). R: A language and environment for statistical computing, Vienna, Austria. https://www.R-project.org/

156. THL. The Vaccination Register. Published online 2021. Accessed March 14, 2021. https://thl.fi/en/web/vaccination/vaccination-coverage/national-vaccination-register

157. R-INLA project. Accessed August 15, 2021. https://www.r-inla.org/

158. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society, Series B*. 2009;71(2):319-392.

159. Toropainen M, Nyholm O, Siira L, et al. Genomic epidemiology and antimicrobial resistance in invasive serotype 19A pneumococci before and after introduction of 10-valent pneumococcal conjugate vaccine (PCV10) in Finland. In: 11th International symposium on pneumococci and pneumococcal diseases. 2018.

160. Frazao N, Brito-Avo A, Simas C, et al. Effect of the seven-valent conjugate pneumococcal vaccine on carriage and drug resistance of *Streptococcus pneumoniae* in healthy children attending day-care centers in Lisbon. *Pediatr Infect Dis*. 2005;24(3):243-252.

161. Auranen K, Rinta-Kokko H, Goldblatt D, et al. Colonisation endpoints in *Streptococcus pneumoniae* vaccine trials. *Vaccine*. 2014;32:153-158.

162. Wouters I, Desmet S, Heirstraeten L, et al. How nasopharyngeal pneumococcal carriage evolved during and after a PCV13-to-PCV10 vaccination program switch in Belgium, 2016 to 2018. *Euro Surveill*. 2020;25(5).

163. Lipsitch M, Kahn R. Interpreting vaccine efficacy trial results for infection and transmission. *Vaccine*. 2021;39(30):4082-4088.

164. Gilbert J, Self S, Rao M, Naficy A, Clemens J. Sieve analysis: methods for assessing from vaccine trial data how vaccine efficacy varies with genotypic and phenotypic pathogen variation. *J Clin Epid*. 2001;54:68-85.

165. Man I, Auranen K, Wallinga J, Bogaards J. Capturing multi-type interactions into practical predictors of type replacement following human papillomavirus vaccination. *Phil Trans R Soc B*. 2019;374:20180298.

166. Whitney C, Pilishvili T, Farley M, et al. Effectiveness of seven-valent pneumococcal conjugate vaccine against invasive pneumococcal disease: a matched case-control study. *Lancet*. 2006;368:1495-1502.

167. Guevara M, Ezpeleta C, Gil-Setas A, et al. Reduced incidence of invasive pneumococcal disease after introduction of the 13-valent conjugate vaccine in Navarre, Spain, 2001-2013. *Vaccine*. 2014;32:2553-2562.

168. Keiding N. Age-specific incidence and prevalence: a statistical perspective. *J R Statist Soc*. 1991;154(3):371-412.

169. Thompson M, Myers J, Kriebel D. Prevalence odds ratio or prevalence ratio in the analysis of cross sectional data: what is to be done? *Occup Eviron Med*. 1998;55:272-277.

170. Zocchetti C, Consonni C, Bertazzi P. Relationship between prevalence rate ratios and odds ratios in cross-sectional studies. *Int J Epidemiol*. 1997;26(1):220-223.

171. Granat S, Mia Z, Ollgren J, et al. Longitudinal study on pneumococcal carriage during the first year of life in Bangladesh. *Pediatr Infect Dis J*. 2007;26(4):319-324.

172. Gray M, Turner M, Dillon H. Epidemiologic studies of *Streptococcus pneumoniae* in infants. *Am J Epid*. 1982;116(4):692-703.

173. Hill P, Cheung Y, Akisanya A, et al. Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian infants: a longitudinal study. *Clin Infect Dis*. 2008;46(6):807-814.

174. Auranen K, Rinta-Kokko H, Goldblatt D, et al. Design questions for *Streptococcus pneumoniae* vaccine trials with a colonization endpoint. *Vaccine*. 2014;32:159-164.