CLINICAL EPIDEMIOLOGY OF VACCINES AND VACCINE-PREVENTABLE DISEASES USING ADMINISTRATIVE HEALTHCARE DATA

Dissertation

for the academic grade of

"*Doctor rerum naturalium"* (Dr. rer. nat.)

University of Bremen

Faculty of Human and Health Sciences

Submitted by

Kathrin Thöne (née Hillebrand)

Bremen, March 2018

Leibniz Institute for Prevention Research and Epidemiology - BIPS

First Reviewer: Prof. Dr. Hajo Zeeb

Second Reviewer: Prof. Dr. Ulrike Haug

Academic Supervisor: Prof. Dr. Edeltraut Garbe

Date of Colloquium: August 9th, 2018

ACKNOWLEDGEMENTS

During my time as a PhD student I was supported by many people whom I'd like to thank.

First of all, I would like to thank my supervisor Prof. Dr. Edeltraut Garbe. As my academic mentor and due to her encouraging support, she had a decisive impact on my scientific development.

Secondly, I would like to thank Prof. Dr. Ulrike Haug and Prof. Dr. Hajo Zeeb for their agreement to review this thesis.

Furthermore, I would like to thank the Leibniz Institute for Prevention Research and Epidemiology – BIPS for the opportunity to work with the German Pharmacoepidemiological Research Database (GePaRD) and my colleagues at BIPS for their support, their helpful ideas, fruitful discussions and the great time we had together. Especially, I would like to thank the co-authors with whom I wrote the research articles cited in this thesis. It was always a great pleasure to work in such sublime project teams.

Finally, I would like to thank my parents, my family and my friends for their support. Especially, I would like to thank my husband Björn for his unconditional support as well as my son Jonte for motivating me just by smiling at me and keeping the nights quiet (most of the time). Thank you both for always being there for me.

TABLE OF CONTENT

ABBREVIATIONS AND ACRONYMS

AGW	Anogenital wart
AMI	Acute myocardial infarction
ARI	Acute respiratory illness
ASHIP	Associations of Statutory Health Insurance Physicians
ASIA	Autoimmune/inflammatory syndrome induced by adjuvants
$AS01_B$	Adjuvant system 01_B
ATC	Anatomical Therapeutic Chemical (system, code, etc.)
AVR	Adverse vaccine reaction
BMI	Body mass index
BIPS	Leibniz Institute for Prevention Research and Epidemiology – BIPS
CHMP	Committee for Medical Products for Human Use
CI	Confidence interval
COX-2	Cyclooxygenase-2
CRPS	Complex regional pain syndrome
CPR	Central pharmaceutical reference
CVD	Cardiovascular disease
DALYs	Disability-adjusted life years
DEGS	German Health Interview and Examination Survey for Adults [Ger.: Studie zur Gesundheit von Erwachsenen in Deutschland]
DDD	Defined Daily Dose
DMP	Disease Management Program
DRG	Diagnosis-Related Group
EBM	Physician's fee schedule [Ger.: Einheitlicher Bewertungsmaßstab]
EMA	European Medicines Agency
EPI	Expanded Program on Immunization
EU	European Union
FDA	U.S. Food and Drug Administration
FSME	Tick-borne encephalitis (TBE) [Ger.: Frühsommer-Meningoenzephalitis]
G-BA	Joint Federal Committee [Ger.: Gemeinsamer Bundesausschuss]
GePaRD	German Pharmacoepidemiological Research Database
Ger.	German language
GPs	General practitioners
GRADE	Grading of recommendations, assessment, development and evaluation
GSK	GlaxoSmithKline plc.

Hib	*Haemophilus influenzae* type b
HIT	Herd immunity threshold
HPV	Human papillomavirus
HZ	Herpes zoster
HZO	Herpes zoster ophthalmicus
ICD-10-GM	International statistical classification of diseases and related health problems, 10th revision, German modification
IfSG	Protection against Infection Act [Ger.: Infektionsschutzgesetz]
IgG	Immunoglobulin G
IR	Incidence rate
IRR	Incidence rate ratio
KEE	Kindergarten entrance examinations
KiGGS	German Health Interview and Examination Survey for Children and Adolescents [Ger.: Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland]
MRSA	Methicillin-resistant *Staphylococcus aureus*
tNSAID	Traditional nonsteroidal anti-inflammatory drug
OA	Original article
OPS	Operation and Procedure Key [Ger.: Operationen- und Prozedurenschlüssel]
OPV	Oral polio vaccine
OR	Odds ratio
PEI	Paul-Ehrlich-Institute
PHN	Postherpetic neuralgia
PICO	Patient, intervention, comparator and outcome
POTS	Postural orthostatic tachycardia syndrome
PRAC	Pharmacovigilance Risk Assessment Committee
PY	Person-years
PZN	Pharmaceutical number [Ger.: Pharmazentralnummer]
RCT	Randomized controlled trial
SARS	Severe acute respiratory syndrome
SCCS	Self-controlled case-series
SEE	School entrance examinations
SHIs	Statutory health insurance providers
SIRs	Standardized incidence rates
STIKO	German Standing Committee on Vaccination [Ger.: Ständige Impfkommission]
STIs	Sexually transmitted infections
TND	Test-negative case-control study design

WHO	World Health Organization
VZV	Varicella zoster virus

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

For centuries, infectious diseases have been among the top 10 leading causes of death. In 2015, they accounted for about 11% of 56.4 million deaths worldwide indicating a global public health relevance.

Vaccines provide an opportunity to eliminate or even eradicate infectious diseases. In order to maximize the benefit of vaccines while minimizing their risks, it is important to investigate infectious disease etiology as well as to continuously monitor and evaluate direct vaccination effects as well as indirect vaccination effects through herd immunity. Since vaccines are usually administered to healthy people to prevent infectious diseases, the monitoring of the safety of vaccines is of high importance. For their evaluation, there is usually no trade-off between risk of disease and risk of its treatment as is usually the case for the treatment of severe diseases. Vaccine safety is also essential for the acceptance of vaccines in the population and thus for high vaccine uptake to enable herd immunity.

Epidemiological observational studies are a powerful tool to investigate the burden of vaccine-preventable diseases, direct and indirect vaccination effects as well as the safety of vaccines in a real-world setting, meaning they provide important data as they include—often in contrast to randomized controlled trials—older or immunocompromised people as well as children or pregnant women, which are most often the target population groups for vaccinations.

Administrative data are a valuable data source for epidemiological observational studies and are increasingly used for studies on vaccines and vaccine-preventable diseases. However, a comprehensive knowledge of the healthcare system itself, including reimbursement policies, but also of the data source and the containing information depth is required.

This thesis investigates different aspects of vaccines and vaccine-preventable diseases. Thus, in a first study, the burden of the vaccine-preventable disease of herpes zoster (HZ) and its complications is investigated and in a separate study, the risk of stroke complication after HZ infection. Vaccine uptake of the human papillomavirus (HPV) vaccine at the population level as well as its indirect impact after vaccine recommendation is assessed. Furthermore, this thesis discusses the nested case-control design with respect to its potential use for direct effectiveness and safety studies of vaccines. Relevant methodological challenges when using different observational study designs based on administrative healthcare data as well as methods to control confounding or to reduce bias are elucidated and discussed. Finally, this thesis gives outlook on potential challenges of future studies on vaccines and vaccine-preventable diseases, especially with regard to newly developed therapeutic vaccines for chronic diseases.

ZUSAMMENFASSUNG

Seit Jahrhunderten gehören Infektionskrankheiten zu den 10 häufigsten Todesursachen der Menschheit weltweit. Im Jahr 2015 waren rund 11% der ca. 56,4 Millionen Todesfällen auf Infektionskrankheiten zurückzuführen, was ein weltweites Public-Health-Problem darstellt. Impfstoffe bieten die Möglichkeit der Elimination oder Eradikation von Infektionskrankheiten. Um den Nutzen von Impfstoffen zu maximieren und gleichzeitig Risiken zu minimieren, ist es wichtig, die Ätiologie von Infektionskrankheiten zu untersuchen, sowie die direkte Wirksamkeit als auch die indirekte Wirksamkeit von Impfstoffen durch Bevölkerungseffekte (z.B. Herdenimmunität) kontinuierlich zu überwachen und zu bewerten. Da Impfungen üblicherweise gesunden Personen—insbesondere Kindern—verabreicht werden, ist zudem die Überwachung der Impfstoffsicherheit von großer Bedeutung. Anders als bei der Behandlung schwerer Erkrankungen besteht bei der Bewertung der Impfstoffsicherheit normalerweise kein Kompromiss zwischen dem Krankheitsrisiko und dem Behandlungsrisiko. Die Impfstoffsicherheit ist deshalb auch für die Akzeptanz von Impfstoffen in der Bevölkerung und somit für die Durchimpfungsrate wichtig.

Epidemiologische Beobachtungsstudien auf Grundlage von GKV-Routinedaten sind ein wirksames Instrument, um die Krankheitslast von impfpräventablen Erkrankungen, die direkte und indirekte Wirksamkeit von Impfungen sowie die Sicherheit von Impfstoffen im Rahmen von real-world Daten zu untersuchen. Im Gegensatz zu randomisiert kontrollierten Studien können Daten von älteren oder immungeschwächten Menschen, Kindern und schwangere Frauen—die oftmals die Zielgruppe für Impfungen darstellen—analysiert werden. GKV-Routinedaten stellen dabei eine wertvolle Datenquelle für Studien zu Impfstoffen und impfpräventable Erkrankungen dar. Die Durchführung solcher Studien setzt ein umfassendes Wissen über das Gesundheitssystem (inkl. der Erstattungsrichtlinien) und die Informationstiefe der Daten voraus.

Die vorliegende Arbeit untersucht verschiedene Aspekte von Impfungen und impfpräventablen Erkrankungen. So wird die Krankheitslast von Herpes Zoster (HZ) und ihrer Komplikationen und in einer separaten Studie das Risiko von Schlaganfall nach einer HZ Infektion untersucht. Es wird die Inanspruchnahme der humanen Papillomavirus (HPV)-Impfung sowie die indirekte Wirksamkeit der HPV-Impfung nach der Impfempfehlung durch die STIKO untersucht. Das Design einer eingebettete Fall-Kontroll-Studie wird hinsichtlich der potenziellen Anwendung für direkte Wirksamkeits- und Sicherheitsstudien von Impfungen diskutiert. Weiterhin werden relevante methodische Herausforderungen in der Anwendung verschiedener Studiendesigns auf Basis von GKV-Routinedaten sowie Methoden zur Kontrolle von Confounding und zur Reduktion von Bias aufgezeigt und diskutiert. Abschließend gibt diese Arbeit einen Ausblick auf mögliche Herausforderungen zukünftiger Impfstudien, insbesondere im Hinblick auf neu entwickelte therapeutische Impfstoffe gegen chronische Erkrankungen.

PREFACE

This cumulative doctoral thesis was prepared based on five research articles which are listed below and are attached in the appendix.
By giving the necessary background information (chapter 1), the following sections will place the respective articles in the overall context of clinical epidemiology of vaccines and vaccine-preventable diseases. The results of the individual research articles will be elucidated subsequently (chapter 2). Methodological issues on the use of administrative data will be discussed, particularly with regard to the strengths and limitations of the different applied study designs (e.g., cohort study, nested case-control study, self-controlled case-series design and ecologic study) as well as with regard to examples of bias and confounding that needed to be addressed within the studies (chapter 3). A final conclusion and an outlook for future research on vaccines and vaccine-preventable diseases will be given (chapter 4).

1. **Hillebrand K,** Bricout H, Schulze-Rath R, Schink T, Garbe E. Incidence of herpes zoster and its complications in Germany, 2005-2009. *Journal of Infection* 2014; 70(2):178-186. [https://doi.org/10.1016/j.jinf.2014.08.018]

2. Schink T, Behr S, Bricout H, **Thöne K**, Garbe E. Risk of Stroke after Herpes Zoster-Evidence from a German Self-Controlled Case-Series Study. *PloS One* 2016; 11(11): e0166554. [https://doi.org/10.1371/journal.pone.0166554]

3. Hense S, **Hillebrand K**, Horn J, Mikolajczyk R, Schulze-Rath R, Garbe E. HPV vaccine uptake after introduction of the vaccine in Germany: an analysis of administrative data. *Human Vaccines & Immunotherapeutics* 2014; 10(6):1729-33. [https://doi.org/10.4161/hv.28450]

4. **Thöne K**, Horn J, Mikolajczyk R. Evaluation of vaccination herd immunity effects for anogenital warts in a low coverage setting with human papillomavirus vaccine – An interrupted time series analysis from 2005 to 2010 using health insurance data. *BMC Infectious* Diseases 2017; 17(1):564. [https://doi.org/10.1186/s12879-017-2663-7]

5. **Thöne K**, Kollhorst B, Schink T. Non-steroidal anti-inflammatory drug use and the risk of acute myocardial infarction in the general German population: a nested case-control study. *Drugs- Real World Outcomes 2017;* 4:127–137 [https://doi.org/10.1007/s40801-017-0113-x]

Furthermore, a co-authored research article cited in this thesis assesses the representativeness of the database on which the five research articles of this thesis are based.

6. Ohlmeier C, Langner I, **Hillebrand K**, Schmedt N, Mikolajczyk R, Riedel O, Garbe E. Mortality in the German Pharmacoepidemiological Research Database (GePaRD) compared to national data in Germany: results from a validation study. *BMC Public Health* 2015; 15:570. [https://doi: 10.1186/s12889-015-1943-7].

1 BACKGROUND

1.1 Vaccination - An Opportunity to Eradicate Vaccine-Preventable Diseases

Infectious disease epidemics have been documented throughout history[1]. Until the late 18th century, infectious diseases constituted the most frequent health risks and causes of death[2]. Due to the improvement of social and hygienic conditions and the isolation of infected patients, high infection rates could be reduced[3,4]. However, reliable infection prevention was lacking until the discovery of active immunization by Edward Jenner in 1796[5]. Jenner found that by inoculating a human host with cowpox, a protection against smallpox (Latin: vaccinia virus) could be induced. Due to this discovery, for the first time in human history, the spread of an infectious disease was widely prevented.

The continuous quality improvement of the smallpox vaccine over time and a consequent vaccination campaign conducted by the World Health Organization (WHO) from 1967 to 1977[6] led to a sufficiently high vaccination coverage of at least 80% worldwide, rendering smallpox the first eradicated human disease in 1979[7].

Resulting from this achievement, vaccinations have gained a high social and health policy acceptance as they are one of the most successful and economical global health interventions to prevent diseases and save lives[8].

WHO initiated the Expanded Program on Immunization (EPI) in May 1974 with the objective of developing and expanding global immunization programs and of pursuing the major goal of reducing child/infant morbidity and mortality against vaccine-preventable diseases[6]. Owing to the successful development of vaccines and the EPI program vaccines against 26 serious vaccine-preventable infections are available to date and many common severe infectious diseases, such as diphtheria, tetanus, pertussis, poliomyelitis and measles, are nowadays kept under control with the exception of some local outbreaks[9].

However, despite the undeniable success of vaccines against serious vaccine-preventable diseases, vaccine opponents are questioning the efficacy, effectiveness and safety of vaccines—even though evidence shows that the benefits of preventing morbidity and mortality from infectious diseases outweigh the risks[10]. Furthermore, misguided safety concerns in some countries have reduced vaccination coverage, resulting in recurrent disease outbreaks of, e.g., pertussis and measles[11,12]. This indicates that public trust in vaccines is an important factor for successful vaccination strategies.

During the vaccine development process, vaccines are extensively checked for their efficacy, immunogenicity and safety in pre-clinical and clinical studies (Phase I–III trials) under ideal conditions[13]. But also after a vaccine has been licensed and/or has been introduced in an

immunization program or schedule, it has to be continuously monitored in real-world settings[14]. Observational studies based on administrative data can be used to analyze the burden of vaccine-preventable diseases and to evaluate direct and indirect vaccine effects after introduction of a vaccine program on the individual and population level as well as to monitor the uptake and safety of immunization programs or schedules[15]. This helps to guide vaccination strategies and to ensure that vaccination targets are being reached[15]. Furthermore, the fact that these studies are carried out and published may improve vaccination coverage by promoting trust in vaccines' impact and safety.

1.2 Clinical Epidemiology of Vaccine-Preventable Diseases

Clinical epidemiology provides the methodological basis for the investigation of vaccines and vaccine-preventable diseases as it composes a conjunction between clinical medicine and epidemiology[16].

The following chapter describes the importance of observational studies based on administrative data on the burden of disease to expand, e.g., the empirical basis for decisions on vaccinations.

1.2.1 Quantifying Vaccine-Preventable Diseases: Incidence and Prevalence

The decision-making process regarding the development and introduction of vaccines into an immunization program or vaccination schedule is always supported by information on the burden of vaccine-preventable diseases and the public health needs in the population[8]. A vaccine will be most accepted and the achievement will be greatest if the vaccine-preventable disease is a visible and distinguished disease in the community[8].

There are several ways to express the burden of disease by assessing (i) the number of cases (e.g., incidence or prevalence), (ii) the number of deaths, (iii) the occurrence of (severe) complications after infection, (iv) the number of hospitalizations or (v) disability-adjusted life years (DALYs)[17].

The incidence and prevalence estimation of a disease is among the basic measures of disease burden calculations in epidemiology[18].

While the incidence rate of an infectious disease measures new events of that disease *per unit of time*, the prevalence is the proportion of persons with the disease at *a specific time point*. Both measurements are related to each other by the duration of the disease[19].

Prevalence of a disease is not only determined by the causes/risk factors of the disease, but also by the determinants of disease survival and is therefore less suitable for studies of etiologic research[20].

Incidence estimation makes is easier to describe and investigate cause and effect as it is certain that the exposure has preceded the outcome.

Therefore, for burden of disease, vaccine effectiveness and safety cohort studies, incidence measures are more useful (i) to study causes of diseases as their investigation is based on (comparing) incidence rates among subgroups with different risk factors or different risk exposures[18] or (ii) to investigate changing patterns of a disease within a population over time. Furthermore, as vaccination usually occurs *before* the disease occurs, the main interest of vaccine studies is to investigate potential changes (reduction) of the incidence rate of the targeted disease *after* vaccination (chapter 1.3.2 and 1.3.3). In the context of vaccine effectiveness and safety studies, it is possible to investigate a presumable causal relationship of the occurrence of new adverse events *after* vaccination (chapter 1.3.4).

An example for a burden of disease study included in this thesis is the cohort study on incidence rates of herpes zoster (HZ) and its complications[21]. This study presents age-, sex- and immune status-specific incidence estimates on HZ and its complications as well as proportions of postherpetic neuralgia (PHN). Results of this study may help to define population subgroups with the highest disease burden and the greatest need of being targeted by vaccination strategies.

1.2.2 Risks of Complications after Infectious Diseases

Short- and long-term consequences (i.e., complications and sequelae) of an infection reflect the risk of severe diseases and/or premature mortality as well as conditions that limit the quality of life like PHN. Hence, besides protection against the infectious agent itself, vaccination particularly confers protection against severe complications and diseases. Especially immunocompromised and older patients are at increased risk of complications after an infection.

For a long time, vaccine-preventable diseases and chronic diseases were considered independent entities. But in recent years, it has been recognized that many infectious agents can lead to severe complications and chronic diseases like cardiovascular diseases (CVDs), chronic respiratory and endocrine diseases as well as cancer[22]. Since these diseases are globally of major public health relevance, their prevention is highly important and the opportunity to vaccinate against severe chronic diseases or cancer would have a major public health impact.

However, it is often still difficult to demonstrate the causal relation between a virus infection and the development of particular cancer types due to long latency. This is due to the fact that multiple risk factors like environmental, biologic and lifestyle factors may contribute jointly or independently to carcinogenesis[23]. However, the causal association of an infection with the human papillomavirus (HPV) and the development of cervical cancer has been

ascertained by several molecular studies as well as epidemiological studies[24,25]. Currently, the finding that HPV infection (i.e., most frequently high-risk types 16 and 18) is a necessary cause of cervical cancer development is considered scientific evidence[24,26]. Furthermore, non-cancer causing types of HPV (especially low-risk types 6 and 11) are associated with the development of anogenital warts (AGWs)[27]. Considering the long latency between HPV infection and cervical cancer development it would require an extremely long follow-up and enormous study size for observational studies to estimate the HPV-vaccination impact regarding the burden of the HPV-related infectious disease outcome cervical cancer. Instead, surveillance of outcomes with less lead time—as is the case with AGWs—provides early outcome measures for vaccine effectiveness (chapter 1.3.2) and impact studies (chapter 1.3.3) on the population level[28]. Furthermore, mathematical predictive models provide a valuable tool to examine long-term effects of vaccination strategies that cannot be investigated during the study period of observational studies[29].

An example for a putative cardiovascular complication after a varicella zoster virus (VZV) infection is stroke, which occurs as VZV invades cerebral arteries and induces VZV vasculopathy[30]. Even though causality between HZ and stroke has been suggested[31], the adequate adjustment for potential confounding factors, such as myocardial infarction, transient cerebral ischemic attack and antithrombotic medications, etc. is challenging in observational studies[32]. In the SCCS-study[33] included in this thesis, we used the SCCS design to investigate an increased risk of stroke within 3–4 weeks after an HZ infection, adjusted for both, time-variant as well as time-invariant confounding factors[33].

The examples of complications (e.g., AGWs as an earlier outcome for cervical cancer, stroke after HZ infection) caused by infectious agents (e.g., HPV, VZV) show the large potential and novel opportunities for primary prevention interventions aiming to reduce the burden of infectious diseases.

1.3 Vaccination as the Primary Prevention of Vaccine-Preventable Diseases

Effective immunization programs/schedules need to vaccinate a sufficiently high proportion of susceptible persons to gain herd immunity[34]. Furthermore, such vaccination strategies need to boost vaccination rates before immunity drops below a protective level to prevent potential outbreaks[35].

There are two types of vaccines to combat vaccine-preventable diseases: Live attenuated vaccines and inactivated vaccines[36,37] (Table 1).

Table 1: Overview of different types of vaccines with examples and respective contents.

Types of vaccines	Examples of vaccines against specific vaccine-preventable diseases	Vaccine contents
Live attenuated vaccines	E.g., measles, mumps, rubella, varicella, rotavirus, herpes zoster (Oka/Merck), influenza (nasal spray), polio (oral) or yellow fever vaccine	Derived from weakened or altered disease-causing pathogens, that have been processed under laboratory conditions
Inactivated vaccines	E.g., polio vaccine (injection)	Made from microorganisms which have been killed by heat or chemical processes
	E.g., diphtheria and tetanus	Contain inactivated toxins
	E.g., influenza (injection), human papillomavirus, herpes zoster (HZ/su; GlaxoSmithKline (GSK), pertussis or pneumococcal vaccine for children	Contain only segments of the pathogens

While live attenuated vaccines are often able to induce a strong and long-lasting immune response, a reversion to virulent wild-type strains may occasionally occur—especially in immunocompromised patients—causing the disease[38]. One example for a very rare but possible reversion to the more virulent profile of wild virus is the oral polio vaccine (OPV)[39]. However, while inactivated vaccines cannot exhibit reversion to virulence they tend to produce a weaker immune response and therefore require multiple booster injections. Furthermore, they contain immunologic adjuvants to enhance vaccine efficacy, but they are also suitable for immunocompromised people[40].

In the case of a new inactivated herpes zoster vaccine (HZ/su; GSK vaccines), efficacy of 97.2% against HZ[41] vs. 51.3% of the live attenuated zoster vaccine Zostavax[42] throughout all age groups has been reported and is suggested to be due to the new adjuvant system ($AS01_B$) of the inactivated vaccine[41]. Adjuvants, in turn, are often suspected of causing adverse vaccine reactions[43–45] (e.g., the autoimmune/inflammatory syndrome induced by adjuvants (ASIA))[46]. While all licensed adjuvant vaccines have shown a favorable benefit-risk ratio[47], the choice whether a vaccine should be recommended within a particular target population is crucial.

The following chapter will shortly introduce the work of the German Standing Committee on Vaccination [Ger.: Ständige Impfkommission (STIKO)] regarding the decision-making process to recommend vaccines for immunization schedules in specific target groups. Furthermore, by evaluating vaccines and their impact on the population level, the next chapter will also provide information on direct and indirect vaccination measures as well as examples for debates on adverse vaccination reactions which might impede vaccine uptake.

1.3.1 Vaccine Recommendation by the German Standing Committee on Vaccination [Ger.: Ständige Impfkommission (STIKO)]

The World Health Organization (WHO) recommends several routine vaccination schedules globally. However, depending on the burden of disease in the population, the availability of an efficacious, effective and safe vaccine, economic factors and the level of priority that is placed on the vaccine-preventable disease, the immunization focus varies from country to country. Therefore, national vaccination schedules are developed, too.

In Germany, the national vaccination schedule is developed by an independent committee of experts called STIKO, which was established in 1972[48]. STIKO investigates the individual risk-benefit ratio, the population-based epidemiology as well as effects of a nationwide vaccination schedule and defines criteria for the assessment of adverse vaccination events. Together with the Protection against Infection Act [Ger.: Infektionsschutzgesetz (IfSG)], STIKO was incorporated into German law in 2001. Vaccination is not compulsory in Germany. However, STIKO recommendations are considered the medical standard. Furthermore, since 2007, vaccines recommended by STIKO according to section (§) 20 (3) of IfSG have been the basis for the immunization guidelines of the Joint Federal Committee (G-BA) and reimbursement is—by inclusion in the guidelines—compulsory for statutory health insurance providers (SHIs) in Germany.

Before the recommendation of a new or updated vaccine, STIKO reviews all available data about the vaccine from clinical trials and other available (observational) studies. The STIKO experts conduct a risk-benefit evaluation by investigating the individual benefit of vaccinated persons. Furthermore, population-level benefits due to herd immunity but also negative effects of vaccination programs/schedules (e.g., pathogen strain replacement phenomena, age shift of the disease burden) or cost-effectiveness analyses are evaluated.

For key questions regarding vaccine efficacy, effectiveness and safety, systematic reviews are conducted according to PICO questions (patient, intervention, comparator, outcome) and are ranked according to the GRADE methodology (grading of recommendations assessment, development and evaluation)[48].

For key questions related to e.g. burden of disease or vaccine acceptance systematic and/or exploratory literature searches are conducted, too.

Based on the results and quality of the available evidence obtained from data and information identified from the literature, STIKO debates all relevant key questions and criteria and finally decides whether a recommendation is given. New or updated recommendations are usually published once a year (usually in August) in the national epidemiological bulletin.

1.3.2 Direct Vaccination Effects: Vaccine Efficacy and Effectiveness

The public health importance of a vaccine is, on the one hand, related to the direct protective effect for the person receiving the vaccine and, on the other hand, to the indirect effects for others in the population (chapter 1.3.3). Population-based observational studies provide a valuable approach to measure vaccine effectiveness[49–51].

Vaccine Efficacy

Vaccine efficacy is defined as the percentage reduction of the vaccine-related disease among patients vaccinated according to the immunization schedule compared to unvaccinated patients (placebo group)[52]. This is usually conducted in a randomized controlled trial (RCT) "per protocol" with strict inclusion and exclusion criteria (i.e., excluding persons who are not included in the recommended schedule) as well as close monitoring to investigate the biologic performance capacity of the vaccine in a controlled environment under optimal conditions with usually healthy participants and 100% vaccine uptake[53]. Due to randomization and allocation concealment, RCTs have high internal validity, and bias that could lead to invalid study results is minimized. This, however, often hampers generalizability[54]. As this thesis does not include vaccine efficacy studies, these will not be described in more detail. They have only been mentioned to point out the differences to vaccine effectiveness studies described in the following.

Vaccine Effectiveness

Due to the fact that the application of a placebo vaccine could place the persons in the placebo group at risk of serious complications[55], it would be unethical to perform placebo-controlled RCTs after a vaccine has been introduced into a population. Vaccine effectiveness is therefore usually assessed in observational studies (e.g., cohort studies, case-control studies, nested case-control studies)[49,52]. It measures—similarly to vaccine efficacy—the percentage reduction in the incidence rate among vaccinated and unvaccinated persons but under real-world conditions; namely, in much larger populations over longer periods of time, within different healthcare systems, with concomitant drug use and vaccination application, including pregnant women, healthy children (i.e., no premature children), older persons or persons with underlying medical conditions which is usually not the case in RCTs. However, occasionally some RCTs also include some of these vulnerable population groups, e.g., RCTs on vaccine effectiveness and safety during pregnancy for mother and fetus.

One can compare vaccine effectiveness of a full vaccine series versus an incomplete vaccine series or no vaccination and measure effectiveness of a pathogen-specific outcome (e.g., incidence of clinically defined infections, laboratory-confirmed infection or vaccine-serotype diseases)[49].

Administrative data can be very helpful as prompt evaluation of vaccine effectiveness of, e.g., HZ vaccine effectiveness[56] or pneumococcal polysaccharide vaccine effectiveness[57] is important and these data have the advantage of being timeless and efficient[58]. However, to evaluate vaccine effectiveness can be also very challenging as is the case with influenza vaccine effectiveness due to the antigenic drift and shift of the virus genome. Consequently, vaccine effectiveness can vary from season to season and needs to be monitored accordingly.

The nested case-control study design and the case-control study design can and already have been used to evaluate the effectiveness and safety of licensed vaccines after introduction in public health programs or vaccination programs/schedules[50,51,59–61]. Since this thesis includes a nested case-control study (using risk-set sampling) that investigated drug safety issues[62], this study design will be discussed with a transferred focus on the methodological challenges (i.e., misclassification and selection bias) within the scope of influenza vaccine effectiveness[63,64] based on administrative data (chapter 3.1.2 and 3.2.1). Furthermore, alternative study design options will be discussed that are more suitable to investigate influenza vaccine effectiveness but potentially need additional data by linkage with other data sources (chapter 3.1.3).

Since vaccine effectiveness studies can evaluate the vaccination strategies on the individual level, the respective results help to optimize vaccine uptake and guide political decision processes by, e.g., identifying preferred vaccine product classes or by stimulating the improvement of vaccines[65]. However, vaccine effectiveness studies cannot evaluate the impact of a vaccination program/schedule within the population.

1.3.3 Indirect Vaccination Effects: Herd Immunity Effects Resulting from Vaccine Uptake and Coverage - Measured by Vaccine Impact Studies

After a vaccination program/schedule has been implemented on the population level, it is important to monitor the impact of the vaccine on the target disease. This way, the effectiveness and the potential benefit of the vaccine can be quantified.

Vaccine Uptake and Coverage

Vaccination coverage is used as an indicator for the success of vaccination strategies and is helpful for vaccine effectiveness and adverse vaccination signal interpretation[66–68]. Coverage estimates are usually measured by routine administrative data, sources from immunization information systems, vaccination cards or coverage surveys. The estimation of the percentage vaccination coverage relying on administrative data is based on the number of vaccinated persons during a specific period (numerator information) divided by the total

number of persons eligible for vaccination (denominator information)[67]. Figure 1 shows the immunization coverage for some important vaccine-preventable diseases in Germany.

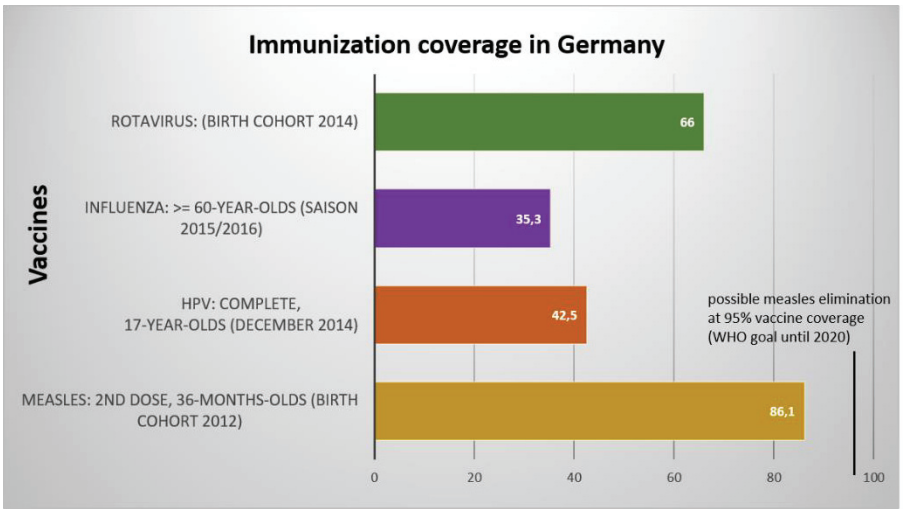


Figure 1: German immunization coverage for measles, human papillomavirus (HPV), influenza, and rotavirus vaccine (reference source: epidemiological bulletin 01/2017)[69].

In Germany, data from surveys or vaccination cards are the sole primary data source for vaccine coverage estimation[68]. However, one study compared external validated vaccine coverage estimation from administrative health insurance data with survey data and found good agreement between the data sources, meaning that analysis of administrative health insurance data can be a useful data source for the monitoring of vaccination coverage or vaccine uptake studies already shortly after a new vaccine has been introduced or for the investigation of the general acceptance of already recommended vaccines[70].

In one study[71] included in this thesis, an estimation of the vaccination coverage rather than the vaccine uptake was not possible, since knowledge about the vaccination status before the study period was not available. Consequently, vaccine uptake has been calculated by dividing the number of females who received at least one dose of HPV vaccine by the number of insurees in the respective age group, demonstrating 32.2% uptake in the recommended age groups (12–26 years) with a peak uptake in females aged 14–16 years[71], which is low compared to other countries[72–74].

Monitoring vaccine coverage is highly important in order to identify areas and groups with low vaccination coverage or high drop-out rates between the first and final dose[75]. Furthermore, it helps to understand reasons for low coverage, so that public health departments, healthcare partners and schools can help to improve vaccination coverage by increasing acceptance and vaccine uptake. Moreover, the potential for outbreaks of vaccine-preventable diseases can be identified.

Monitoring vaccine coverage is not only relevant for the targeted diseases. Introduction of one vaccine can also affect the coverage of other vaccines[76]. On one hand, there may be

catch-up vaccinations which could increase coverage of all routine vaccines. On the other hand, rumors about the safety of a new vaccine could lead to a reduced overall immunization coverage as parents may refuse vaccination of their children[68].

Vaccine Impact

Vaccine impact is defined as the reduction in incidence of the disease in a population targeted by a vaccination strategy[15]. For the estimation of vaccine impact, it is necessary to compare populations with and without an introduced vaccination program/schedule or to compare disease incidence attributable to a new vaccination program/schedule within the same population in the years before and after the intervention ("before/after studies") usually conducted in ecologic studies using interrupted time-series analysis[50,77–80].

One study included in this thesis assessed the HPV vaccine impact by investigating the incidence of anogenital warts (AGWs) before and after HPV vaccine introduction[81] suggesting herd immunity effects among males of similar age groups as that of the vaccine-recommended age groups of females.

Herd Immunity

Herd immunity implies that a rather high level of immunity in a community (vaccine coverage) can serve to protect susceptible (unimmunized) persons[34,82]. Such a collective immunologic protection represents an indirect effect of a vaccine exceeding the individual protection and impacting the population as a whole. The level of population immunity which is needed to interrupt transmission is defined as the herd immunity threshold (HIT) which is given by 1-(1/R0)[83]. R0 as the basic reproduction number is the number of secondary cases which one case would generate on average over the course of its infectious period in a completely susceptible population[84]. In general, the larger the value of R0, the harder it is to control an epidemic. The basic reproduction number is also affected by the duration of the infectious period, the infectiousness of the organism and the number of susceptible people in the population[85]. A vaccination program/schedule does not need to achieve 100% coverage to provide herd immunity against diseases in a community, e.g., to eliminate *Haemophilus influenzae* type b (Hib), a vaccine coverage of less than 70%[86,87] is sufficient. However, in the case of measles, a vaccine coverage of approx. 95% is necessary for the elimination of the disease[88] (Figure 1). If the incidence reduction in vaccinated and unvaccinated persons is higher than the coverage level, there is a strong indication for herd immunity effects. This also applies to incidence reduction in age groups or sexes besides the target vaccination group as has been seen in the HPV-study[81].

For sexually transmitted infections (STIs), a high level of vaccine-induced immunity in one sex can induce herd immunity in the other sex, but this also depends on a complex combination of other factors (e.g., vaccination coverage and sexual behavior)[89–91]. Due to

assortative mixing patterns, STIs will be more concentrated among high-risk persons as these have the greatest potential to infect others[89]. Therefore, in certain populations, high-risk persons of both sexes require immunization by vaccination to gain herd immunity effects[89,91].

A decrease of disease incidence in all age groups may result in an increase of the mean and median age of first infection, which in turn can cause more serious complications if infection takes place later in life (e.g., rubella infection during pregnancy)[92]. Thus, it is necessary for vaccination programs/schedules to maintain a high coverage in subsequent birth cohorts as numbers of susceptible persons will accumulate in older age groups.

Only contagious diseases transmissible among humans can be eradicated by herd immunity due to reduced transmission from one person to another[93]. Therefore, herd immunity cannot be achieved if infectious agents can also use hosts other than human individuals (e.g., tetanus (*Clostridium Tetani*) or tick-borne encephalitis (TBE) [Ger.: Frühsommer-Meningoenzephalitis (FSME-virus)]).

1.3.4 Vaccine Safety Studies of Adverse Vaccination Reactions after Marketing Approval

Potential safety concerns regarding vaccines can have a large effect on vaccine uptake and coverage. Once its efficacy and safety have been demonstrated in a multi-stage test procedure (Phase I–III), is a new vaccine licensed by the European Medicines Agency (EMA) Committee for Medical Products for Human Use (CHMP) for the European Union (EU) and respectively by the German Federal Agency for Sera and Vaccines—the Paul-Ehrlich-Institute (PEI)—for Germany.

As no vaccine is perfectly safe or effective, potential risks accompany its benefits. Vaccines are given to healthy persons, which is why tolerance of adverse reactions is even lower than for pharmaceuticals which are mostly given to ill persons for curative purposes. Serious so-called adverse vaccine reactions (AVRs) after vaccination are very rare[94]. However, vaccine safety concerns attract great attention[95,96]. For AVRs, special reporting obligations apply: According to § 6 (1) of the IfSG, already the suspicion of an AVR which exceeds the usual level of an inoculation reaction is subject to reporting. Furthermore, confirmed and suspicious AVRs are reported to the PEI directly or via the local Health Office. The PEI maintains a database that includes both suspicious messages and confirmed cases of adverse reactions associated with vaccinations. Potential risks are published and, if necessary, further investigated in clinical and epidemiological studies. Although AVRs from spontaneous reporting provide warning signals of risks and can function as a useful tool to provide hypotheses of unknown AVRs, it is not possible to determine the risks and frequency of specific AVRs due to underreporting and missing information on the vaccinated individuals

necessary for denominator assessment[97]. Therefore, vaccine safety issues need to be investigated by epidemiological observational studies in a real-world setting as knowledge on the benefit-risk profile during the approval process is limited[98]. These studies play an important role for a sound evidence base regarding population-based vaccine effects of rare and severe AVRs once a vaccine has been introduced to the population[99].

Awareness of background incidences of possible vaccine-related adverse events can provide valuable information to investigate whether a disease/adverse event occurs after vaccination at disproportionally high rates[100] as AVRs do not exhibit specific characteristics but manifest themselves like any other newly occurring disease. Most serious AVRs are rare and it is challenging to investigate and differentiate causality and temporal association by chance[101]. As already mentioned, spontaneous reporting systems can only "signal" emerging AVRs. If epidemiological data indicate causality of severe AVRs, the vaccine will be removed from the market (e.g., in Germany: FSME vaccine TicoVac in 2001[102], Hexavac® in 2005[103]).

Unfortunately, misleading reports on AVRs may result in a decline or disruption of attendance to vaccination programs/schedules[101]. One example is the false association of the MMR vaccine with the risk of autism which resulted in decreasing uptake of measles vaccine[104–106].

Another example is the erroneously suggested relationship of a teenage death shortly after HPV vaccination in Austria. However, due to the available background rates of sudden deaths in adolescents in Austria, the HPV vaccination program could be continued[101]. During 2015, there were case reports of the Complex Regional Pain Syndrome (CRPS) and Postural Orthostatic Tachycardia Syndrome (POTS) occurring in young women after HPV vaccination[107,108]. However, in November 2015, the European Medicines Agency's Pharmacovigilance Risk Assessment Committee (PRAC) investigated pre- and post-licensure data as well as background incidences of the diseases and concluded no evidence to support causality[109–111].

Guillain-Barré syndrome is another good example of an AVR as several large-scale vaccination programs/schedules were associated with an increased number of cases of Guillain-Barré syndrome[112]. The assessment of such cases before and after vaccination has therefore been of high priority[113].

The inappropriate assessment of vaccine safety data or the non-scientific social-media attention (e.g., on Facebook, etc.) on single individual cases suffering from AVRs could severely distort and undermine the impact of mass vaccination campaigns. In the case of HPV vaccination, an enormous public health debate has been dominated by safety concerns resulting in decreasing vaccination coverage and thus ignoring the fact that an effective vaccine is available against cervical cancer.

Post-marketing observational studies on vaccines based on administrative data can contribute to the monitoring of direct and indirect vaccine effects, as well as of AVRs and boost trust in vaccines' impact and safety.

1.4 Data Sources for Clinical Epidemiology of Vaccine Effectiveness and Safety and on Vaccine-Preventable Diseases

1.4.1 Primary and Secondary Data

Primary as well as secondary data can be used as data sources for epidemiological studies on vaccines and vaccine-preventable diseases[70].

In Germany, primary data on vaccination uptake and coverage and on the immune status have been specifically collected in health surveys for the purpose of the respective study and research question. Solely or combined with secondary data, they provide a valuable contribution to monitoring the coverage or uptake of vaccines and to the evaluation of compliance with vaccination schedules in Germany. A comparison and overview of characteristics and vaccine status assessment of primary and secondary data for studies on vaccines and vaccine-preventable diseases is given in Table 2.

In Germany, there are various primary data sources for studies on vaccines, i.e., primary data from surveys such as kindergarten entrance examinations (KEEs) that are performed merely in a few federal states, population surveys (e.g., the population-based German Health Interview and Examination Survey for Children and Adolescents (KiGGS) or the German Health Interview and Examination Survey for Adults (DEGS)) or school entrance examinations (SEEs)[70].

Secondary databases can be divided into medical record databases and administrative claims databases[114]. While data from medical record databases are derived from electronic patient files, administrative claims databases contain data obtained by health insurance providers or state-financed healthcare systems mostly in the field of healthcare services and quality control[98,115] or other administrative purposes[116]. Therefore, it may not contain all information of interest[115]. The information available in both databases is mostly similar, even though medical record databases often additionally contain information on lifestyle factors (smoking, body mass index (BMI) or alcohol consumption) which is missing in administrative claims data.

Commonly in administrative claims data, information on drug prescriptions as well as outpatient and inpatient data are reported to the patient's health insurance provider for reimbursement purposes. Additionally, the available information on sociodemographic information can be used for the investigation of the respective research questions. As an

advantage, persons who are rarely investigated in RCTs (older or immunocompromised persons, severely diseased or deceased persons, infants or pregnant women) or are difficult or impossible to reach through field studies are included in these data[98]. For these population groups, vaccine effectiveness, impact and safety studies conducted under real-world conditions based on secondary data can provide important information. In addition, since secondary data are already available, analysis can usually be performed promptly and cost effectively[98]. The validity of data related to applied vaccinations is high since the reimbursement of the applied vaccine is directly related to the deduction of the vaccination [Ger.: Impfziffer][117]. However, some limitations need to be mentioned for studies on vaccines with German claims data. This topic will be discussed in more detail within the context of the German Pharmacoepidemiological Research Database (GePaRD) in the next chapter. GePaRD—as an administrative claims database—represents the basis of the individual research articles included in this thesis.

Table 2: Comparison of primary and secondary data sources for studies on vaccine effectiveness and safety and on vaccine-preventable diseases in Germany.

Characteristics	Primary data	Secondary data
Data sources	- KEE, SEE, KiGGS, DEGS	- Obtained from data sources with a different original purpose, e.g., by health insurance providers collected for reimbursement purposes (claims databases) - Electronic medical records from general practitioners or hospitals (medical record databases)
Vaccination status assessment	- Measured, documented or remembered a*d hoc* by the respective individual	- Information on vaccination status and coverage, e.g., from reimbursement of vaccination (necessary detail may not be available, e.g., differentiation of different vaccine types)
Information on lifestyle-related factors (e.g., smoking, alcohol intake, etc.)	- Measured, documented or remembered *ad hoc* by the respective individual	- Partly available in medical record databases, not in claims databases
Vaccination program and campaign assessment	- *Ad hoc* knowledge and attitudes towards vaccination and programs/schedules (KiGGS, DEGS)	- Compliance with vaccination schedules and immunization programs
Time and cost efficiency	- *Ad hoc* data collection requires time and personal/monetary resources	- Data are mostly promptly available and inexpensive
Individual immune status	- Supplied by vaccination cards and serological tests (KiGGS,	- Partly available in medical record databases, not in claims

Characteristics	Primary data	Secondary data
	DEGS)	databases
Regional coverage comparisons	- Vaccination coverage assessment is possible for all recommended childhood vaccinations for a specific cohort of children (SEEs) - Population-based data of selected vaccinations by age or sociodemographic status available (KIGGS, DEGS)	- Population-based coverage data of recommended and reimbursed vaccinations can be assessed by age (data available since the time data collection started, e.g., associations of statutory health insurance physicians (ASHIP) data since 2004)
Trend analyses	- Possible for complete cohorts by all recommended childhood vaccinations	- Follow-up of birth cohorts and vaccination incidence - Cumulative vaccination incidence by a defined age and year - Trends over time
Disease incidence estimation	- Not possible in surveys	- Estimation of disease incidence possible
Vaccine safety assessment	- Generally not possible (sometimes surveys ask for AVRs)	- Estimation of rather acute, late and rare events possible
Bias	- Recall bias possible - Interview bias possible - Selection bias (due to non-responders) possible	- Recall bias not possible - Interview bias not possible - Selection bias (due to non-responder) not possible

Abbreviations: Associations of Statutory Health Insurance Physicians (ASHIPS), kindergarten entrance examination (KEE), school entrance examination (SEE), German Health Interview and Examination Survey for Children and Adolescents (KiGGS), German Health Interview and Examination Survey for Adults (DEGS).

1.4.2 The German Pharmacoepidemiological Research Database (GePaRD)

The German Pharmacoepidemiological Research Database (GePaRD), which was established by and is maintained at the Leibniz Institute for Prevention Research and Epidemiology – BIPS, currently consists of data from four statutory health insurance providers (SHIs) with a total of more than 20 million insured persons over all data collection years[118]. Two SHIs are operating nationwide and the other two, smaller ones in Bremen and in the Northwest German region. GePaRD represents nearly 17% of the German population throughout all federal states in Germany[118]. In addition to sociodemographic core data, GePaRD includes outpatient care data and inpatient diagnoses and procedures as well as prescription data of all individuals enrolled in one of the four SHIs since 2004. The description of the structure of GePaRD and contained information due to the linkage with the central pharmaceutical reference (CPR) database is displayed in Figure 2. Diagnoses are coded according to the International statistical classification of diseases and related health problems, 10th revision, German modification (ICD-10-GM). Inpatient data comprise

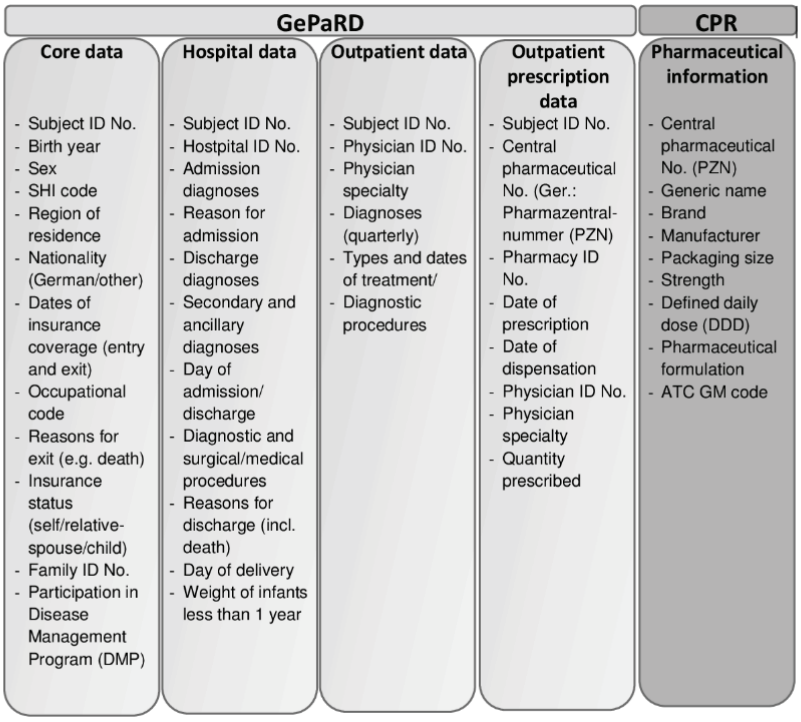
diagnostic and therapeutic procedures as well as outpatient surgeries that are represented via the Operation and Procedure Key (OPS). Further information contained in inpatient data is the date of admission and discharge (incl. the reason for hospital discharge/death) as well as different types of diagnosis (admission, main discharge, secondary and ancillary diagnosis).

Outpatient data comprise types and dates of diagnostic and therapeutic procedures and diagnoses on a quarterly basis including the diagnostic certainty (certain, suspected, excluded, and status post diagnosis). Outpatient therapeutic services are coded according to the physician's fee schedule [Ger.: Einheitlicher Bewertungsmaßstab (EBM)] that defines the content of and the payment for each outpatient service. Outpatient prescription data include information on prescriptions dispensed in a pharmacy and reimbursed by the respective SHI. Further information included in prescription data are the dates of the prescription and dispensation, the number of prescribed packages, the specialty of the prescribing physician, and the central pharmaceutical number [Ger.: Pharmazentralnummer (PZN)]. Via linkage with an internal reference database at BIPS further information such as the defined daily dose (DDD), information on generic and brand name, packaging size, strength and the Anatomical-Therapeutic-Chemical (ATC) code, can be added. But information on prescribed daily doses or the planned duration of therapy is not included.

GePaRD data can only be used after the respective research project has been officially approved by the contributing SHIs and their governing authorities[98,118].

For studies on vaccines, some particularities apply: standard vaccinations are not prescribed to individual patients but included in the so called "Sprechstundenbedarf" (medical products for use at a physician's office only) which cannot be identified in the database. Via outpatient EBM codes that are used for reimbursement of, e.g., administration of recommended vaccines, the application of vaccines can be identified. However, before 2008, there was no uniform EBM coding and EBM codes for administered vaccines varied in 2007/2008 between the regional Associations of Statutory Health Insurance Physicians (ASHIPs). There were also differences among the SHIs in the doses and over time as every regional ASHIP could negotiate individual vaccine agreements [Ger.: Impfvereinbarungen] with the different SHIs. Additionally, EBM codes related to vaccinations are updated (and consequently changed) frequently. Since July 2008, uniform EBM codes for all ASHIPs have been introduced for the identification of the different administered vaccines and for the identification of completed or incomplete vaccine schedules[119]. However, in reality, the new uniform EBM codes were not introduced in all ASHIP regions directly. Furthermore, the EBM codes do not differentiate between different kinds of vaccines if these are administered in the same way. For example, the bivalent (Cervarix®) and the quadrivalent (Gardasil®) HPV vaccination have the same EBM code in Germany. The same is true, e.g., for certain pneumococcal vaccines and

subcutaneous influenza vaccines. Resultingly, the distinction between those vaccines based on administrative claims data is impeded. In addition, in the case of the HPV vaccine, during the introduction period of specific EBM codes, only a few federal states had specific EBM codes for HPV vaccinations in females older than the recommended age groups (> 17 years) as only some SHIs reimbursed the vaccine for these older females. Consequently, not all vaccinations in older females can be identified across Germany for these years.



* Hospital and outpatient diagnoses are coded according to the International statistical classification of diseases and related health problems, 10th. revision, German modification (ICD-10-GM) with at least 4 digits

** Inpatient data contain different types of diagnosis (admission, main discharge, secondary and ancillary diagnosis)

*** Outpatient diagnoses refer to a period of three months, as physicians' services are settled quarterly. Outpatient data comprise types and dates of diagnostic and therapeutic procedures, and diagnosis on a quarterly basis including the diagnostic certainty (certain, suspected, excluded, and status post diagnosis)

Inpatient data comprise diagnostic and therapeutic procedures as well as outpatient surgery that are presented via the Operation and Procedure Key (OPS)

Outpatient therapeutic services are coded according to the uniform assessment standard [Ger.: Einheitlicher Bewertungsmaßstab (EBM)] that defines the content of and the payment for each outpatient service

Anatomical-Therapeutic-Chemical (ATC) Classification System, German modification

$ Provided to SHIs by hospitals

$$ Provided to SHIs by regional associations of statutory health insurance physicians [Ger.: Kassenärztliche Vereinigung]

$$$ Provided to SHIs by pharmacies' electronic data processing centers [Ger.: Apothekenrechenzentren]

Figure 2: Structure and content of GePaRD and of the CPR.

1.5 Objectives

The individual research articles included in this thesis are all based on administrative claims data from health insurance providers contained in the GePaRD database. These studies aimed (i) to investigate the burden of vaccine-preventable diseases and complications to extend the data basis for a vaccine recommendation decision process, (ii) to evaluate vaccine uptake and the impact of an already introduced and recommended vaccine in the general population and (iii) to apply a nested case-control study design to its potential use for vaccine effectiveness and safety studies.

The following objectives were addressed in this thesis:

1. to estimate data on the disease burden of herpes zoster, its complications and postherpetic neuralgia throughout all age groups and stratified by immune status,

2. to assess the risk of stroke as a major complication after the onset of herpes zoster, to investigate the risk of stroke subtypes, the role of herpes zoster location and the time interval between herpes zoster onset and stroke,

3. to describe the HPV vaccine uptake in Germany after reimbursement of the vaccine on a broad regional level,

4. to assess potential vaccination herd immunity effects in a low HPV coverage setting in Germany after the HPV vaccine recommendation and

5. to apply the method of a nested case-control study design with risk-set sampling in a defined cohort to estimate population-based risks of acute myocardial infarction (AMI) for individual and widely used nonsteroidal anti-inflammatory drugs (NSAIDs). Although this study design was not applied to a vaccine-related topic, it is a valuable method for vaccine-related studies as discussed later.

Based on the conducted research articles, various methodological aspects will be described and discussed in the context of clinical epidemiology of vaccines and vaccine-preventable diseases on the basis of administrative claims data.

2 EXECUTIVE SUMMARY OF RESULTS OF INDIVIDUAL RESEARCH ARTICLES

This chapter summarizes the most important results of the individual research articles preceded by a brief introduction of the respective viral disease.

2.1 Burden of Infectious Diseases and Potential Risk of Cardiovascular Diseases after an Infection (Using the Example of Herpes Zoster)

Herpes zoster (HZ) is a viral disease caused by an endogenous reactivation of the varicella zoster virus (VZV)[120]. Primary infection usually occurs in childhood and causes varicella (chickenpox). Afterwards, the virus remains latent within sensory dorsal roots of the cranial and spinal ganglia. More than 95% of all German adults are VZV immunoglobulin G (IgG) antibody positive[121]. Upon a decrease in cell-mediated immunity, the virus can be reactivated decades later and manifests as herpes zoster (shingles)[122]. The lifetime risk of the development of an HZ episode is estimated to be 33% and increases considerably with age[123,124].

There are different manifestations of HZ depending on the affected dermatome. The thorax is mostly affected (50–56% of cases), followed by the face (20% of cases). Other neurological manifestations are, e.g. zoster encephalitis, zoster meningitis or HZ with other neurological system involvements[125]. In addition, there are also various HZ-related complications. The main complication of HZ is PHN, which occurs in 10–20% of HZ cases, but predominantly in patients older than 60 years[126,127]. It is defined as a long-lasting, occasionally recurring pain for 1–3 months after the HZ-related onset rash[128]. Another serious complication of both, primary infection (chickenpox) and VZV reactivation (shingles) is ischemic and hemorrhagic stroke, which occurs by VZV vasculopathy and affects both immunocompetent and immunocompromised individuals[129]. VZV is thought to spread transaxonally along afferent nerve fibers from the cranial nerve ganglia to arteries of the anterior or posterior cerebral circulation causing inflammation, thrombosis and occlusion of the vessels. In addition, VZV vasculopathy can lead to transient ischemic attacks, arterial ectasia, aneurysm and subarachnoid hemorrhage.

The risk of developing HZ, its manifestations or HZ-related complications, e.g., PHN or vasculopathy, is increased in older and immunocompromised individuals. Systemic antiviral therapies are necessary and should be used as early as possible after the onset of symptoms. Especially in these population groups, the current pharmacological treatment options of HZ-related complications are challenging as they might lead to higher complication

rates[130]. Hence, to boost the VZV-specific cell-mediated immunity, a live attenuated VZV vaccine against HZ has been developed. However, to date (March 2018), STIKO has not recommended routine VZV vaccination with the live attenuated VZV vaccine based on their scientific evaluation of safety, efficacy and effectiveness data[131,132]. Furthermore, there is also an inactivated vaccine against HZ, which could also be administered to immunocompromised persons, but the vaccine has not yet been approved by EMA. Given an aging population and the fact that more than 95% of the German population is infected with VZV, the disease burden of HZ infection is supposed to increase.

2.1.1 Incidence of Herpes Zoster and its Complications

OA1 Incidence of herpes zoster and its complications in Germany, 2005-2009.
Hillebrand K, Bricout H, Schulze-Rath R, Schink T, Garbe E. *Journal of Infection* 2014; 70(2):178-186. [https://doi.org/10.1016/j.jinf.2014.08.018]

Methods: In this burden of diseases study, we estimated incidence rates (IRs) of HZ, its manifestations and complications overall, by age, sex and immune status to extend the data on the burden of HZ disease. The proportion of PHN, hospitalizations, the diagnosing physician specialty and systemic antiviral therapy were also assessed. Based on administrative claims data from 2005–2009, a retrospective cohort study including 7 million statutory health insurance members was conducted.

Results: Between 2006 and 2009, the annual standardized IRs ranged between 5.3 and 5.5 per 1000 person-years (PY). Females had higher IRs than males throughout the study period. IRs increased more than threefold (up to 15 per 1000 PY) in the age group 80–84 compared to those 10–44 years of age. About 72% of HZ patients had no complications in 2009, while about 16% suffered from nervous system involvement. During the other respective study years, these numbers were similar. The age-related increase of IRs was higher for HZ complications than for uncomplicated HZ. Immunocompromised patients suffered slightly more complications than immunocompetent patients. The annual PHN proportion among HZ cases increased from 12% in 2005 to 15% in 2009 with a steady age-related elevation for both sexes. About 3% of HZ cases were hospitalized. More than 50% of HZ cases were diagnosed by general practitioners (GPs) and 71% of HZ cases received systemic antiviral treatment.

Conclusion: The presented baseline data contribute to an enrichment of empirical data on the burden of HZ disease. Furthermore, these data can be used for future vaccine impact studies.

2.1.2 Risk of Stroke after Herpes Zoster Infection

OA2 Risk of Stroke after Herpes Zoster- Evidence from a German Self-Controlled Case-Series Study.

Schink T, Behr S, Bricout H, Thöne K, Garbe E. *PloS One* 2016; 11(11): e0166554. [https://doi.org/10.1371/journal.pone.0166554]

Methods: To assess the risk of stroke as a serious complication after HZ onset, we conducted a self-controlled case-series study on a cohort of patients with incident stroke. Furthermore, potential differences in risks between stroke subtypes (ischemic and hemorrhagic) and HZ location, as well as throughout different time periods between HZ onset and stroke were investigated. We used administrative claims data of the years 2004–2011 including 7.7 million statutory health insurance members.

Results: Within the cohort of 124,462 stroke patients, 6,035 (5%) had at least one HZ diagnosis identified either as a main hospital discharge diagnosis or as HZ treated with antivirals. In the three-month risk period after HZ onset, the risk for stroke was about 1.3 times higher compared to control periods (Incidence rate ratio (IRR): 1.29; 95% confidence interval (CI): 1.16–1.44). This was similar in magnitude for ischemic and unspecified stroke. The risk of hemorrhagic stroke was 1.5-fold higher compared to control periods. During the 3 months after herpes zoster ophthalmicus (HZO) onset, the effect on the risk of stroke was slightly higher, with an about 1.6 times higher risk compared to control periods (1.59; 1.10–2.32). During short-time periods (3–4 weeks after HZ onset), the risk for stroke was highest and decreased thereafter.

Conclusion: Our study found an increased risk of stroke after HZ with highest risk estimates 3–4 weeks after HZ onset. Stroke risk was slightly higher after HZO and was also higher for hemorrhagic stroke than for ischemic or unspecified stroke.

2.2 Preventive Public Health Interventions: Vaccine Uptake and Vaccine Impact Studies (Using the Example of the Human Papilloma Virus)

HPV infections are the most frequent sexually transmitted viral infections worldwide affecting both, men and woman[133]. They can result in malignant cancer or benign skin and mucosal tumors, including AGWs. More than 120 HPV types have been identified of which more than 40 affect mostly anogenital epithelium. HPV types 6 and 11 account for about 90% of AGWs; HPV types 16 and 18 are responsible for 70% of all cervical cancers as well as for a large proportion of other anogenital cancers[134]. In 2006, a quadrivalent vaccine against HPV 6, 11, 16 and 18 was approved by the FDA in the United States and EMA/CHMP for the prevention

of cervical cancer. In Germany, the HPV vaccination has been recommended for girls between 12 and 17 years of age by STIKO since March 2007 and since 2014, for 9-year-old to 14-year-old girls. The vaccine is free of charge for this age group. Additionally, some SHIs offer reimbursement for women aged 18–26 years. The bivalent vaccine against HPV types 16 and 18 (recommended since March 2007) as well as the nonavalent HPV vaccine against 6, 11, 16, 18, 31, 33, 45, 52, and 58 (recommended since April 2016) are also available in Germany, but to date the quadrivalent HPV vaccine strongly dominates the German market[135]. However, in 2017, the quadrivalent HPV vaccine was replaced with the nonavalent HPV vaccine.

HPV vaccine efficacy was estimated to be 90–100% for preventing persistent and incident HPV infections as well as AGWs[136]. To date, the follow-up time has been too short to estimate protection against cervical cancer, but AGW incidence rates can be an effective measure for earlier outcomes of HPV vaccine impact, as AGWs develop rapidly after HPV infection.

Recent studies in Australia, Europe and the United States reported an AGW incidence reduction of up to 90% in the vaccine-recommended age group[26]. Some studies also reported decreasing incidence in older age groups of females as well as in males, suggesting effects of herd immunity[74,137]. This would dependent on vaccine coverage and as most of these studies were conducted in countries with high vaccine coverage of 70% to 90%, almost no data on herd immunity effects are reported in countries with low vaccine coverage[72–74].

HPV vaccine uptake in Germany when based on at least one vaccine dose was low in 2008 (about 32.2% in 12- to 17-year-old females)[138]. In 2012, a similarly low vaccine uptake was reported ranging from 6.1% in 12-year-old females to 47.6% in 16-year-old females[139]. Compared to numbers of other countries with implemented HPV vaccination programs/schedules these numbers are low. Reasons for a limited vaccine uptake might be controversial discussions on vaccine effectiveness and safety which may have led to uncertainty among young women, their parents or also among physicians[140]. Furthermore, a school-based vaccination program which has led to higher vaccination rates in other countries is still missing in Germany.

2.2.1 HPV Vaccine Uptake in Germany

OA3 HPV vaccine uptake after introduction of the vaccine in Germany: an analysis of administrative data.

Hense S, Hillebrand K, Horn J, Mikolajczyk R, Schulze-Rath R, Garbe E. *Human Vaccines & Immunotherapeutics* 2014; 10(6):1729-33. [https://doi.org/10.4161/hv.28450]

Methods: To assess the HPV vaccine uptake in 2008 for females aged 12–26 years on a broad regional level in Germany, we conducted a retrospective cohort study with data from one large SHI including about 7 million statutory health insurance members (about 8.5% of the German population).

Results: The overall study population consisted of 317, 234 females, of whom 77,350 received at least one HPV vaccine dose in 2008. Vaccine uptake was 32.2% in the recommended age group of 12- to 17-year-old females and peaked at the age of 14–16 years (36.4%). Among older females aged 18–26 years, the HPV uptake was only 12.3%. HPV vaccination was not officially recommended by STIKO for this age group but the SHI which contributed data for this study offered reimbursement of the HPV vaccine also for this age group. In four federal states, about 66% received the vaccine in the recommended age, while about one third (33.9%) received the vaccine at older ages (18–26 years). Other states could not be investigated due to a lack of specific EBM codes for reimbursement of vaccine within this age group.

Conclusion: The HPV vaccine uptake in 2008 reflects an early status of HPV vaccine uptake after vaccine recommendation in 2007. For timelier monitoring, information on future changes in HPV uptake is needed. This will create a basis for prompt public health reactions and for efforts to adapt immunization programs adequately.

2.2.2 HPV Vaccination Herd Immunity Effects for Anogenital Warts

OA4 Evaluation of vaccination herd immunity effects for anogenital warts in a low coverage setting with human papillomavirus vaccine – An interrupted time series analysis from 2005 to 2010 using health insurance data.

Thöne K, Horn J, Mikolajczyk R. *BMC Infectious* Diseases 2017; 17(1):564. [https://doi.org/10.1186/s12879-017-2663-7]

Methods: With this vaccine impact study, we assessed potential vaccination herd immunity effects among males in a German low HPV vaccine coverage setting. A retrospective open cohort study with data from one large SHI including more than 9 million statutory health insurance members from 2005–2010 was conducted.

Results: About 5 million insurance members aged 11–79 years were included in the cohort for each study year. Overall, 49,214 incident AGW cases were identified. Overall incidence rates of AGWs were relatively stable throughout the study years. In all age groups, between the 1st quarter of 2005 and the 2nd quarter of 2007 incidence was approximately stable. Among 16- to 26-year-old females and 16- and 18-year-old males, incidence decreased

between the 2nd quarter of 2007 and the 4th quarter of 2008. Afterwards, the incidence stabilized at a lower level. In most other age groups, the incidence was relatively stable or increased slightly over the studied period. The incidence rate ratio of AGWs for the post-vaccination period (2009–2010) compared to the pre-vaccination period (2005–2007) showed a u-shaped decrease among the 14- to 24-year-old females and also among males which corresponds well with the reported HPV vaccination uptake in 2008[71]. Within the 16- to 20-year-old females, a reduction of up to 60% was seen while for the 16- to 18-year-old males, a reduction of up to 50% was observed. In younger females of 21–26 years of age, a reduction of 10–20% was seen. This was also found for males but estimates did not achieve significance. The reduction in the youngest age group of 12- to 15-year-olds was about 20–30% but with wide confidence intervals.

Conclusion: The slightly less pronounced relative reduction among males in approximately the same age group as females who received the HPV vaccination suggests herd immunity resulting from assortative mixing by age. However, the early decrease among males may be reduced over time due to partner change.

2.3 Nested Case-Control Study Design with Risk-Set Sampling (Using the Example of Non-Steroidal Anti-Inflammatory Drugs)

The nested case-control design is a method that consolidates advantages of both, a case-control study and a cohort study where cases and controls are sampled from a pre-defined cohort[97,141]. For each case, the risk-set is defined and one (or more) controls are randomly selected from each risk-set. This study design is sometimes used in vaccine effectiveness and safety studies, which will be described in detail within the discussion chapter 3.1.2. Within this thesis, the nested case-control design has been applied in the context of a drug safety study, the content of which will be shortly described in the following.

NSAIDs are among the most frequently used therapeutics in the general population[142]. They have a wide range of clinical indications, e.g. short- or long-term pain states, fever and a range of chronic inflammatory and degenerative joint diseases such as rheumatic arthritis and osteoarthritis[143]. For cyclooxygenase-2 (COX-2) selective NSAIDs, an increased risk of adverse cardiovascular events has been reported, resulting in the withdrawal of rofecoxib in 2004[144] and valdecoxib in 2005[145]. Consequently, COX-2 selective NSAIDs are contraindicated in the EU for patients suffering from coronary heart disease, cerebrovascular disease and peripheral arterial disease[146].

During the last decade, several European[147,148] and international[149,150] observational studies as well as meta-analyses[151–153] indicated an elevated risk of AMI for both traditional NSAIDs (tNSAIDs) and COX-2 selective NSAIDs. Several studies support findings that tNSAIDs may

increase the risk of heart failure or stroke in patients with or without heart disease or risk factors for heart disease[143,154]. However, estimates of the increased risk varied and little is known about the cardiovascular risk profile of tNSAIDs and COX-2 selective NSAIDs as well as the influence of concomitant drug use or co-morbidities.

OA5 Non-steroidal anti-inflammatory drug use and the risk of acute myocardial infarction in the general German population: a nested case-control study.

Thöne K, Kollhorst B, Schink T. *Drugs- Real World Outcomes 2017;* 4:127–137 [https://doi.org/10.1007/s40801-017-0113-x]

Methods: A case-control study nested in a cohort of 3,476,931 new NSAID users of the years 2004–2009 was conducted. Population-based AMI risks for individual and widely used NSAIDs, for the cumulative amount of NSAID use and for patients with and without a prior history of cardiovascular risk factors were estimated. Cohort members had to be continuously insured for at least 12 months before the first NSAID dispensation. In order to include only incident NSAID users, only patients without any NSAID prescription within these 12 months were included in the cohort. Cohort entry was the first NSAID dispensing date between 2005 and 2009. All patients were followed from their first NSAID dispensation in the study period until either interruption of insurance status for more than 3 days, termination of insurance including death, diagnosis of malignant cancer or the end of the study period/longest available follow-up in the database, whichever came first. All patients with a first hospitalization with a main discharge diagnosis of AMI or subsequent MI were identified as cases. The hospital admission day was defined as the index date of the case. Up to 100 controls were randomly matched by age, sex, SHI and length of follow-up using risk-set sampling. Length of follow-up was defined by assigning an index date to each control that resulted in the same duration of follow-up as the corresponding case. Patients might have served as controls for more than one case and were eligible to be selected as controls until they became a case. Exposure status was classified into current, recent or past users, where past users were used as reference. Multivariable conditional logistic regression was applied to estimate ORs and 95% CIs. Duration of NSAID use was calculated by the cumulative amount of dispensed DDDs and stratified analyses were conducted for potential effect modifiers.

Results: Overall, 17,236 AMI cases were matched to 1,714,006 controls. For the most frequently used NSAIDs, ibuprofen and diclofenac, a 40–50% increased risk of AMI (1.54, 1.43–1.65 and 1.43, 1.34–1.52, respectively) was observed. No association was seen for the COX-2 selective NSAIDs celecoxib and lumiracoxib but the number of current users was low and both confidence intervals included the null value. A low cumulative NSAID amount was

associated with a higher AMI risk for ibuprofen, diclofenac and indometacin. The risk associated with current use of diclofenac, fixed combinations of diclofenac with misoprostol, etoricoxib or ibuprofen was highest in the younger age group (<60 years) and similar for patients with or without major cardiovascular risk factors.

Conclusion: Among the 15 investigated individual NSAIDs, relative AMI risk estimates differed. The most frequently used NSAIDs—diclofenac and ibuprofen—were associated with a 40–50% increased relative risk of AMI, even for low cumulative NSAID amount. The relative AMI risk in patients with and without cardiovascular risk factors was similarly elevated.

3 METHODOLOGICAL ISSUES

As described in chapter 1, observational studies provide important methods to assess, e.g., the burden of vaccine-preventable diseases and their complications, to evaluate direct vaccination effects (e.g., vaccine effectiveness), indirect vaccination effects through herd immunity or the safety of vaccines in a real-world setting.

To investigate these different research topics, several study designs can be used, while the preferred design is always dependent on context and data source [155]. For example, cohort studies, case-control studies and case-only designs are commonly used for studies either on the burden of vaccine-preventable disease, for vaccine effectiveness and for vaccine safety assessment[156]. While nested case-control studies are also sometimes conducted to investigate direct vaccination effects and safety issues, vaccine impact can be best evaluated by using an ecologic time-trend study design with a segmented regression analysis. Furthermore, based on the underlying data source each design has advantages and disadvantages for the investigation of the respective research questions. However, each design differs in terms of its susceptibility to confounding and bias.

In this section, methodological aspects for studies on vaccines with regard to the underlying administrative data source will be discussed. First, the different study designs and analyses that were applied within the individual research articles included in this thesis are introduced and discussed with respect to their advantages and disadvantages (chapter 3.1). Second, examples of bias and confounding that might have occurred and that were addressed within the studies are discussed (chapter 3.2).

3.1 Study Designs and Analyses

3.1.1 Cohort Studies

Cohort studies are a valuable tool in order to investigate, e.g., vaccine-preventable disease burden, vaccine effectiveness or vaccine safety issues.

The general principle of a cohort study is to classify disease-free cohort members according to their exposure status (e.g., vaccination or no vaccination, dose schedules of vaccination (fully vs. one (or two) doses) or different brands of vaccines) and to follow them over time in order to examine the frequency of first (incident) outcome occurrence (e.g., vaccine-preventable target disease, AVR)[157].

Depending on the follow-up time in relation to the point in time when the study is conducted, cohort studies can be classified into prospective (concurrent) or retrospective (non-

concurrent or historical) studies. A retrospective cohort study can be assembled from data of, e.g., administrative claims data in the past and followed-up to the present (or the longest available follow-up). While a single-arm cohort study is sufficient to estimate incidences[21], a multi-arm cohort study is necessary for risk estimation. In a matched cohort study, patients are paired by specific baseline characteristics (e.g., age, sex) but with different exposures (e.g., vaccination vs. no vaccination or vs. different vaccines of the same type of vaccine).

To conduct a cohort study based on administrative data, it is important to correctly define the study population (the denominator) by a set of inclusion and exclusion criteria. For example, in order to investigate only incident cases during the observation period, a prior pre-defined observation baseline period must be free of diagnoses of the disease of interest. The length of the pre-observation period is usually chosen depending on the etiology of the disease to avoid misclassification of prevalent cases as incident ones (chapter 3.2.2).

An open population is dynamic, which is often the case for cohort studies based on administrative claims data, where individuals enter and exit the cohort at different points in time (Figure 3).

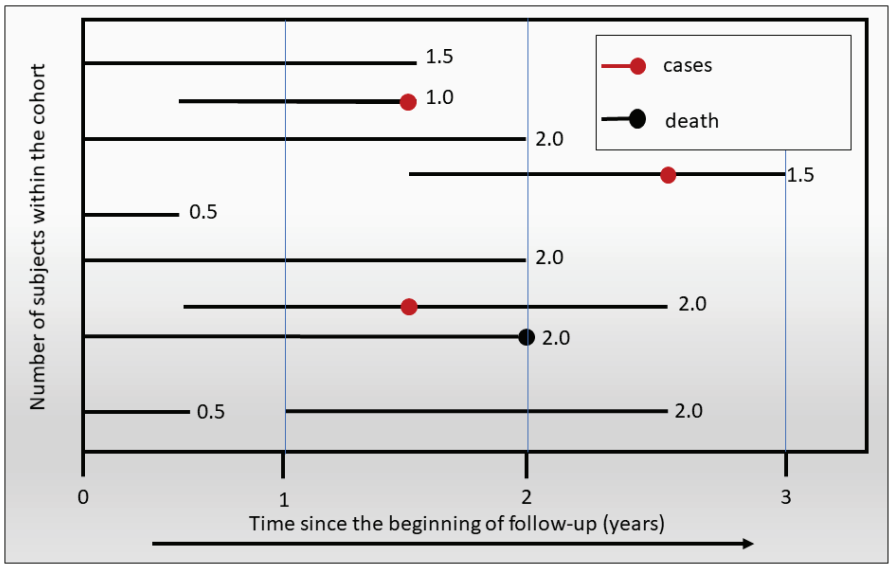


Figure 3: Cohort time contributed by 9 cohort members in an open cohort study (subjects enter and exit at different times, numbers behind bars indicate the person-time in years).

Due to the population dynamic, it is of great importance to exactly define and assess cohort entry and exit dates for all cohort members in order to calculate the individual person-time at risk. Cohort entry of a cohort study based on administrative data can be defined as either a freely selected date in time, a specific age[158] or the exact date of a specific diagnosis, drug prescription or vaccine administration[97]. The exit of a cohort terminates the end of follow-up within the cohort and thus the contribution of person-time at risk and is usually defined as the end of the study period, the occurrence of the outcome of interest or the end of follow-up due to the end of the insurance membership including death (Figure 3). In the open cohort studies included in this thesis[21,81], patients were included if they had at least 12 months of

continuous insurance time during the study period and no diagnosis of the target disease within the 12 months prior to cohort entry. This way, only incident cases were included. In the HZ-study[21], for HZ and its complications, incidence rates were calculated based on person-time at risk that ended with the first respective HZ or its complication diagnosis. However, for descriptive purposes of the ascertainment of HZ treatment and PHN complication following HZ, the follow-up of HZ cases continued beyond the first day of an HZ episode to identify an additional PHN diagnosis or a prescription of a pain medication during the quarter of the HZ diagnosis or the following quarter indicating treatment for PHN and thus to capture all possible PHN cases. Consequently, for PHN, the proportion was calculated and not based on person-time.

Based on the cohort entry and exit definitions, a cohort re-entry after a pre-defined disease-free period is sometimes possible in order to re-enter the cohort as a disease-free cohort member, which is often used for vaccine-preventable diseases that have a short latency and a possibility for relapse. This was also the case in the HZ-study[21], where a cohort re-entry after a repetitive diagnosis-free interval of 12 months after an HZ diagnosis was possible in order to capture all incident HZ cases within the cohort.

To estimate incidence rates in open cohort studies based on administrative data, the numerator is defined by counting the incident number of cases in the cohort within a given interval of time. The denominator for the incidence rate is defined by the person-time at risk which is a more exact definition as it accounts for the length of time each cohort member spent in the population at risk of disease occurrence.

Using administrative claims data presents specific challenges for conducting retrospective cohort studies. Due to the dynamic of the underlying insurance collective in GePaRD, where the insured persons may end or interrupt insurance or switch SHIs, left and right truncation of the data occurs. This is relevant for the denominator definition and consequently can best be taken into account by estimating incidence rates. Furthermore, diagnosis and procedure data are generally available in GePaRD, but only from 2004 onwards, resulting in a left truncation of the data, which hampers obtaining a complete history of previous diseases or medication intake. This is especially important in order to identify prevalent cases within a defined pre-observation period and to exclude them from entering the cohort. The challenge of preventing such outcome misclassification resulting in overestimation of the disease incidence, was addressed in the HZ-study[21] and will be discussed in more detail in chapter 3.2.2.

The cohort study design (as well as RCTs) is the only design that facilitates direct measurement of risks based on person-time. For example, the incidence rate ratio (incidence density ratio) as a measure of the relative risk is directly estimated by the ratio of the

incidence rate in the exposed group to the incidence rate in the unexposed group[159], which are important estimates, e.g., for benefit-risk evaluation in vaccine safety studies[160].

Incidence rates of vaccine-preventable diseases often vary by age, sex, season or calendar year. When incidence comparisons need to be conducted, it is often necessary to calculate stratified or weighted incidence rates instead of only investigating crude incidence rates[161]. These methods imply that the contribution of the cohort members to the numerator and denominator needs to be assigned to the respective stratum. For example, in the HZ-study[21], we estimated yearly age- and sex-adjusted standardized incidence rates (SIRs) using direct standardization referring to a common standard population, which will be valuable when e.g., a comparison with future incidence rates becomes necessary. SIRs can be also used to compare incidences between, e.g., years, geographical regions, immune status or specific time periods, in order to investigate if a specific vaccine is particularly needed in a target population or subgroup[15]. Another example for the comparison of event rates against reliable data from other sources is the GePaRD mortality validation study[162]. Here, age- and sex-standardized mortality rates (direct standardization) in GePaRD were compared to mortality rates from the German Federal Statistical Office. The results showed that death is reliably recorded in GePaRD. Some detected discrepancies may be attributable to differences in the socioeconomic status of the SHI study population in comparison to the total German population[162].

A major advantage of retrospective cohort studies based on administrative data is that they are well suited to investigate the association of a rare exposure (e.g., specific vaccines) on one or multiple outcomes[159]. Due to the temporal framework of this study design, an association between exposure and incidence of an outcome could indicate possible causality[157].

Furthermore, based on administrative claims databases retrospective cohort studies offer the opportunity to include a large cohort sample size with a potentially long follow-up over several years that facilitate incidence estimation usually with sufficient power even for rare events/AVRs[18,163].

However, due to the enormous data size needed, a disadvantage of cohort studies based on administrative data compared to other observational study designs is the very complex and challenging data analysis. For example, to estimate absolute and relative risks, both, the numerator and denominator and also confounding factors need to be measured. To achieve this, the whole cohort needs to be enumerated and kept under surveillance to identify the outcome during follow-up, which becomes especially costly and challenging in the case of time-variant exposure and covariate status[65]. Furthermore, differences between exposure groups (e.g., vaccinated and non-vaccinated cohort members) by, e.g., socioeconomic

status, frailty or healthcare-seeking behavior are difficult to control in cohort studies based on administrative data and can lead to biased estimates (chapter 3.2.1).

3.1.2 Nested Case-Control Studies

In a nested case-control study based on administrative data, the selection of cases (individuals with the disease) and controls (individuals without the disease) is usually sampled by using risk-set sampling from a prior well-defined cohort (source population) with a known sample size that is mostly defined by similar characteristics, e.g., geographical area, birth year, (new) drug/vaccine users[164]. As differences in the odds of the preceding exposure (e.g., vaccination or no vaccination) can be investigated[159], this is an important and valuable epidemiological study design which is sometimes used for vaccine effectiveness[158,165,166] or safety studies[49,167–170].

To design a nested case-control study based on administrative claims data generally requires four steps: (i) the cohort needs to be defined with inclusion and exclusion criteria as well as with specific cohort entry and exit dates as explained for cohort studies. In our NSAID-study[62], cohort entry was defined as the patient's first dispensation of an NSAID (new user cohort). Interruption of insurance status for more than 3 days, end of insurance including death, diagnosis of malignant cancer or the end of the study period/longest available follow-up in the database, whichever came first were defined as cohort exit criteria. We used this design to include only new NSAID users in the cohort as the potential risk of AMI has been suggested to be higher soon after NSAID use than during long-term use.

Furthermore, the primary time axis of the cohort needs to be defined which can either be calendar time or follow-up time.

By using calendar time as time axis, cohort members are ranked chronologically according to their date of cohort entry (e.g., date of first vaccination in a user cohort), which is therefore called a variable-entry cohort.

The time axis defined by follow-up time could be based on disease duration or drug/vaccination exposure (e.g., follow-up time since first vaccination). For example, in the NSAID-study[62], cohort members were ranked with regard to the duration of follow-up time between exposure (first NSAID use) and outcome in the study (diagnosis of AMI) (fixed-entry cohort) (Figure 4a).

The next step is (ii) to select all cases in the cohort at first outcome occurrence. This is different from cohort studies, where cohort members are included based on the presence or absence of the exposure. In an ideal nested case-control study, the cases are the same individuals that would have been identified as cases in the respective cohort study of the same population[171]. To select eligible controls, the third step is (iii) to define the risk-set of possible controls for each case. The risk-set consists of all cohort members who are at risk of

developing the outcome at the date of the outcome occurrence of the respective case (index date), meaning they are disease-free members of the cohort at that index date. Finally (iv), a predefined number of controls are randomly selected from each risk-set (risk-set sampling)[97]. In the NSAID-study[62], up to 100 controls from the cohort of new NSAID users were randomly selected for each case and matched by age at index date, sex and SHI using risk-set sampling which can be also applied in vaccine studies. By the definition of risk-sets, cases and controls have the same index date resulting in the same duration of follow-up and individuals might serve as controls for more than one case and are eligible to be selected as controls until they become a case[172] (Figure 4b).

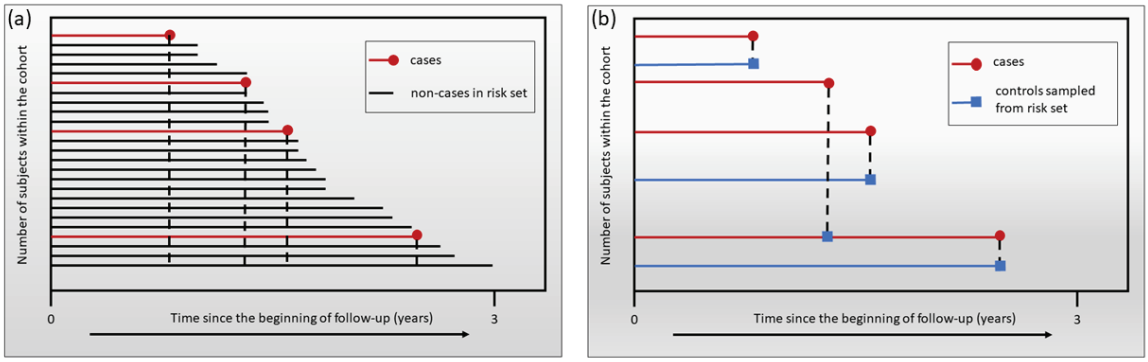


Figure 4a: Illustration of a fixed-entry cohort ranked chronologically by follow-up time, with new risk-sets for the four cases (modified from Suissa 2006, Textbook of Pharmacoepidemiology[173]). Figure 4b: Nested case-control sample of one control per case from the cohort in Figure 2 using risk-set sampling (modified from Suissa 2006, Textbook of Pharmacoepidemiology[173]).

This approach is similar to that in a cohort study, where a case contributes to both, the numerator and denominator for incidence estimation. In the NSAID-study[62], cohort members who were hospitalized for any reason at the index date of the case were excluded from the set of potential controls, since they were not at risk of being hospitalized because of an AMI event.

With the sampling technique in this more modern epidemiological study design, practically the same results for measuring exposure-outcome associations can be achieved as with cohort studies. The matching is accounted for by the statistical analysis, which uses conditional logistic regression to provide OR estimates. The risk-set sampling has the advantage that the resulting OR is an estimator of the relative risk (incidence rate ratio)—that is directly measured in cohort studies—without the necessity of the "rare disease assumption"[171], as the exposure odds of the cases divided by the exposure odds of the controls approximate the incidence rate ratio of the source population[171].

For administrative database studies, the nested case-control design offers several advantages. A major advantage is that the exposure assessment (e.g., vaccination status at the time of the event) is only necessary for cases and matched controls and not for all cohort members[65]. This is more cost- and time-efficient[97] compared to the enumeration of a

complete cohort while the analysis has nearly all the statistical power of traditional cohort studies[157,171]. Another advantage of a nested case-control study is that the calculation of person-time is only needed for the identification of the index date of the cases and controls for risk-set sampling and not for the entire cohort. Furthermore, vaccinated and unvaccinated population groups often differ by further characteristics, e.g., frailty, health status or healthcare-seeking behavior[174]. The possibility to assess more easily information on confounding variables within pre-defined time window(s) before the index date compared to the complete follow-up time in a full cohort study is another advantage of the sampling technique. Therefore, it is also more practical to assess time-variant confounders or time-variant exposure status with different start and end dates than in cohort studies. While in a full cohort analysis, all cases and controls are necessary for analysis[97,175], in a nested case-control study, only a random sample is required. For example, in the NSAID-study[62], potential confounders (comorbidities) were assessed in the 12 months before cohort entry and drug use was determined in the 12 months or 90 days before index date. Classification of exposure status of NSAID use was based on the period (in days) between the index date and the end of supply of the most recent dispensing before the index date based on the DDDs. The user status at the index date was then categorized as current, recent or past use (defined as reference) for cases and controls.

Furthermore, the identification of eligible controls from the cohort in which the case-control study is nested is relatively straightforward and reduces the potential for selection bias[97,156] as the well-defined underlying cohort (e.g., birth cohort, new user cohort), that is mostly defined by similar characteristics is the source of both, cases and controls.

The nested case-control design is especially useful for diseases or outcomes that are relatively rare yet their investigation is of vital importance[65] (e.g., target diseases after highly efficacious vaccines, AVRs after vaccination). It is also more efficient than a cohort study for diseases with long latency or induction periods.

The major disadvantage of nested case-control studies is that they are generally inefficient for the investigation of rare exposures and are limited to one outcome under investigation. In turn, the nested case-control design facilitates the simultaneous evaluation of multiple risk factors for the outcome of interest[176].

However, a study[96] based on GePaRD investigating the occurrence of an AVR, i.e., febrile convulsion within a pre-specified individual risk window shortly after exposure to different brands of childhood vaccination (MMR, MMR+V, MMRV) has been conducted as a matched cohort study which might be sometimes a more appropriate study design because the exposure risk window can be defined better than in a nested case-control setting where a “look-back period” needs to be defined for both cases and controls.

Without information on important confounding factors in the database, it is impossible to conduct vaccine safety or effectiveness studies based on administrative data which is the case for influenza vaccine effectiveness with respect to frailty bias (chapter 3.2.1). Consequently, alternative study designs have to be applied which will be discussed in the next chapter.

3.1.3 Case-Control Studies

In traditional case-control studies, the key challenge is the sampling/matching of the controls from the source population independently of their exposure status. The controls should provide an estimate of the distribution of exposure and covariates in the source population leading to the cases, in order to exploit the benefits of a case-control study over a cohort design[171]. In contrast to a nested case-control study, a traditional case-control study samples cases and controls usually from a source population of unknown size and without similar underlying characteristics.

It is not possible to estimate exposure-specific incidence rates in traditional case-control studies[177], which is in contrast to nested case-control studies (chapter 3.1.2). However, if the disease under investigation is rare—which is usually the case in case-control studies—ORs are a close estimate of the relative risk[159].

Both, case-control studies and nested case-control studies may be less efficient whenever vaccine safety or effectiveness within populations with a very high or low vaccination coverage (e.g., <10% or >80%, respectively) is investigated[49]. For example, in settings with high vaccine coverage and high vaccine efficacy, a decline of the target disease might represent a major challenge for the sampling of an efficient number of suitable cases to investigate vaccine effectiveness with enough precise statistical power and would probably require an increase of the sample size[65]. Furthermore, in this scenario, unvaccinated individuals might differ from the vaccinated ones and consequently from the general population in several characteristics that might be associated with the outcome, independently of vaccination[65]. Then, a person-time analysis from a cohort might provide more valuable information.

Compared to the nested case-control design, a traditional case-control study is more prone to selection bias as cases and controls might differ in several characteristics as they are not sampled from a pre-defined underlying cohort by using risk-set sampling.

Matching in case-control studies is intended to control for confounding and to gain efficiency compared to cohort studies[178,179]. This thesis discusses the case-control design in order to explain the test-negative design as an alternative study design to conduct an influenza vaccine effectiveness study which is not possible in GePaRD.

The test-negative design (TND)

Cohort, nested case-control and case-control studies based on administrative data investigating, i.e., influenza vaccine effectiveness—a vaccine that is recommended for older patients or for patients with comorbid conditions[180]—are particularly vulnerable to selection bias, namely frailty bias which could lead to a healthy vaccine effect[181,182] (chapter 3.2.1). Furthermore, the potential mismatch of the seasonal, circulating viral strain with the antigens in the vaccine increases the complexity of influenza effectiveness studies.

The test-negative design (TND), which shares some similarities with the case-control design, has been suggested as a more valuable design for the investigation of the influenza vaccine effectiveness in order to reduce frailty bias[183]. In a case-control study, cases are sampled based on the disease occurrence and controls are usually sampled based on the absence of the disease. In a TND, the selection for inclusion in the study occurs before the case status is known but only from the group of individuals seeking medical care for symptoms of, e.g., an acute respiratory illness (ARI)[183]. The patients are not defined as cases until they are tested for laboratory-confirmed influenza. Those who are tested negative are defined as controls. In principle, within the TND the vaccine status between influenza test-positive cases and test-negative controls who all sought medical care for an ARI is compared by calculating vaccine effectiveness adjusted for potential confounders[184].

The TND is only applicable if the data source includes information on influenza test and vaccination status or if the required information can be obtained by linkage to other datasets[185]. Unfortunately, laboratory data are usually not available in administrative databases and specifically not in GePaRD. Furthermore, there are several other factors that can affect the vaccine effectiveness, like virus and host factors, prior exposure to the disease, waning of efficacy and boosting, addition of adjuvants, the match to circulating strains, the individual immune status, etc.[58,186,187]. Some of this information is also not available in GePaRD and consequently represents an additional challenge for the conduction of vaccine effectiveness studies based on administrative databases.

For the estimation of influenza vaccine effectiveness, the TND has an advantage compared to other observational study designs mainly because it is less costly and easier to implement as cases and controls are recruited in one process[185,188]. Furthermore, the TND is reported to be less susceptible to frailty bias as one can assume that the test-negative patients were more similar to the test-positive patients in their health status and healthcare-seeking behavior[183] as well as in their comorbidities than randomly sampled controls in traditional sampled case-control studies[189] since controls would have been defined as cases if they had the outcome of interest. Furthermore, outcome misclassification of influenza is reduced as only patients with a positive influenza laboratory test result are defined as cases. This might be also different from traditional case-control studies, where controls are often defined as

individuals without influenza diagnosis but without confirmed test results. Furthermore, it has been reported that estimates of the TND were closer to RCT estimates than those generated with traditional observational study designs[190].

However, by using, e.g., influenza-specific study endpoints, this design presupposes that test sensitivity and specificity of diagnostic methods have to be high to avoid misclassification[80].

Generally, careful interpretation of vaccine effectiveness estimates is needed if a high vaccination coverage results in high herd immunity effects as these would increase protection against the viral agent among vaccinated and unvaccinated individuals by reduced viral transmission, which could lead to an underestimation of vaccine effectiveness[191]. Therefore, to investigate indirect vaccine effects, other study designs are more appropriate which will be discussed in the next chapter.

3.1.4 Ecologic Studies

An ecologic study compares exposure and outcome measures between populations or groups, rather than between individuals[192]. Exposure measures (e.g., vaccine introduction) are often aggregated measures (e.g., proportions) or global measures (e.g., densities). Outcomes can be mortality rates, prevalence data or, like in the HPV-impact-study[81] incidence rates of AGWs that have been stratified by populations groups, e.g., 1-year age groups and sex to observe potential indirect HPV vaccine effects. Also, strata of race or socioeconomic status are possible in ecologic studies.

Within an ecologic study it is possible to estimate aggregated individual-level data like the incidence rates of the target disease at time points before vaccine recommendation and compare them with incidence rates and transmission rates (including the unvaccinated individuals) of the period after vaccination uptake was established (“before/after studies”). This can be done either by graphical displays and/or by segmented regression models (chapter 3.1.5). In the HPV-impact-study[81], in order to capture all vaccination effects, including indirect effects, we studied changes in AGW incidence rates in a study population which was restricted to 11- to 30-year-old females (the target vaccination groups) and males immediately following the recommendation of vaccination to investigate potential herd immunity effects. Therefore, two time intervals were defined: i) time before the recommendation of vaccination and ii) after uptake of vaccination was established.

This study design is an important part of the evidence base for several vaccines after their introduction within the general population[99]. In contrast to other study designs, where different patient groups are compared at a specific point in time, an ecologic time trend study can be considered a valuable research method due to the longitudinal natural experiment as the patients are followed over time to investigate potential changes after a vaccine introduction[193]. For example, the main interest is often to investigate herd immunity effects,

which is a group effect rather than an individual effect and can only occur if the vaccine uptake is sufficiently high as explained in chapter 1.3.3.

There are some advantages of ecologic studies. Based on administrative data, vaccine impact studies can be generated in a less time-consuming way in order to generate a rapid evaluation of the impact of a vaccine at the population level shortly after vaccine introduction. Furthermore, if applied to an administrative data source, this design offers the advantage of including very large population groups facilitating even the investigation of small changes in incidences or within specific subgroups of individuals. However, even in large databases numbers of specific strata might be sometimes too low to achieve statistical significance.

A disadvantage of an ecologic study is that it is not possible to conduct and quantify cause-and-effect analyses among individuals as the data are averaged measurements of groups, and ecologic bias (e.g., ecologic fallacy) can be introduced[194]. This can be interpreted as “the failure of ecologic associations to reflect the biological effect at the individual level”[195]. This definition also implies the impossibility of estimating confounding effects. The underlying problem of this ecologic bias is heterogeneity of exposure and covariates within groups that cannot be captured with ecologic data as information on the individual level is missing[194]. In the HPV-impact-study[81], heterogeneity of the exposure level within specific age groups might explain observed indirect vaccination effects in older female age groups or in males despite quite low coverage overall. This might have occurred as vaccination is voluntary in Germany (although recommended) and the HPV vaccination can occur at any age from 12–17 years (since 2015 from 9–14 years), which is different to immunization programs where a percentage of girls of a certain age is vaccinated. In consequence, coverage in Germany is low for the full age group of 12- to 17-year-olds, but in those older than 17 years coverage cumulates over time as more and more women get vaccinated. Such heterogeneity might also explain why effects in males can be observed despite overall low coverage.

Furthermore, results about potential causal inference within ecological studies based on administrative claims data can be hampered due to very long, varying or even unknown latency between intervention and outcome[194]. In the HPV-impact-study[81], the incidence of AGWs started to decrease among the 16- to 26-year-olds approximately 3 months after the vaccine recommendation by STIKO in Germany in March 2007. The short delay with which the effects of vaccination on AGW incidence were observed is consistent with the biology of infection. As HPV immunity is reported already after one vaccine dose and AGWs develop after a medium incubation time of about 3 months, the corresponding time lag of decreasing incidence of 1–2 quarters of a year after the vaccine recommendation for females is plausible.

3.1.5 Segmented Regression Analysis

For a segmented regression analysis (also called interrupted time-series analysis or piecewise regression), sequential and regular outcome measures are needed at equally spaced intervals and with a sufficient number both before and after the intervention (e.g., vaccination)[196]. Then, a time series of outcome measurements (e.g., incidence rates) can be used to define an underlying presumably stable time trend before the vaccine intervention, which is (potentially) disrupted by the vaccine introduction at a specific time point (breaking point)[197]. However, it is especially important to know the data quality in order to understand the potential impact of coding changes, particularly when these are introduced at the same time as the vaccine[197]. Since we investigated incidence rates of AGWs three years before and after vaccine introduction in the HPV-impact-study[81] and observed approximately stable incidence rates throughout the study years as well as we were not aware of a change of ICD-10-GM coding of the AGW outcome or other conspicuous changes in the data, an interrupted time-series analysis could be conducted to investigate vaccination herd immunity effects of AGWs based on the underlying data source.

The study period can be subdivided by a breaking point into a pre-intervention period and a post-intervention period, creating segments[196] (Figure 5). The impact of the intervention can be examined by the change in level, e.g., a change in the incidence rates occurring between the pre- and post-intervention period[198,199] or by a change in trend defined by a change in the slope of the segment after the intervention compared with the segment before the intervention[196]. However, it is sometimes difficult to define the exact time point when the intervention began. For vaccine impact studies, a breaking point is often assumed at the time of vaccine recommendation/introduction. In a first analysis[200], based on administrative data, a Poisson regression model was used that includes age, calendar time in quarters of the year, calendar time within one year, and calendar time in relation to the breaking point, to test all possible breaking points throughout the study period and to finally select the one with the best goodness of fit based on the Bayesian Information Criterion (BIC)[201]. This analysis[200], suggesting evidence of a change in trend in the AGW incidence in the second quarter of 2007 (HPV vaccine recommendation by STIKO) based on the lowest BIC (breaking point). Based on this, changes in AGW incidence immediately following the recommendation of vaccination were studied[81,200]. For example, in the HPV-impact-study[81], we investigated stable incidence rates in all age-groups between the 1st quarter of 2005 and the 2nd quarter of 2007. Among 16- to 26-year-old females and 16- and 18-year-old males AGW incidence decreased between the 2nd quarter of 2007 and the 4th quarter of 2008 and stabilized at a lower level afterwards. Administrative claims data are a reliable data source for segmented regression analyses, as they provide data over long time series with a high number of regularly assessable time points which increases the power of analysis.

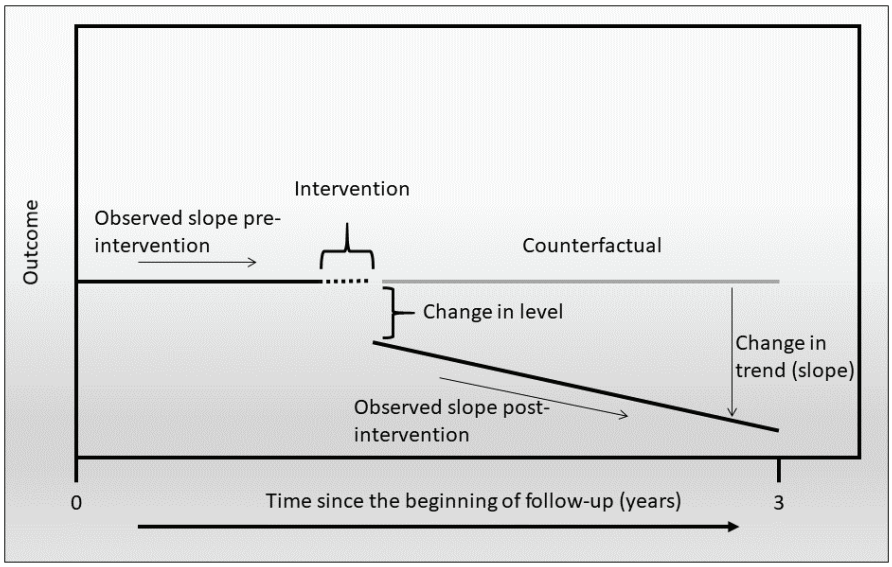


Figure 5: An effect of vaccination observed by a decrease of outcome trends in level and slope. Bold black lines indicate observed rates, the dashed line indicates intervention introduction, the grey line indicates expected rates if the intervention had not been introduced. (modified from Schneeweiss et al. 2001, Health Policy[202])

However, a good understanding of the validity and reliability of the data is necessary with regard to the potential of misclassification (chapter 3.2.2) and confounding (chapter 3.2.3). One major strength of the segmented regression analysis is the fact that it is relatively unaffected by time-invariant confounders (e.g., socioeconomic status). Furthermore, short-term outcomes that occur soon after intervention introduction are the most appropriate outcomes to measure an intervention impact as the timing between intervention and outcome is clearer[197]. For example, in the HPV-impact-study[81], changes of an endpoint occurring earlier (AGW incidence reduction) were used to investigate vaccine impact instead of the primary outcome that develops slowly and requires long-term follow-up (e.g., cervix carcinoma which is the intended preventive goal of the HPV vaccine).

The major advantage of time series designs is that they are valuable study designs to investigate a longitudinal impact of an intervention (i.e. vaccination schedule/program) or for the estimation of intervention effects in non-randomized settings and that they are less likely to generate misleading results compared to other observational study designs[196].

For example, vaccine effectiveness studies conducted as cohort, nested case-control studies or case-control studies do not measure the indirect effects of a vaccine introduction at the population level[65] and are more prone to confounding owing to group differences[197]. Furthermore, other designs are not as suitable as interrupted time-trend studies for the investigation of vaccine impact. This is due to the fact that the dynamics of the vaccine target disease can be easily displayed in time-trend studies and with the design it is possible to control for pre-existing trends[196]. Furthermore, the effect size can not only be estimated at different time points but also as a trend over time.

However, as a disadvantage, it is not possible to control for individual-level covariates that would confound the segmented regression results[196]. Furthermore, in the HPV-impact-study[81], due to data protection reasons, only the birth year but not the exact date of birth was available in the database and some persons may already have been 1 year older according to the definition of age. Consequently, for the analysis of trends over time, the age of each patient was kept the same throughout all four quarters of the year, which led to an average aging cohort from the 1st to the 4th quarter of a year resulting in higher incidences in the 4th quarter than in the 1st quarter of each year most prominently for 16- to 20-year-olds. Furthermore, exact diagnosis dates in the outpatient sector are not available in the database, making the estimation of incidence less precise. It is not possible to control for time-variant confounding or to investigate that the vaccine effect actually occurred in vaccinated individuals (on the individual level) which might produce misleading results[192].

Besides using an ecologic design with time-trend analysis in the HPV-impact-study[81], it would have been also possible to conduct a HPV vaccine effectiveness study as a cohort or nested case-control study based on administrative data under certain conditions. In this case, the assessment of the individual vaccination status of the target vaccination group would have been required. In Germany, since March 2007, the bivalent and quadrivalent HPV vaccinations have been recommended by STIKO for girls between 12 and 17 years of age, and since 2015 for girls between 9 and 14 years. The decision on which of the two available vaccines (bivalent or quadrivalent) should be used is jointly made by physicians and patients. However, the German market is strongly dominated by the quadrivalent HPV vaccine Gardasil® (90% of the market share)[135], so that the effectiveness of Gardasil® in preventing AGWs could theoretically be assessed in vaccinated girls with GePaRD. But during 2007 and partly during 2008, uniform and specific EBM codes for HPV vaccination were not adequately applied. Consequently, it is difficult to exactly define the lack of Gardasil® vaccination and therefore the fraction of non-vaccinated females, hampering the conduction of HPV vaccine effectiveness studies in GePaRD for the time shortly after vaccine recommendation in 2007 and for females older than 17 years (chapter 1.4.2). Therefore, future effectiveness analyses should be limited to federal states with the respective available codes for older females and with longer time periods of more recent years. This will provide information to clarify uncertainties regarding the initial introduction period of the HPV vaccine when specific EBM codes were not available yet. However, for the HPV-impact-study[81], the impossibility to distinguish between the bivalent and the quadrivalent HPV vaccination as well as the fact that non-vaccinated females could not be exactly defined by specific EBM codes during the years 2007 and 2008 was not a limitation for the investigation of the impact of the HPV vaccine in the general population.

3.1.6 Self-Controlled Case-Series Design

The self-controlled case-series (SCCS) design was primarily developed to investigate acute (temporary) adverse events after vaccination (transient exposure) using only data from cases[203,204]. In principle, the SCCS method compares the incidence of the outcome within a pre-defined risk period after exposure to the incidence of the outcome during other times after exposure defined as control periods[97,205].

To conduct an SCCS study for the investigation of events after vaccination based on administrative data, the study time windows are defined first, either by boundaries of specific age-ranges or calendar time periods[97]. The length of the study time windows depends on the vaccine exposure under study and its biological mechanism which might lead to the event. The boundaries of the study time windows should be chosen in such a manner that an individual can experience both, risk and control periods[97]. All individuals who have experienced an event of interest (cases) are selected. The observation period of each case is determined and the exact date of vaccination history is assessed[97]. Risk periods during the observation period after vaccination are defined as, e.g., a pre-defined fixed number of days after vaccination within which an event would likely occur. In the SCCS-study[33], we investigated the stroke risk after an HZ infection and the beginning of the risk period was defined as the date of HZ onset. The end of the risk period was defined as 91 days thereafter since a patient was considered to have an elevated risk of stroke during this risk period. In a secondary analysis, the end of the risk period was further divided into intervals of 1–14 days, 14 days to 1 month, 2–3 months, 4–6 months and 7–12 months after HZ onset to assess the temporal pattern of the HZ-related risk. Remaining time periods of the cases' observation period (more than 12 months) were defined as control periods. Consequently, in the SCCS design, the entire observation period of an individual case is divided into fixed risk and control periods, meaning individuals act as their own controls[97]. The events that are regarded as random (e.g., stroke in the SCCS-study[33]) are "mapped" in relation to one of these different periods. Therefore, the only variability is whether the event occurs in a risk or control period. This information is then used to estimate the exposure effect. By using conditional Poisson regression the relative incidence of the outcome is estimated[97,206]. There are three main assumptions of the SCCS design: (i) events arise in a non-homogeneous Poisson process, (ii) the SCCS design requires both, independent recurrent events or rare non-recurrent events and (iii) the observation period must be independent of the event date, meaning the outcome must not censor or terminate the observation period[206]. If events always occurred at exactly the same time or age, the SCCS design would fail as there would be no within-individual variation[207].

Applied to administrative databases, the SCCS design is a valuable and appropriate method if exact information on the denominator is lacking or difficult to be defined and it is most

attractive to account for all time-invariant confounders (e.g., genetic and socioeconomic factors) as it is self-controlled[207]. Furthermore, for investigating vaccine safety issues or risks of severe complications after an infection there may be substantial differences between exposure groups. As comparisons in the SCCS are made within individuals rather than between individuals, this is a valuable advantage.

Based on administrative data it is challenging to investigate whether the main assumptions of the SCCS design are violated. In the SCCS-study[33], the latter two assumptions were violated when a stroke resulted in death. Therefore, sensitivity analyses were conducted and patients with fatal strokes were excluded to evaluate the impact of fatal strokes on the study results[206]. However, the exclusion did not change the estimate of the IRRs nor the corresponding CIs, supporting the robustness of our results.

Compared to cohort, case-control or nested case-control studies, the SCCS design reduces unmeasured and residual confounding, for which adjustment is difficult in other observational study designs. This was especially valuable in the SCCS-study[33], since multifactorial risk factors that are not represented in GePaRD data can lead to stroke. However, it is often necessary to control for other time-variant confounders especially if follow-up time is long (chapter 3.2.3). This was done in the SCCS-study[33] as the observation period was split into further time-periods to account for respective changes of confounding factors over time[205].

The SCCS design comprises some principles of a cohort study, for example, based on the fixed vaccination history, the individuals are followed over time until events occur (or not). However, unlike in a cohort study, follow-up time is not censored at the time of event occurrence[205].

With the SCCS design, it is also possible to use more than one risk period and to investigate the association between an event and multiple exposures or vice versa as well. Acute events and shorter risk periods relative to control periods are favored over longer and therefore more indefinite risk periods as these are more susceptible to confounding between, e.g., age (or other relevant time-variant factors) and exposure (vaccination) effects[205].

Furthermore, SCCS designs are considered suitable or even equivalent to cohort studies for investigating the influence of temporal, explicitly defined vaccinations on AVRs or of infections on acute events, but they are also considered susceptible to exposure misclassification (chapter 3.2.2). Thus, in administrative databases and for individuals, who have multiple outpatient diagnostic codes within the same quarter and/or diagnostic codes during hospitalization, it is difficult to determine if the outcome subordinated the exposure. It is therefore necessary to assess a date-specific exposure and outcome status, which was done in the SCCS-study[33] (chapter 3.2.2).

3.2 General Biases

A bias can be defined as a systematic error of an observational study design that occurred during the design and conduct of the study[208]. As it is difficult to obtain adjustment for bias at the stage of analysis, it is important to consider and control which bias might be introduced in the study in order to avoid or minimize its extend on the validity of the observational study results.

There are several types of bias in clinical epidemiological studies, but most of them can be assigned to two major groups: selection bias and information bias.

Within this chapter, (I) a specific type of selection bias, namely frailty bias, will be discussed. This bias has been reported to occur in influenza vaccine effectiveness studies conducted with cohort, nested case-control or case-control designs based on administrative data; (II) outcome and exposure misclassification will be discussed and examples of addressing this form of information bias within the underlying data source will be given; (III) time-variant and time-invariant confounding and ways to address them by using the appropriate study design to investigate the respective research questions will be discussed.

3.2.1 Selection Bias

Selection bias in descriptive studies (incl. surveys) occurs if the selection, the self-selection (bias due to sampling), no participation (non-response bias) or the survival (survivor bias) of individuals for study participation is not representative of the population that was originally intended to be analyzed[208]. Selection bias in analytic epidemiological studies (observational studies) based on administrative data can occur (i) due to non-representative “unexposed” groups of the source population, (ii) due to differential loss to follow-up among exposed and unexposed individuals, (iii) if the selected cases are not derived from a well-defined source population or (iv) if the selected controls are not providing an unbiased sample of the exposure distribution in the source population[208,209].

Selection bias would lead to a different exposure-outcome association in the study population compared to individuals who are not participating in the study but are potentially eligible to be included. The observed effects may then be biased by factors determining participation and/or outcome[161].

Frailty Bias

In influenza vaccine effectiveness studies using a cohort or case-control design, selection bias may occur as vaccinated and non-vaccinated persons might differ with respect to frailty[63]. Frailty is defined by Fried et al. as “a biologic syndrome of decreased reserve and resistance to stressors, resulting from cumulative declines across multiple physiologic

systems, and causing vulnerability to adverse outcomes"[210]. This definition distinguishes frailty from disability or comorbidity and assessing accurate indicators of frailty is therefore quite difficult[181].

Because influenza vaccination is recommended mostly for the elderly or in populations with poor health and as vaccination is voluntary, a frailty bias occurs if patients choosing to be vaccinated have a better baseline health or have better health behaviors than patients not being vaccinated[181,211,212].

The influenza vaccine should reduce the occurrence of flu symptoms, hospitalizations for pneumonia and mortality[193]. In the presence of frailty bias, an overestimation of influenza vaccine effectiveness has been reported to occur in cohort or case-control studies, resulting in large biased reductions in all-cause mortality of ~50%[165,213], while excess mortality attributable to influenza has been reported to be less than 10%[214,215]. Furthermore, several cohort studies not accounting for frailty bias reported less pneumonia-related hospitalizations for influenza-vaccinated older patients compared to non-influenza vaccinated older patients[216–218]. Accurate indicators of frailty and consequently differences regarding frailty between the vaccinated and non-vaccinated groups are generally hard to characterize in administrative databases as frailty is not captured by ICD-10-GM codes[181]. Consequently, matching or adjusting for frailty is either not possible or rather difficult.

Therefore, the cohort, case-control or nested case-control study design using administrative databases for influenza vaccine effectiveness is prone to bias, especially when investigating unspecific outcomes (e.g., all-cause mortality, pneumonia)[219]. Instead, studies of alternate designs like the TND have been reported to be less susceptible to frailty bias in influenza vaccine effectiveness estimation[190] as they reduce differences between vaccinated and unvaccinated persons during the study design phase resulting in less selection bias[184] (chapter 3.1.3).

An area of application for observational cohort studies based on administrative database without the possibility to control adequately for frailty bias is the comparison of influenza vaccine effectiveness estimates obtained "off-influenza season" to estimates from the influenza season to detect frailty bias since usually no vaccine effect should be present during the "off-season"[220]. If vaccine effectiveness were detectable during this control period, this would indicate bias due to frailty.

3.2.2 Information Bias

Information bias occurs if systematic differences regarding exposure or outcome classification between study participants arise. Such classification errors can be introduced either by the observer in RCTs (observer bias), the study participant (responder bias) or by the measurement instruments (instrument bias). Errors in measurement itself can be defined

as misclassifications and can be distinguished into differential and non-differential misclassification[221]. A differential misclassification occurs if the information errors differ between exposed and non-exposed individuals or between cases and controls which can lead to an over- or underestimation of the study effects. A non-differential misclassification occurs if misclassification is similar between the exposure groups (in cohort studies) or outcome groups (in case-control studies). This happens, if the exposure is unrelated to other variables (incl. the outcome) or vice versa and can usually lead to bias towards the null[222]. For example, the intake of OTC pharmaceuticals (e.g., Aspirin) can only be defined as "non-use" in GePaRD as OTC pharmaceuticals are not recorded in the database. In reality, the intake would need to be defined as "use" for patients obtaining it OTC.

Therefore, the correct identification of the study outcome and the exposure status is of crucial importance. As administrative claims data are primarily collected for reimbursement purposes and comprise a dynamic structure of insured individuals, exposure and outcome assessments based on this data source involve specific challenges[138] that will be discussed in the following.

Outcome misclassification

The outcome identification impacts the estimation of the disease burden or the identification of a disease as an outcome in effectiveness or safety studies. The aim of an accurate assessment of the study outcome is to avoid misclassification, e.g., to erroneously classify a healthy individual as "diseased" or a diseased individual as "healthy".

Outcome identification in administrative claims data is mostly based on the information from outpatient and inpatient diagnoses. This means that only diseases for which the patients were seeking medical care (also including chance findings [Ger.: Zufallsdiagnosen]) and which could be coded by a respective diagnostic ICD-10-GM code can be identified in the GePaRD database.

Since 2003, inpatient diagnoses are coded by uniform coding guidelines and are generally used for reimbursement purposes based on diagnosis-related groups (DRGs) which combine the principal discharge diagnosis, diagnostic codes, and clinical procedures[223]. As SHIs cover the operating costs via DRGs, inpatient diagnoses are checked and audited regularly by the SHIs. Therefore, it can be assumed that inpatient diagnoses accurately reflect patients' clinical conditions as well as the provided care[223].

Outpatient diagnoses depict the physician's medical opinion at the time when care is given to the patient without a required laboratory confirmation of the diagnosis classification[223,224]. However, sometimes medical information and diagnoses may change during the course of the disease and a final reliable diagnosis can only be made if several pieces of information are assembled over time. To indicate the medical decision-making process, the coding of

outpatient diagnoses requires physicians to code diagnostic certainty with the categories ‘certain’, ‘suspected’, ‘excluded’ and ‘status post’. In contrast to regular reviews of inpatient diagnoses, outpatient diagnoses are not checked regularly for coding quality as they are not related to the amount of reimbursement[223]. Therefore, there may be uncertainties regarding accurate diagnostic coding, and thus the validity of the information[225]. However, EBM and OPS codes are checked and audited as they are directly related to reimbursement, and hence give valid information on coding of in- and outpatient operations and procedures[223].

Due to German data protection regulation, it is almost impossible to validate cases against patient charts. However, to validly identify an outcome it has been suggested (i) to develop case-identifying algorithms with different sensitivity and specificity, (ii) to provide information on outcome misclassification, (iii) to compare event rates against external data or (iv) to build longitudinal patient profiles to better understand the clinical context[223].

One example for applying different outcome definitions is the HZ-study[21], in which HZ episodes were identified via inpatient main discharge diagnoses as well as via outpatient diagnoses coded as ‘certain’ or with missing diagnostic certainty. In 2004/2005, coding of diagnostic certainty was not yet fully implemented and it has been reported that only about 22% of outpatient diagnoses have been coded with their diagnostic certainty[225]. If the diagnostic certainty was only partly captured (from introduction until 2005) and no missing diagnostic certainty was included in the outcome definition, the burden of disease would be underestimated. On the other hand, if the diagnostic certainty was captured wrong, e.g., if a ‘suspected’ or ‘excluded’ diagnosis was coded as a certain diagnosis, misclassification could be introduced and an overestimation of the incidence rates were to occur. To determine potential misclassification due to missing diagnostic certainty in the year 2005, a second outcome definition with higher specificity and lower sensitivity was introduced, that included not only ‘certain’ but also ‘suspected’ outpatient diagnoses for the years 2006-2009, and confirmed the possible overestimation in 2005 due to missing diagnostic certainties. Another example for the application of different outcome definition algorithms is the identification of PHN cases in the HZ-study[21], where in addition to ICD-10-GM diagnoses, also specific treatment options recommended for zoster pain by the Guideline of the German Dermatological Society were included[226].

Because patients with chronic diseases may have regular contact with the healthcare system, their diagnoses are regularly coded in the data. In this case, a shorter time period may be sufficient to identify them as prevalent cases. For infectious diseases like HZ, a longer pre-observation period of one or two years is mostly chosen in order to identify all prior HZ diseases and to avoid misclassifying prevalent cases as incident ones. However, an incomplete history of the disease before cohort entry is a frequent problem within

administrative database studies due to left truncation of the data[97] especially if the duration of the disease correlates with disease severity (e.g., diabetes mellitus).

Outcome algorithms with different sensitivity and specificity have been also applied in some other GePaRD-based studies[96,113]. For example, in a study assessing the risk of febrile convulsion after vaccination with Priorix-Tetra™, MMR and MMR + V, one outcome algorithm included hospitalized cases of febrile convulsion and excluded alternative plausible causes which might have led to febrile convulsion (e.g., an infection or a neurologic condition) to gain high specificity, while the other outcome algorithm with higher sensitivity only excluded hospitalizations for febrile convulsion with a neurological condition as main discharge diagnosis[96].

In another study that investigated background incidence rates of the Guillain-Barré Syndrome (GBS), one outcome algorithm included primary diagnostic inpatient codes of GBS while the second—more specific—algorithm additionally included OPS codes as recommended in the German guidelines for GBS diagnosis and treatment[113].

Preference of the respective algorithm is context dependent. However, regarding the public health relevance, a more specific outcome definition is likely closer to the “true” incidence in the population and is therefore a more appropriate measure.

Exposure misclassification/misclassification of vaccination status

The identification of the vaccination status influences the estimation of vaccine impact, vaccine effectiveness and vaccine safety issues or the risks of subsequent infectious disease-related complications. An inadequate definition of the vaccination status can lead to misclassification.

Administrative claims data have been reported to be the gold standard for assessing drug exposure as there is no occurrence of recall bias, i.e., if cases remembered their exposures better than controls[227]. Furthermore, due to the almost complete lack of selection effects among participants because of a routine care situation reflected in the data, complete information on prescriptions (if they were filled in the pharmacy to get reimbursed by SHIs) or procedure codes is available[227]. Also information on prescriptions of subpopulations that are difficult to recruit for field studies, e.g., children, older persons or pregnant women, is available in administrative claims data[227]. Vaccinations can also be identified in administrative claims data by outpatient codes used for reimbursement of the administration of vaccines (ASHIP-specific EBM codes). However, vaccines are usually not prescribed to individual patients but included in the so called “Sprechstundenbedarf” (medical products for use at a physician’s office only) which cannot be identified in the database.

Misclassification of vaccination status can also occur in administrative claims data as no information on the antibody level is available (this is also true for all other data sources not

containing laboratory data on antibody level). If vaccinated patients do not show a serological antibody response, these patients could be misclassified as vaccinated with an adequate immune response but have in reality no immune response and are therefore still prone to infection.

Furthermore, for vaccine safety studies, the definition of the temporal relation between vaccination and systemic immunization reactions is challenging as this varies between vaccines (e.g., a few hours after the pertussis vaccine; up to several days for the measles vaccine), brands (due to different antigens and adjuvants) or individuals (individual immune response/reaction).

Vaccination status misclassification can also occur for vaccinations that have not been recommended by STIKO or for vaccinations upon “self-pay” prescription as reimbursement of them is SHI-dependent (e.g., for travel vaccinations) and information on these vaccinations is thus not completely available in the database which hampers the feasibility of vaccine studies for these vaccinations.

Within the HPV-uptake-study[71], only vaccination uptake and not a true coverage could be calculated due to the restricted study period of the year 2008 only, for which data were available. For coverage estimations, knowledge about the vaccination status before this study period would have been necessary, which was not possible due to the lack of specific EBM codes. Additionally, the study data refer to a time period shortly after introduction of the vaccine in 2007 and thus cannot directly depict the situation in more recent years. However, since 2008, HPV coverage remained almost stable over time and the estimated HPV vaccine uptake in our study has recently been confirmed in a coverage study based on ASHIP data[228]. Furthermore, a comparison of the HPV uptake data with the official sales figures can facilitate quantification of misclassification effects. Despite the limitation to estimate HPV coverage data within our study, the estimation of the HPV vaccination coverage per age cohort for more recent years would be possible with the GePaRD database by using the introduced uniform and specific EBM codes for HPV vaccination across Germany.

To determine the time interval between an infection of interest (exposure) and the risk of a following complication (outcome), it is necessary to know the exact date of both, exposure and outcome. While inpatient diagnoses are related to an exact date, outpatient data are not. Instead, they can only be allocated to a quarter of a calendar year as outpatient physician visits are reimbursed on a quarterly basis. If the exact date of a diagnosis in outpatient data is missing, one may use—according to the respective research question—the prescription date of a disease-specific drug treatment, EBM or OPS code or an inpatient diagnosis to reliably determine an exact date.

In the SCCS-study[33], the main inpatient discharge diagnosis was used to identify an HZ infection and the exact date of disease onset was defined by the admission date. For

outpatient data, diagnoses of HZ qualified as 'certain' or 'suspected' by the respective physician were considered and the date of disease onset was defined as the date of systemic antiviral prescription (acyclovir, brivudin, famciclovir, valaciclovir). Patients with a first outpatient HZ diagnosis and no accompanying prescription had to be excluded as no exact date of the exposure could be assessed. Using an approximate date would have hampered the definition of exact risk and control periods in the study design (chapter 3.1.6).

3.2.3 Confounding

Confounding occurs by "the distortion of the estimated effect of an exposure on an outcome due to the association of the exposure with other factors that influence the occurrence of the outcome"[208]. A factor can be defined as a confounder by fulfilling three assumptions: (i) the confounder must be associated with the exposure and outcome under study in the source population, (ii) the confounder must be a causal risk factor for the disease in the unexposed cohort and (iii) a confounder must not lie in the causal pathway (intermediate cause) between exposure and the disease[208]. For example, the socioeconomic status is a potential confounder in, e.g., estimating the effectiveness of 13-valent pneumococcal conjugate vaccine (PCV13) on pneumococcal disease in children since the socioeconomic status is associated with poor health outcomes and vaccination status[229].

In observational studies, it is important to adequately consider confounding factors if the information on an individual level is available[161]. As administrative data are not primarily collected for research purposes, important potential confounders regarding the respective research questions may be either missing or incomplete, e.g., body mass index (BMI), history of smoking, alcohol consumption, lifestyle factors, physical activity, socioeconomic status (chapter 1.4.2). If these factors are confounders, unmeasured and/or residual confounding could distort the association between exposure and outcome. Consequently, this would lead to misinterpretation of the study results. It is therefore highly important to check in advance if the potential research questions can be accurately investigated based on the available information in the database to achieve valid results.

Time-invariant confounding

For time-invariant measured confounding, it is possible to restrict the study population to individuals who are similar in relation to the confounder or to match the selected controls to cases by confounders. This way, a similar distribution of confounders among cases and controls could be achieved. Beyond matching by length of follow-up in nested case-controls studies in GePaRD, cases and controls are often matched by age at the index date and by sex but also by SHI. In the case of vaccine effectiveness or safety studies, matching by SHI would be important as differences regarding the "vaccination behavior" ("Impfverhalten")

might exist between different SHIs. Vaccine refusal is associated with factors related to a higher socioeconomic status[230] and the socioeconomic status differs between the patient populations insured by the respective SHIs in GePaRD.

During the analyses, it is possible to control for confounding by standardization, stratification, restriction or adjustment in multivariable analyses. Furthermore, the current "gold standard" in observational research is to control confounding by using propensity score (PS) methods which can be defined as the probability of a patient receiving the treatment under investigation. The PS is used for different procedures: (i) matching on the PS, (ii) building the inverse probability of treatment weighting (IPTW), (iii) stratification on PS strata or (iv) regression adjustment for the PS[231]. For unmeasured confounding, other techniques like high-dimensional propensity score (HDPS) analyses or instrumental variables may be applied to reduce potential confounding during analysis. However, this will not be further discussed as this topic is outside the realm of this thesis.

For unmeasured time-invariant confounders, case-only designs (e.g., case-crossover design or self-controlled case-series (SCCS) design) are valuable methods to investigate vaccine-related research questions and thus a case-only design was applied in the SCCS-study[33] (chapter 3.1.6). The SCCS design implicitly controls for time-invariant confounders as every case represents its own control. This is very helpful if not all confounding factors are available in the administrative database.

Time-variant confounding

The more complicated form of confounding refers to variables that may vary over time. However, effects of potentially time-variant confounders (e.g., age or seasonality) can also be investigated using the SCCS design by dividing the observation periods according to, e.g., age groups or any other relevant time-variant factors.

Within the SCCS-study[33], the follow-up time was stratified by potential cardiovascular risk factors that varied over time (e.g., myocardial infarction, transient cerebral ischemic attack, atrial fibrillation and flutter identified from inpatient and outpatient diagnoses; use of antithrombotics, use of statins identified from outpatient dispensations or age). Furthermore, it is possible to analyze stratified estimates according to these risk factors. To investigate whether the confounding factors were time-dependent, changes in the prevalence during baseline and follow-up were investigated.

Other examples of time-variant confounders that may bias estimates in vaccine impact studies are seasonality, levels of infectious diseases that are prone to outbreaks, weather, etc.[197] (chapter 3.2.3).

While a segmented regression analysis is relatively unaffected by time-invariant confounders, time-variant confounders can mask or exaggerate the vaccine impact[197]. Examples of time-variant confounders are other events that occur around the same time as

the intervention. This could either be (i) other simultaneous interventions targeting the same outcome, (ii) seasonal changes in the outcome during the intervention or (iii) changes in diagnostic coding patterns or reporting patterns[197]. It is challenging to disentangle changes in incidence rates due to vaccine introduction from changes caused by unrelated factors such as secular trends or changes in healthcare utilization, reporting patterns, access to care or diagnostic coding practices and the underlying health of the population[232].

If time-variant confounders are known, they can be included in the regression model. However, to separate intervention effects from unmeasured or unknown time-variant confounding effects the use of control groups may be necessary. This can be done either by choosing a group of subjects that is similar to the study group but does not experience the intervention and is followed over the same time period as the group with the intervention[196] or by using baseline data from the outcome during a period when the vaccine had not yet been introduced. In this instance, the comparison of the effect between intervention and control groups makes it possible to disentangle the real intervention effect from confounding effects that could have occurred during the same time as the intervention. In the HPV-impact-study[81], no decrease in age groups other than in the vaccine-recommended female age groups was seen which could support the assumption of causality for the decreasing incidence in the younger age groups following vaccine introduction. This would also invalidate the assumption that other events could have led to a change in AGW incidence as AGWs are a specific outcome for HPV-types 6 and 11 infection, that are responsible for 90% of AGWs[233]. This also indicates that specificity of the outcome is highly important in investigating vaccine impact in time-trend studies[99].

It is also possible to investigate the same group of subjects while comparing changes of the target disease against incidence rates of other diseases during the same time period[78,234]. For example, the effect of PCV13 use on pneumococcus-related hospital admissions has been assessed while urinary tract infections and hospital admissions for any reason were used as control outcomes[234]. However, selecting a comparison disease is challenging as both the target and comparison disease should share similar biases and causal factors while the comparison disease should not be influenced by the intervention[99,235]. Retrospectively, the simultaneous estimation of incidences of STIs other than AGWs would have been interesting in the HPV-impact-study[81] and could have supported our results.

A further approach to control for possible time-variant, concurrent events is the multiple baseline design. Here, the intervention (e.g., vaccination) is introduced in different geographical areas at different times. For example, multiple population subgroups receive the intervention time-delayed (e.g., several weeks or months) to observe potential outcome changes in relation to intervention introduction at a different time point[236]. Another approach is to add an additional phase, e.g., if the intervention is introduced, withdrawn and introduced

again afterwards, in order to investigate a reversal of the effect by the withdrawal[197,236]. Additionally, more complex statistical analyses like splines can be used to control for time-variant confounders but this will not be further discussed[197].

4 CONCLUSION AND FUTURE PERSPECTIVES

This thesis has shown that administrative healthcare data are an important data source for epidemiological research of vaccines and vaccine-preventable diseases. Studies conducted with this data source require detailed knowledge of the structure of the database and the depth of the available information. Furthermore, as it is highly important to avoid bias and confounding, careful study planning under the consideration of methodological challenges is necessary, some of which have been highlighted and discussed in this thesis. This thesis further has demonstrated that not all studies of vaccines or vaccine-preventable diseases can be conducted using administrative databases. It is therefore important to precisely evaluate the feasibility of a study in advance, e.g., the estimation of the influenza vaccine effectiveness is not possible with GePaRD. To continue and improve future observational epidemiological studies of vaccines and vaccine-preventable diseases based on administrative data sources for the investigation and evaluation of vaccine-related effects in a real-world setting, it is necessary to be aware of methodological and data source-dependent challenges.

In the 21st century, vaccine-preventable diseases are relevant national and international public health concerns and different challenges regarding vaccines and vaccine-preventable diseases may arise in the future. New challenges in the battle against infectious diseases are emerging in the course of globalization and related social, demographic and environmental changes. The increased mobility and density of the population also increases the rapid spread of infectious diseases which has been demonstrated during the hazardous Ebola outbreak 2014–2015 or during the severe acute respiratory syndrome (SARS) outbreak 2002–2003. The potential of the viruses to infect a large number of persons across different countries is of great concern and requires coordinated surveillance strategies between different countries[237]. Additionally, ongoing epidemics of tuberculosis, human immunodeficiency virus (HIV), malaria and influenza as well as new zoonotic pathogens constitute major clinical management challenges worldwide[238]. Furthermore, the increased occurrence of drug- or vaccine-resistant strains (e.g., methicillin-resistant *Staphylococcus aureus* (MRSA), drug-resistant tuberculosis bacteria or drug-resistant malaria parasites)[40] will create difficulties for the treatment of infections with resistant organisms. These examples indicate the importance of developing new vaccines and vaccination strategies (e.g., the new vaccine rVSV-ZEBOV against Ebola), the need to expand the use of existing vaccines and the tremendous significance of continuously evaluating faster and repeated efficacy and effectiveness studies in the future to combat, e.g., pathogen-resistant strains. Additionally, new endpoint definitions (e.g., diseases) and different confounding factors might be required in future vaccine studies.

Fortunately, the progress made in laboratory and medical techniques into pathogen-host interactions, pathogenesis and inflammatory pathways as well as the understanding of the host's immune response are leading to the identification and the development of new therapies[238]. In addition to the success of traditional vaccines against infectious diseases, there are now also some vaccines available that target infectious agents causing cancer, like the HPV vaccine and hepatitis B vaccine. However, the investigation of vaccination effects on the primary outcomes of cervical cancer or liver cancer, respectively, among adults will take decades. To date, mathematical predictive models can be very helpful to examine the long-term effects of vaccination strategies over several decades.

However, there is also a newer concept of vaccinations. Instead of targeting infectious agents, it comprises therapeutic vaccination against chronic diseases. For example, in the United States, the cancer vaccine Sipulencel-T (Provenge®) has been approved by the FDA for the treatment of metastatic prostate cancer. Against this background, future vaccine studies involving therapeutic vaccines against chronic diseases or cancer will face different challenges than studies on traditional vaccines as the spectrum shifts. Meaning, while in the case of traditional vaccines, the target disease can be linked to the respective pathogen (e.g., 99% of cervical carcinoma cases can be attributed to HPV), there is no such association for vaccines against chronic diseases. Only a fraction of chronic diseases can be referred to the pathogen as most of them are multifactorial, e.g., modifying lifestyle behavior or comorbidities that might interact over very long-term periods/decades and develop over time can lead to the disease. Other future challenges of vaccine studies could be safety concerns, e.g., cancer following SV40-contaminated poliovaccine[239] or false suggestions regarding long-term AVRs that might result in immediate reductions of vaccine uptake. Consequently, the investigation of effectiveness and safety of vaccines against chronic diseases is especially challenging as it will take several decades before the respective outcomes will occur and it might be even more complex than is the case for current traditional vaccines due to the multi-causal nature of chronic diseases.

These examples show that an administrative database covering data of several decades with information on interacting risk factors over long time periods is urgently needed for the execution of vaccine studies either on traditional vaccines against infectious diseases but also on vaccines targeting chronic diseases.

Fortunately, the immense importance of such databases has also been acknowledged by German legislation. The legal foundation of such databases has been reformed due to the introduction of a new European data privacy law. In order to preserve databases such as the GePaRD database, an amendment of § 75 SGB X (use of social security data for research purposes) will become effective in May 2018. As a result, it will be possible to investigate

future long-term vaccine effectiveness, impact and safety issues in a real-world setting in Germany.

The increasingly large amounts of available data require the use of new methods. For example, additional information by linkage of electronic healthcare databases with other data sources (e.g., surveillance/outbreak systems, disease registries, Disease Management Programs (DMPs), biobanks, global laboratory data, etc.) will be necessary for future vaccine studies, especially for vaccines against multifactorial chronic diseases. The resulting linked data would offer a valid opportunity to rapidly monitor the acute epidemiological situation. Furthermore, they could provide additional information (e.g., laboratory parameters on individual serological tests and antibody titers, virus strains, biomarkers, biomaterial, etc.) for studies on vaccines against chronic diseases requiring specific and different outcome definitions (e.g., disease occurrence, remission, competing risks events, overall and cause-specific survival) or for the development of multivariate algorithms to combine information on multiple interacting risk factors for chronic diseases over long time periods, even decades[240] (e.g., lifestyle behavior, comorbidities, medication intake, biomarkers).

The electronic health record [Ger.: elektronische Gesundheitsakte] could be a valuable tool to provide such additional important data. In the form of a longitudinal electronic record system, it is envisioned to integrate the medical history, treatment data, medication, allergies, vaccination certificates, laboratory data and other health data of the SHI sector as well as cross-sectoral and non-medical information (OTC medication, diets, physical activity, etc.)[241]. In theory, this considerably extensive data depth should be completely available in real time which could then be used for future vaccine effectiveness and safety studies, particularly regarding chronic diseases. However, creating the necessary infrastructure, the resources and the technical support[242] combined with specific data protection regulations will be a tremendous challenge in the future.

A two-phase design used in electronic healthcare databases could also be a valuable and helpful tool to solve some of the above-mentioned challenges of the investigation of future vaccine effects. Here, more detailed information from linkage with other databases (e.g., DMPs, disease registries, biobanks) could be included, but only for a subgroup of patients (phase 2). This additional information could then be used by including data from the entire study population (phase 1) to obtain valid results in future studies of vaccines and vaccine-preventable diseases.

5 REFERENCES

1. Nelson, E. Kenrad WCM, ed. Early history of infectious disease: epidemiology and control of infectious diseases. In: *Nelson, E. Kenrad, Williams, C. M., Infectious Disease Epidemiology: Theory and Practice*. 3rd ed. Burlington: Jones & Bartlett Learning; 2014:3-18.
2. Omran AR. The epidemiologic transition. A theory od the epidemiology of population change. *Bull World Health Organ*. 1971;49(4):509-538. doi:S0042-96862001000200011 [pii].
3. McKeown RE. The Epidemiologic Transition: Changing Patterns of Mortality and Population Dynamics. *Am J Lifestyle Med*. 2009;3(1 Suppl):19S-26S. doi:10.1177/1559827609335350.
4. Frieden TR. The Future of Public Health. *N Engl J Med*. 2015;373(18):1748-1754. doi:10.1056/NEJMsa1511248.
5. Riedel S. Edward Jenner and the history of smallpox and vaccination. *BUMC Proceeding*. 2005;18:21-25. doi:10.1159/000322101.
6. World Health Organization (WHO). The expanded programme on immunization. Immunization, vaccines and biologicals. www.who.int/immunization/programmes_systems/supply_chain/benefits_of_immunization/en/. Published 2013. Accessed July 5, 2017.
7. World Health Organization (WHO). Smallpox vaccines. Emergencies, preparedness, response. http://www.who.int/csr/disease/smallpox/vaccines/en/. Published 2016. Accessed July 5, 2017.
8. World Health Organization (WHO). Principles and considerations for adding a vaccine to a national immunization programme- From decision to implementation and monitoring. Department of immunization, vaccines and biologicals. www.who.int/immunization/documents%0A. Published 2014. Accessed July 5, 2017.
9. World Health Organization (WHO). Vaccines and disease. Immunization, vaccines and biologicals. www.who.int/immunization/diseases/en/. Published 2017. Accessed July 5, 2017.
10. World Health Organization (WHO). Vaccination greatly reduces disease, disability, death and inequity worldwide. Bulletin of the World Health Organization. www.who.int/bulletin/volumes/86/2/07-040089/en/. Published 2008. Accessed July 5, 2017.
11. Atkinson P, Cullinan C, Jones J, Fraser G MH. Large outbreak of measles in London: reversal of health inequalities. *Arch Dis Child*. 2005;90:424–425. doi:: 10.1136/adc.2003.048892.

12. Phadke VK, Bednarczyk RA, Salmon DA, Omer SB. Association Between Vaccine Refusal and Vaccine-Preventable Diseases in the United States: A Review of Measles and Pertussis. *JAMA*. 2016;315(11):1149-1158. doi:10.1001/jama.2016.1353.
13. Leroux-Roels G, Bonanni P. Vaccine development. In: Garçon N, Leroux-Roels G CW-F, ed. *Understanding Modern Vaccines: Perspectives in Vaccinology*. Vol 1. ; 2011:115-150. doi:10.1016/j.pervac.2011.06.002.
14. King C, Beard J, Crampin AC, et al. Methodological challenges in measuring vaccine effectiveness using population cohorts in low resource settings. *Vaccine*. 2015;33(38):4748-4755. doi:10.1016/j.vaccine.2015.07.062.
15. Lahariya C. Vaccine epidemiology: A review. *J Fam Med Prim Care*. 2016;5(1):7. doi:10.4103/2249-4863.184616.
16. Sackett DL. Clinical epidemiology: What, who, and whither. *J Clin Epidemiol*. 2002;55(12):1161-1166. doi:10.1016/S0895-4356(02)00521-8.
17. Ezzati M, Hoorn S.V, Lopez A.D, Danaei G, Rodgers A, Mathers C.D MCJL. Global Burden of Disease and Risk Factors. In: Lopez A. D, Mathers C. D , Ezzati M., Jamison D. T MCJ, ed. *Comparative Quantification of Mortality and Burden of Disease Attributable to Selected Risk Factors*. New York: Oxford University Press; 2006.
18. Ward MM. Estimating disease prevalence and incidence using administrative data: Some assembly required. *J Rheumatol*. 2013;40(8):1241-1243. doi:10.3899/jrheum.130675.
19. Nelson EK. Epidemiology of Infectious Disease: General Principles. In: Nelson EK, Williams CM, eds. *Infectious Disease Epidemiology: Theory and Practice*. 3rd ed. Burlington: Jones & Bartlett Learning; 2014:19-44.
20. Greenland S, Rothman KJ. Measures of occurence. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008:32-50.
21. Hillebrand K, Bricout H, Schulze-Rath R, Schink T, Garbe E. Incidence of herpes zoster and its complications in Germany, 2005-2009. *J Infect*. 2015;70(2):178-186. doi:10.1016/j.jinf.2014.08.018.
22. World Health Organization (WHO). *Global Action Plan for the Prevention and Control of Noncommunicable Diseases 2013-2020*.; 2013.
23. Franco EL,Rohan TE VL. Epidemiologic evidence and human papillomavirus infection as a necessary cause of cervical cancer. *J Natl Cancer Inst*. 1999;91(6):506-511.
24. Spence AR, Franco EL, Ferenczy A. The role of human papillomaviruses in cancer: Evidence to date. *Am J Cancer*. 2005;4(1):49-64. doi:10.2165/00024669-200504010-00004.
25. Hausen H zur. Papillomavirus infections — a major cause of human cancers. *Biochim*

Biophys Acta - Rev Cancer. 1996;1288(2):F55-F78. doi:10.1016/0304-419X(96)00020-0.

26. Harper DM, Franco EL, Wheeler CM, et al. Sustained efficacy up to 4·5 years of a bivalent L1 virus-like particle vaccine against human papillomavirus types 16 and 18: follow-up from a randomised control trial. *Lancet*. 2006;367(9518):1247-1255. doi:10.1016/S0140-6736(06)68439-0.
27. Muñoz N, Bosch FX, de Sanjosé S, et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med*. 2003;348(6):518-527. doi:10.1056/NEJMoa021641.
28. Garland SM, Molesworth EG, Machalek DA, Cornall AM, Tabrizi SN. How to best measure the effectiveness of male human papillomavirus vaccine programmes? *Clin Microbiol Infect*. 2015;21(9):834-841. doi:10.1016/j.cmi.2015.05.038.
29. Horn J, Damm O, Kretzschmar MEE, et al. Estimating the long-term effects of HPV vaccination in Germany. *Vaccine*. 2013;31(19):2372-2380. doi:10.1016/j.vaccine.2013.03.006.
30. Nagel M a, Mahalingam R, Cohrs RJ, Gilden D. Virus vasculopathy and stroke: an under-recognized cause and treatment target. *Infect Disord Drug Targets*. 2010;10(2):105-111. doi:BSP/ID-DT/E-Pub/0015-10-2 [pii].
31. Marra F, Ruckenstein J, Richardson K. A meta-analysis of stroke risk following herpes zoster infection. *BMC Infect Dis*. 2017;17(1):198. doi:10.1186/s12879-017-2278-z.
32. Min-Chul Kim, MD, Sung-Cheol Yun, PhD, Sun U. Kwon, MD, Han-Bin Lee, MD, Pil Hyung Lee, MD, Seung-Whan Lee, MD, Sang-Oh Lee, MD, Sang-Ho Choi, MD, Yang Soo Kim, MD, Jun Hee Woo, MD, Sung-Han Kim M. Increased Risks for Stroke and Acute Myocardial Infarction After Herpes Zoster: A Population-Based Propensity Score-Matched Study. 2016:3 (suppl_1): 913. doi:10.1093/ofid/ofw194.78.
33. Schink T, Behr S, Thöne K, Bricout H, Garbe E. Risk of stroke after herpes zoster - Evidence from a German self-controlled case-series study. *PLoS One*. 2016;11(11):1-11. doi:10.1371/journal.pone.0166554.
34. Fine P, Eames K, Heymann DL. “Herd immunity”: A rough guide. *Clin Infect Dis*. 2011;52(7):911-916. doi:10.1093/cid/cir007.
35. Rouderfer V, Becker NG, Hethcote HW. Waning Immunity and its effects on vaccination schedules. *Math Biosci*. 1994;124(1):59-82. doi:10.1016/0025-5564(94)90024-8.
36. Minor PD. Live attenuated vaccines: Historical successes and current challenges. *Virology*. 2015;479-480:379-392. doi:10.1016/j.virol.2015.03.032.
37. Sanders B, Koldijk M, Schuitemaker H. Inactivated Viral Vaccines. In: Nunnally B, Turula V, Sitrin R, eds. *Vaccine Analysis: Strategies, Principles, and Control.* Springer,

Berlin, Heidelberg; 2015:45-80.

38. Liljeqvist S, Ståhl S. Production of recombinant subunit vaccines: Protein immunogens, live delivery systems and nucleic acid vaccines. *J Biotechnol*. 1999;73(1):1-33. doi:10.1016/S0168-1656(99)00107-8.
39. Wattigney W a, Mootrey GT, Braun MM, Chen RT. Surveillance for poliovirus vaccine adverse events, 1991 to 1998: impact of a sequential vaccination schedule of inactivated poliovirus vaccine followed by oral poliovirus vaccine. *Pediatrics*. 2001;107(5):E83. doi:10.1542/peds.107.5.e83.
40. Halloran ME, Longini IM. Emerging, evolving, and established infectious diseases and interventions. *Science (80-)*. 2014;345(6202):1292-1294. doi:10.1126/science.1254166.
41. Cunningham AL, Lal H, Kovac M, et al. Efficacy of the Herpes Zoster Subunit Vaccine in Adults 70 Years of Age or Older. *N Engl J Med*. 2016;375(11):1019-1032. doi:10.1056/NEJMoa1603800.
42. Oxman MN, Levin MJ, Johnson GR, et al. A Vaccine to Prevent Herpes Zoster and Postherpetic Neuralgia in Older Adults. *N Engl J Med*. 2005;352(22):2271-2284. doi:10.1056/NEJMoa051016.
43. Frost L, Johansen P, Pedersen S, Veien N, Aabel??stergaard P, Nielsen MH. Persistent Subcutaneous Nodules in Children Hyposensitized with Aluminium???Containing Allergen Extracts. *Allergy*. 1985;40(5):368-372. doi:10.1111/j.1398-9995.1985.tb00249.x.
44. Netterlid E, Hinds??n M, Siemund I, et al. Does allergen-specific immunotherapy induce contact allergy to aluminium? *Acta Derm Venereol*. 2013;93(1):50-56. doi:10.2340/00015555-1409.
45. Petrovsky N. Comparative Safety of Vaccine Adjuvants: A Summary of Current Evidence and Future Needs. *Drug Saf*. 2015;38(11):1059-1074. doi:10.1007/s40264-015-0350-4.
46. Jara LJ, García-Collinot G, Medina G, et al. Severe manifestations of autoimmune syndrome induced by adjuvants (Shoenfeld's syndrome). *Immunologic Research*. 2016:1-9.
47. Pasquale A, Preiss S, Silva F, Garçon N. Vaccine Adjuvants: from 1920 to 2015 and Beyond. *Vaccines*. 2015;3(2):320-343. doi:10.3390/vaccines3020320.
48. Robert Koch-Institut (RKI). Ständige Impfkommission (STIKO). www.rki.de/DE/Content/Kommissionen/STIKO/stiko_node.html. Published 2017. Accessed July 5, 2017.
49. Expanded Programme on Immunization (EPI). Measuring impact of Streptococcus pneumoniae and Haemophilus influenzae type b conjugate vaccination. 2012.

http://apps.who.int/iris/bitstream/10665/75835/1/WHO_IVB_12.08_eng.pdf.

50. World Health Organization (WHO). Evaluation of influenza vaccine effectiveness: a guide to the design and interpretation of observational studies. 2017.
51. European Centre for Disease Prevention and Control (ECDC). Protocol for case-control studies to measure pandemic and seasonal influenza vaccine effectiveness in the European Union and European Economic Area Member States. 2009. https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/0907_TED_Influenza_AH1N1_Measuring_Influenza_Vaccine_Effectiveness_Protocol_Case_Control_Studies.pdf.
52. Halloran ME, Haber M, Longini IM, Struchiner CJ. Direct and indirect effects in vaccine efficacy and effectiveness. *Am J Epidemiol*. 1991;133(4):323-331. doi:10.1002/ajmg.a.33992.
53. Suvarna V. Phase IV of Drug Development. *Perspect Clin Res*. 2010;1(2):57-60. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3148611&tool=pmcentrez&rendertype=abstract.
54. Spieth PM, Kubasch AS, Penzlin AI, Illigens BMW, Barlinn K, Siepmann T. Randomized controlled trials ? A matter of design. *Neuropsychiatr Dis Treat*. 2016;12:1341-1349. doi:10.2147/NDT.S101938.
55. Rid A, Saxena A, Baqui AH, et al. Placebo use in vaccine trials: Recommendations of a WHO expert panel. *Vaccine*. 2014;32(37):4708-4712. doi:10.1016/j.vaccine.2014.04.022.
56. Langan SM, Smeeth L, Margolis DJ, Thomas SL. Herpes Zoster Vaccine Effectiveness against Incident Herpes Zoster and Post-herpetic Neuralgia in an Older US Population: A Cohort Study. *PLoS Med*. 2013;10(4). doi:10.1371/journal.pmed.1001420.
57. Jackson L a, Neuzil KM, Yu O, et al. Effectiveness of pneumococcal polysaccharide vaccine in older adults. *N Engl J Med*. 2003;348(18):1747-1755. doi:10.1056/NEJMoa022678.
58. Savage RD, Winter AL, Rosella LC, et al. Strengths and limitations of assessing influenza vaccine effectiveness using routinely collected, passive surveillance data in Ontario, Canada, 2007 to 2012: balancing efficiency versus quality. *Euro Surveill Bull Eur sur les Mal Transm = Eur Commun Dis Bull*. 2015;20(16). doi:10.2807/1560-7917.ES2015.20.16.21100.
59. Crowe E, Pandeya N, Brotherton JML, et al. Effectiveness of quadrivalent human papillomavirus vaccine for the prevention of cervical abnormalities: case-control study nested within a population based screening programme in Australia. *BMJ*. 2014;348(19):g1458. doi:10.1136/bmj.g1458.

60. Whitney CG, Pilishvili T, Farley MM, et al. Effectiveness of seven-valent pneumococcal conjugate vaccine against invasive pneumococcal disease: a matched case-control study. *Lancet*. 2006;368(9546):1495-1502. doi:10.1016/S0140-6736(06)69637-2.
61. Oberle D, Pavel J, Mayer G, Geisler P, Keller-Stanislawski B. Retrospective multicenter matched case–control study on the risk factors for narcolepsy with special focus on vaccinations (including pandemic influenza vaccination) and infections in Germany. *Sleep Med*. 2017;34:71-83. doi:10.1016/j.sleep.2017.02.026.
62. Thöne K, Kollhorst B, Schink T. Non-Steroidal Anti-Inflammatory Drug Use and the Risk of Acute Myocardial Infarction in the General German Population: A Nested Case–Control Study. *Drugs - Real World Outcomes*. 2017;4(3):127-137. doi:10.1007/s40801-017-0113-x.
63. Jackson LA, Jackson ML, Nelson JC, Neuzil KM, Weiss NS. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. *Int J Epidemiol*. 2006;35(2):337-344. doi:10.1093/ije/dyi274.
64. Haber M, An Q, Foppa IM, Shay DK, Ferdinands JM, Orenstein WA. A Probability Model for Evaluating the Bias and Precision of Influenza Vaccine Effectiveness Estimates from Case-Control Studies. *Epidemiol Infect*. 2015;143(7):1417-1426. doi:10.1017/S0950268814002179.A.
65. Verani JR, Baqui AH, Broome C V., et al. Case-control vaccine effectiveness studies: Preparation, design, and enrollment of cases and controls. *Vaccine*. 2017;35(25):3295-3302. doi:10.1016/j.vaccine.2017.04.037.
66. Centers for Disease Control and Prevention (CDC). VaxView. https://www.cdc.gov/vaccines/vaxview/index.html. Published 2016. Accessed August 22, 2017.
67. World Health Organization (WHO). Immunization coverage. Immunization, Vaccines and Biologicals. http://www.who.int/immunization/monitoring_surveillance/routine/coverage/en/. Published 2017. Accessed August 22, 2017.
68. Rieck T, Feig M, Eckmanns T, Benzler J, Siedler a, Wichmann O. Vaccination coverage among children in Germany estimated by analysis of health insurance claims data. *Hum Vaccines Immunother*. 2014;10(2):476-484. doi:10.4161/hv.26986.
69. Robert-Koch-Institut. Impfquoten der Rotavirus-, Masern-, HPV-und Influenza-Impfung in Deutschland. *Epidemiol Bull*. 2017;(1):1-16. https://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2017/Ausgaben/01_17.pdf?__blob=publicationFile.
70. Siedler A, Rieck T, Reuss A, et al. Estimating vaccination coverage in the absence of

immunisation registers--the German experience. *Euro Surveill Bull Eur sur les Mal Transm = Eur Commun Dis Bull*. 2012;17(17). doi:20152 [pii].

71. Hense S, Hillebrand K, Horn J, Mikolajczyk R, Schulze-Rath R, Garbe E. HPV vaccine uptake after introduction of the vaccine in Germany: an analysis of administrative data. *Hum Vaccin Immunother*. 2014;10(6):1729-1733. doi:10.4161/hv.28450.
72. Baandrup L, Blomberg M, Dehlendorff C, Sand C, Andersen KK, Kjaer SK. Significant decrease in the incidence of genital warts in young Danish women after implementation of a national human papillomavirus vaccination program. *Sex Transm Dis*. 2013;40(2):130-135. doi:10.1097/OLQ.0b013e31827bd66b.
73. Donovan B, Franklin N, Guy R, et al. Quadrivalent human papillomavirus vaccination and trends in genital warts in Australia: analysis of national sentinel surveillance data. *Lancet Infect Dis*. 2011;11(1):39-44. doi:10.1016/S1473-3099(10)70225-5.
74. Howell-Jones R, Soldan K, Wetten S, et al. Declining genital Warts in young women in england associated with HPV 16/18 vaccination: an ecological study. *J Infect Dis*. 2013;208(9):1397-1403. doi:10.1093/infdis/jit361.
75. Cutts FT, Claquin P, Danovaro-Holliday MC, Rhoda DA. Monitoring vaccination coverage: Defining the role of surveys. *Vaccine*. 2016;34(35):4103-4109. doi:10.1016/j.vaccine.2016.06.053.
76. Hyde TB, Dentz H, Wang SA, et al. The impact of new vaccine introduction on immunization and health systems: A review of the published literature. *Vaccine*. 2012;30(45):6347-6358. doi:10.1016/j.vaccine.2012.08.029.
77. Hariri S, Johnson ML, Bennett NM, et al. Population-based trends in high-grade cervical lesions in the early human papillomavirus vaccine era in the United States. *Cancer*. 2015;121(16):2775-2781. doi:10.1002/cncr.29266.
78. Grijalva CG, Nuorti JP, Arbogast PG, Martin SW, Edwards KM, Griffin MR. Decline in pneumonia admissions after routine childhood immunisation with pneumococcal conjugate vaccine in the USA: a time-series analysis. *Lancet*. 2007;369(9568):1179-1186. doi:10.1016/S0140-6736(07)60564-9.
79. Dijkstra F, Donker GA, Wilbrink B, Van Gageldonk-Lafeber AB, Van Der Sande MAB. Long time trends in influenza-like illness and associated determinants in The Netherlands. *Epidemiol Infect*. 2009;137(4):473-479. doi:10.1017/S095026880800126X.
80. Hanquet G, Valenciano M, Simondon F, Moren A. Vaccine effects and impact of vaccination programmes in post-licensure studies. *Vaccine*. 2013;31(48):5634-5642. doi:10.1016/j.vaccine.2013.07.006.
81. Thöne K, Horn J, Mikolajczyk R. Evaluation of vaccination herd immunity effects for anogenital warts in a low coverage setting with human papillomavirus vaccine—an

interrupted time series analysis from 2005 to 2010 using health insurance data. *BMC Infect Dis*. 2017;17(1):564. doi:10.1186/s12879-017-2663-7.

82. John TJ, Samuel R. Herd immunity and herd effect: New insights and definitions. *Eur J Epidemiol*. 2000;16(7):601-606. doi:10.1023/A:1007626510002.
83. Moss WJ, Ota MO. Measles control, elimination, and eradication. In: Nelson E, Williams C, eds. *Infectious Disease Epidemiology: Theory and Practice*. 3rd ed. Burlington: Jones & Bartlett Learning; 2014:485-507.
84. Dietz K. The estimation of the basic reproduction number for infectious diseases. *Stat Methods Med Res*. 1993;2(1):23-41. doi:10.1177/096228029300200103.
85. Heffernan JM, Smith RJ, Wahl LM. Perspectives on the basic reproductive ratio. *J R Soc Interface*. 2005;2(4):281-293. doi:10.1098/rsif.2005.0042.
86. Adegbola RA, Secka O, Lahai G, et al. Elimination of Haemophilus influenzae type b (Hib) disease from The Gambia after the introduction of routine immunisation with a Hib conjugate vaccine: A prospective study. *Lancet*. 2005;366(9480):144-150. doi:10.1016/S0140-6736(05)66788-8.
87. Moulton LH, Chung S, Croll J, Reid R, Weatherholtz RC, Santosham M. Estimation of the indirect effect of Haemophilus influenzae type b conjugate vaccine in an American Indian population. *Int J Epidemiol*. 2000;29(4):753-756. doi:10.1093/ije/29.4.753.
88. Plans P. New preventive strategy to eliminate measles, mumps and rubella from Europe based on the serological assessment of herd immunity levels in the population. *Eur J Clin Microbiol Infect Dis*. 2013;32(7):961-966. doi:10.1007/s10096-013-1836-6.
89. Garnett GP. Role of herd immunity in determining the effect of vaccines against sexually transmitted disease. *J Infect Dis*. 2005;191 Suppl:S97-106. doi:10.1086/425271.
90. Lowy DR, Schiller JT. Reducing HPV-associated cancer globally. *Cancer Prev Res*. 2012;5(1):18-23. doi:10.1158/1940-6207.CAPR-11-0542.
91. Lenzi A, Mirone V, Gentile V, et al. Rome Consensus Conference - statement; human papilloma virus diseases in males. *BMC Public Health*. 2013;13(1):117. doi:10.1186/1471-2458-13-117.
92. Panagiotopoulos T, Antoniadou I, Valassi-Adam E. Increase in congenital rubella occurrence after immunisation in Greece: retrospective survey and systematic review. *BMJ*. 1999;319(7223):1462-1467. doi:10.1136/bmj.319.7223.1462.
93. Thwaites CL, Loan HT. Eradication of tetanus. *Br Med Bull*. 2015;116(1):69-77. doi:10.1093/bmb/ldv044.
94. Maglione M a., Das L, Raaen L, et al. Safety of Vaccines Used for Routine Immunization of US Children: A Systematic Review. *Pediatrics*. 2014;134(2):325-337.

doi:10.1542/peds.2014-1079.

95. Chen RT, DeStefano F, Davis RL, et al. The Vaccine Safety Datalink: Immunization research in health maintenance organizations in the USA. *Bull World Health Organ*. 2000;78(2):186-194.
96. Schink T, Holstiege J, Kowalzik F, Zepp F, Garbe E. Risk of febrile convulsions after MMRV vaccination in comparison to MMR or MMR+V vaccination. *Vaccine*. 2014;32(6):645-650. doi:10.1016/j.vaccine.2013.12.011.
97. Garbe E, Suissa S. Pharmacoepidemiology. In: Ahrens W, Pigeot I, eds. *Handbook of Epidemiology*. Second Edi. New York: Springer Science+Business Media; 2014:1876-1915.
98. Garbe E, Pigeot I. Der Nutzen großer Gesundheitsdatenbanken für die Arzneimittelrisikoforschung. *Bundesgesundheitsblatt - Gesundheitsforsch - Gesundheitsschutz*. 2015;58(8):829-837. doi:10.1007/s00103-015-2185-7.
99. Lipsitch M, Jha A, Simonsen L. Observational studies and the difficult quest for causality: Lessons from vaccine effectiveness and impact studies. *Int J Epidemiol*. 2016;45(6):2060-2074. doi:10.1093/ije/dyw124.
100. Berlin JA, Glasser SC, Ellenberg SS. Adverse event detection in drug development: Recommendations and obligations beyond phase 3. *Am J Public Health*. 2008;98(8):1366-1371. doi:10.2105/AJPH.2007.124537.
101. Black S, Eskola J, Siegrist CA, et al. Importance of background rates of disease in assessment of vaccine safety during mass immunisation with pandemic H1N1 influenza vaccines. *Lancet*. 2009;374(9707):2115-2122. doi:10.1016/S0140-6736(09)61877-8.
102. Arzneimittelkommission der deutschen Ärzteschaft, Wissenschaftlicher Fachausschuss der Bundesärztekammer. Rote-Hand-Brief zu TicoVac. https://www.akdae.de/Arzneimittelsicherheit/RHB/Archiv/2001/TicoVac.pdf. Published 2001. Accessed October 23, 2017.
103. Arzneimittelkommission der deutschen Ärzteschaft, Wissenschaftlicher Fachausschuss der Bundesärztekammer. Rote-Hand-Brief zu Hexavac®. https://www.akdae.de/Arzneimittelsicherheit/RHB/Archiv/2005/64_20050921.pdf. Published 2005. Accessed October 23, 2017.
104. Offit PA, Coffin SE. Communicating science to the public: MMR vaccine and autism. *Vaccine*. 2003;22(1):1-6. doi:10.1016/S0264-410X(03)00532-2.
105. Gerber JS, Offit P a. Vaccines and Autism: A Tale of Shifting Hypotheses. *Clin Infect Dis*. 2009;48(table 1):456-461. doi:10.1086/596476.
106. The Editors of The Lancet. Retraction - Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet*.

2010;375(9713):445. doi:10.1016/S0140-6736(10)60175-4.

107. Martínez-Lavín M. Hypothesis: Human papillomavirus vaccination syndrome—small fiber neuropathy and dysautonomia could be its underlying pathogenesis. *Clin Rheumatol*. 2015;34(7):1165-1169. doi:10.1007/s10067-015-2969-z.
108. Palmieri B, Poddighe D, Vadalà M, Laurino C, Carnovale C, Clementi E. Severe somatoform and dysautonomic syndromes after HPV vaccination: case series and review of literature. *Immunologic Research*. 2016:1-11.
109. Gee J, Weinbaum C, Sukumaran L, Markowitz LE. Quadrivalent HPV vaccine safety review and US safety monitoring plans for nine-valent HPV vaccine. *Hum Vaccin Immunother*. 2016;5515(April):1-12. doi:10.1080/21645515.2016.1168952.
110. Head MG, Wind-Mozley M, Flegg P. Inadvisable anti-vaccination sentiment: Human Papilloma Virus immunisation falsely under the microscope. *npj Vaccines*. 2017;2(6):1. doi:10.1038/s41541-017-0004-x.
111. European Medicines Agency (EMA). Review concludes evidence does not support that HPV vaccines cause CRPS or POTS- Reports of CRPS and POTS after HPV vaccination are consistent with what would be expected in this age group. http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2015/11/WC500196352.pdf. Published 2015. Accessed July 1, 2017.
112. Van Der Maas NAT, Kramer MA, Jacobs BC, et al. Guillain-Barré syndrome: Background incidence rates in the Netherlands. *J Peripher Nerv Syst*. 2011;16(3):243-249. doi:10.1111/j.1529-8027.2011.00356.x.
113. Hense S, Schink T, Kreisel SH, et al. Estimation of background incidence rates of guillain-barré syndrome in Germany-A retrospective cohort study with electronic healthcare data. *Neuroepidemiology*. 2014;43(3-4):244-252. doi:10.1159/000369344.
114. Strom BL. Overview of Automated Databases in Pharmacoepidemiology. In: *Pharmacoepidemiology, Fifth Edition*. ; 2012:158-162. doi:10.1002/9781119959946.ch11.
115. Andersohn F, Garbe E. Pharmakoepidemiologische forschung mit routinedaten des gesundheitswesens. *Bundesgesundheitsblatt - Gesundheitsforsch - Gesundheitsschutz*. 2008;51(10):1135-1144. doi:10.1007/s00103-008-0648-9.
116. Cadarette SM, Wong L. An Introduction to Health Care Administrative Data. *Can J Hosp Pharm*. 2015;68(3):232-237. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4485511/pdf/cjhp-68-232.pdf.
117. Poggensee G, Reuss A, Reiter S, Siedler A. [Overview and assessment of available data sources to determine incidence of vaccine preventable diseases, vaccination coverage, and immune status in Germany]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2009;52(11):1019-1028.

doi:10.1007/s00103-009-0952-z.

118. Leibniz-Institut für Präventionsforschung und Epidemiologie – BIPS GmbH. Die Deutsche Pharmakoepidemiologische Forschungsdatenbank. https://www.bips-institut.de/forschung/forschungsinfrastrukturen/versichertenstichprobe.html. Published 2017. Accessed August 24, 2017.
119. Federal Joint Commitee. Richtlinie des Gemeinsamen Bundesausschusses über Schutzimpfungen nach § 20 d, Abs. 1 SGB V (Schutzimpfungs-Richtlinie/SI-RL) in der Fassung vom 18. Oktober 2007, zuletzt geändert am 24. November 2011., 2012. https://www.g-ba.de/downloads/62-492-581/SI-RL_2011-11-24.pdf. Published 2011. Accessed August 19, 2017.
120. Gershon AA, Breuer J, Cohen JI, et al. Varicella zoster virus infection. *Nat Rev Dis Prim*. 2015;(July):15016. doi:10.1038/nrdp.2015.16.
121. Robert Koch-Institut (RKI). Windpocken, Herpes zoster (Gürtelrose). RKI-Ratgeber für Ärzte. https://www.rki.de/DE/Content/Infekt/EpidBull/Merkblaetter/Ratgeber_Varizellen.html#doc2374554bodyText3. Published 2016. Accessed August 24, 2017.
122. Gilden D, Mahalingam R, Nagel M., Pugazhenthi S, Cohrs RJ. The neurobiology of varicella zoster virus infection. *Neuropatholy*. 2011;37(5):441-463. doi:10.1111/j.1365-2990.2011.01167.x.
123. Studahl M, Petzold M, Cassel T. Disease burden of herpes zoster in Sweden - predominance in the elderly and in women - a register based study. *BMC Infect Dis*. 2013;13(1):586. doi:10.1186/1471-2334-13-586.
124. Yawn BP, Gilden D. The global epidemiology of herpes zoster. *Neurology*. 2013;81(10):928-930. doi:10.1212/WNL.0b013e3182a3516e.
125. Schmidt SAJ, Kahlert J, Vestergaard M, Schønheyder HC, Sørensen HT. Hospital-based herpes zoster diagnoses in Denmark: rate, patient characteristics, and all-cause mortality. *BMC Infect Dis*. 2016;16(1):99. doi:10.1186/s12879-016-1369-6.
126. Johnson RW, Alvarez-Pasquin M-J, Bijl M, et al. Herpes zoster epidemiology, management, and disease and economic burden in Europe: a multidisciplinary perspective. *Ther Adv vaccines*. 2015;3(4):109-120. doi:10.1177/2051013615599151.
127. Brisson M, Gay NJ, Edmunds WJ, Andrews NJ. Exposure to varicella boosts immunity to herpes-zoster: Implications for mass vaccination against chickenpox. *Vaccine*. 2002;20(19-20):2500-2507. doi:10.1016/S0264-410X(02)00180-9.
128. Dworkin RH, Portenoy RK. Pain and its persistence in herpes zoster. *Pain*. 1996;67(2-3):241-251. doi:10.1016/0304-3959(96)03122-3.
129. Nagel MA, Gilden D. Update on varicella zoster virus vasculopathy. *Curr Infect Dis Rep*. 2014;16(6). doi:10.1007/s11908-014-0407-z.

130. Johnson RW, Wasner G, Saddier P, Baron R. Herpes zoster and postherpetic neuralgia: optimizing management in the elderly patient. *Drugs Aging*. 2008;25(12):991-1006. doi:10.2165/0002512-200825120-00002.
131. Robert-Koch-Institut (RKI). Mitteilung der Ständigen Impfkommission am Robert Koch-Institut (RKI) Empfehlungen der Ständigen Impfkommission (STIKO) am Robert Koch-Institut – 2017/2018. *Epidemiol Bull*. 2017;34/2017.
132. Warren-Gash C, Forbes H, Breuer J. Varicella and herpes zoster vaccine development: lessons learned. *Expert Rev Vaccines*. 2017;16(12):1191-1201. doi:10.1080/14760584.2017.1394843.
133. Palefsky JM. Human Papillomavirus-Related Disease in Men: Not Just a Women's Issue. *J Adolesc Heal*. 2010;46(4 SUPPL.). doi:10.1016/j.jadohealth.2010.01.010.
134. Braaten KP, Laufer MR. Human Papillomavirus (HPV), HPV-Related Disease, and the HPV Vaccine. *Rev Obstet Gynecol*. 2008;1(1):2-10. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2492590&tool=pmcentrez&rendertype=abstract.
135. Schwabe U, Paffrath D. Arzneiverordnungsreport 2009. *Heidelb Springer*. 2009.
136. Giuliano AR, Palefsky JM, Goldstone S, et al. Efficacy of quadrivalent HPV vaccine against HPV Infection and disease in males. *N Engl J Med*. 2011;364(5):401-411. doi:10.1056/NEJMoa0909537.
137. Ali H, Donovan B, Wand H, et al. Genital warts in young Australians five years into national human papillomavirus vaccination programme: national surveillance data. *BMJ*. 2013;346(April):f2032. doi:10.1136/bmj.f2032.
138. Hense S, Hillebrand K, Horn J, Mikolajczyk R, Schulze-Rath R, Garbe E. HPV vaccine uptake after introduction of the vaccine in Germany: An analysis of administrative data. *Hum Vaccines Immunother*. 2014;10(6):1729-1733. doi:10.4161/hv.28450.
139. Rieck T, Feig M, Deler?? Y, Wichmann O. Utilization of administrative data to assess the association of an adolescent health check-up with human papillomavirus vaccine uptake in Germany. *Vaccine*. 2014;32(43):5564-5569. doi:10.1016/j.vaccine.2014.07.105.
140. Perkins RB, Clark JA. What Affects Human Papillomavirus Vaccination Rates? A Qualitative Analysis of Providers' Perceptions. *Women's Heal Issues*. 2012;22(4). doi:10.1016/j.whi.2012.04.001.
141. Ahrens, Wolfgang; Pigeot I. *Handbook of Epidemiology*. Vol 101. Springer New York Heidelberg Dordrecht London; 2014. doi:DOI 10.1007/978-0-387-09834-0.
142. Ong CKS, Lirk P, Tan CH, Seymour RA. An evidence-based update on nonsteroidal anti-inflammatory drugs. *Clin Med Res*. 2007;5(1):19-34. doi:10.3121/cmr.2007.698.
143. Schmidt M, Lamberts M, Olsen AMS, et al. Cardiovascular safety of non-aspirin non-

steroidal anti-inflammatory drugs: Review and position paper by the working group for Cardiovascular Pharmacotherapy of the European Society of Cardiology. *Eur Heart J*. 2016;37(13):1015-1023. doi:10.1093/eurheartj/ehv505.

144. Bresalier RS, Sandler RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *N Engl J Med*. 2005;352(11):1092-1102. doi:10.1056/NEJMoa050493.
145. Postmarket Drug Safety Information for Patients and Providers - Public Health Advisory - FDA Announces Important Changes and Additional Warnings for COX-2 Selective and Non-Selective Non-Steroidal Anti-Inflammatory Drugs (NSAIDs).
146. European Medicines Agency (EMA). European Medicines Agency concludes action on COX-2 inhibitors. News and press releases. http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/news/2010/01/news_detail_000969.jsp&mid=WC0b01ac058004d5c1. Published 2005. Accessed August 26, 2017.
147. Fosbøl EL, Folke F, Jacobsen S, et al. *Cause-Specific Cardiovascular Risk Associated with Nonsteroidal Antiinflammatory Drugs among Healthy Individuals.* Vol 3.; 2010. doi:10.1161/CIRCOUTCOMES.109.861104.
148. Ray WA, Varas-Lorenzo C, Chung CP, et al. Cardiovascular risks of nonsteroidal antiinflammatory drugs in patients after hospitalization for serious coronary heart disease. *Circ Cardiovasc Qual Outcomes*. 2009;2(3):155-163. doi:10.1161/CIRCOUTCOMES.108.805689.
149. Solomon DH, Avorn J, Stürmer T, Glynn RJ, Mogun H, Schneeweiss S. Cardiovascular outcomes in new users of coxibs and nonsteroidal antiinflammatory drugs: high-risk subgroups and time course of risk. *Arthritis Rheum*. 2006;54(5):1378-1389. doi:10.1002/art.21887.
150. Graham DJ, Campen D, Hui R, et al. Risk of acute myocardial infarction and sudden cardiac death in patients treated with cyclo-oxygenase 2 selective and non-selective non-steroidal anti-inflammatory drugs: Nested case-control study. *Lancet*. 2005;365:475-481. doi:10.1016/S0140-6736(05)17864-7.
151. Bally M, Dendukuri N, Rich B, et al. Risk of acute myocardial infarction with NSAIDs in real world use: bayesian meta-analysis of individual patient data. *BMJ*. 2017;357:j1909. doi:10.1136/bmj.j1909.
152. Varas-Lorenzo C, Riera-Guardia N, Calingaert B, et al. Myocardial infarction and individual nonsteroidal anti-inflammatory drugs meta-analysis of observational studies. *Pharmacoepidemiol Drug Saf*. 2013;22(6):559-570. doi:10.1002/pds.3437.
153. Trelle S, Reichenbach S, Wandel S, et al. Cardiovascular safety of non-steroidal anti-inflammatory drugs: network meta-analysis. *BMJ*. 2011;342:c7086.

doi:10.1136/bmj.c7086.

154. Arfè A, Scotti L, Varas-Lorenzo C, et al. Non-steroidal anti-inflammatory drugs and risk of heart failure in four European countries: nested case-control study. *Bmj*. 2016;354:i4857. doi:10.1136/bmj.i4857.
155. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. 2008;61(4):344-349. doi:10.1016/j.jclinepi.2007.11.008.
156. Andrews N. Epidemiological designs for vaccine safety assessment: Methods and pitfalls. *Biologicals*. 2012;40(5):389-392. doi:10.1016/j.biologicals.2011.08.010.
157. Miller AB, Goff Jr. DC, Bammann K, Wild P. Cohort Studies. In: Ahrens W, Pigeot I, eds. *Handbook of Epidemiology*. Second Edi. New York: Springer Science+Business Media; 2014:260-289.
158. Diaz J, Terrazas S, Bierrenbach AL, et al. Effectiveness of the 10-valent pneumococcal conjugate vaccine (PCV-10) in Children in Chile: A nested case-control study using nationwide pneumonia morbidity and mortality surveillance data. *PLoS One*. 2016;11(4). doi:10.1371/journal.pone.0153141.
159. Strom BL, Kimmel SE, Sean H. *Textbook of Pharmacoepidemiology*. Second Edi. (Strom BL, Kimmel SE, Hennesy S, eds.). Chichester, UK: John Wiley & Sons Ltd; 2013.
160. Chen RT. Evaluation of vaccine safety after the events of 11 September 2001: Role of cohort and case-control studies. In: *Vaccine*. Vol 22. ; 2004:2047-2053. doi:10.1016/j.vaccine.2004.01.023.
161. Rothman KJ, Greenland S. Basic concepts. In: Ahrens W, Pigeot I, eds. *Handbook of Epidemiology*. Second Edi. New York: Springer Science+Business Media; 2014:77-119. doi:10.1007/978-0-387-09834-0.
162. Ohlmeier C, Langner I, Hillebrand K, et al. Mortality in the German Pharmacoepidemiological Research Database (GePaRD) compared to national data in Germany: results from a validation study. *BMC Public Health*. 2015;15:570. doi:10.1186/s12889-015-1943-7.
163. Gavrielov-Yusim N, Friger M. Use of administrative medical databases in population-based research. *J Epidemiol Community Health*. 2014;68(3):283-287. doi:10.1136/jech-2013-202744.
164. Biesheuvel CJ, Vergouwe Y, Oudega R, Hoes AW, Grobbee DE, Moons KGM. Advantages of the nested case-control design in diagnostic research. *BMC Med Res Methodol*. 2008;8. doi:10.1186/1471-2288-8-48.
165. Hak E, Buskens E, van Essen GA, et al. Clinical effectiveness of influenza vaccination

in persons younger than 65 years with high-risk medical conditions: the PRISMA study. *Arch Intern Med*. 2005;165(3):274-280. doi:10.1001/archinte.165.3.274.

166. Casanova L, Cortaredona S, Gaudart J, et al. Effectiveness of seasonal influenza vaccination in patients with diabetes: protocol for a nested case–control study. *BMJ Open*. 2017;7(8):e016023. doi:10.1136/bmjopen-2017-016023.
167. Mommers M, Weishoff-Houben M, Swaen GMH, et al. Infant immunization and the occurrence of atopic disease in Dutch and German children: A nested case-control study. *Pediatr Pulmonol*. 2004;38(4):329-334. doi:10.1002/ppul.20089.
168. Black S, Shinefield H, Ray P, et al. Risk of hospitalization because of aseptic meningitis after measles- mumps-rubella vaccination in one- to two-year-old children: An analysis of the Vaccine Safety Datalink (VSD) project. *Pediatr Infect Dis J*. 1997;16(5):500-503. doi:10.1097/00006454-199705000-00009.
169. Black C, Kaye JA, Jick H. MMR vaccine and idiopathic thrombocytopaenic purpura. *Br J Clin Pharmacol*. 2003;55(1):107-111. doi:10.1046/j.1365-2125.2003.01790.x.
170. DeStefano F, Mullooly JP, Okoro C a, et al. Childhood vaccinations, vaccination timing, and risk of type 1 diabetes mellitus. *Pediatrics*. 2001;108(6):E112. doi:10.1542/peds.108.6.e112.
171. Rothman KJ, Greenland S, Lash TL. Case-Control Studies. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. Second Edi. Philadelphia: Lippincott Williams & Wilkins; 2008:111-127.
172. Wacholder S, Mclaughlin JK, Silverman DT, et al. Selection of Controls in Case-Control Studies I. Principles. *Am J Epidemiol*. 1992;135(9):1019-1028. doi:10.1002/bdra.20388.
173. Suissa S. Novel approaches to pharmacoepidemiology study design and statistical analysis. In: Strom BL, Kimmel SE, eds. *Textbook of Pharmacoepidemiology*. John Wiley & Sons, Ltd; 2006.
174. Glanz JM, McClure DL, Xu S, et al. Four different study designs to evaluate vaccine safety were equally validated with contrasting limitations. *J Clin Epidemiol*. 2006;59(8):808-818. doi:10.1016/j.jclinepi.2005.11.012.
175. Andersohn F. Identifizierung unerwünschter Arzneimittelwirkungen nach Marktzulassung (Habilitationsschrift). *Charité-Universitätsmedizin Berlin*. 2010.
176. Kronman MP, Gerber JS, Newland JG, Hersh AL. Database research for pediatric infectious diseases. *J Pediatric Infect Dis Soc*. 2015;4(2):143-150. doi:10.1093/jpids/piv007.
177. Benichou J, Palta M. Rates, Risks, Measures of Association and Impact. In: Ahrens W, Pigeot I, eds. *Handbook of Epidemiology*. Second Edi. New York: Springer Science+Business Media; 2014:125-179.

178. Faresjö T, Faresjö Å. To match or not to match in epidemiological studies - Same outcome but less power. *Int J Environ Res Public Health*. 2010;7(1):325-332. doi:10.3390/ijerph7010325.
179. Bloom MS, Schisterman EF, Hediger ML. The use and misuse of matching in case-control studies: the example of polycystic ovary syndrome. *Fertil Steril*. 2007;88(3):707-710. doi:10.1016/j.fertnstert.2006.11.125.
180. Desai RJ, Glynn RJ, Wang S, Gagne JJ. Performance of disease risk score matching in nested case-control studies: A simulation study. *Am J Epidemiol*. 2016;183(10). doi:10.1093/aje/kwv269.
181. Jackson LA, Nelson JC, Benson P, et al. Functional status is a confounder of the association of influenza vaccine and risk of all cause mortality in seniors. *Int J Epidemiol*. 2006;35(2):345-352. doi:10.1093/ije/dyi275.
182. Lone NI, Simpson C, Kavanagh K, et al. Seasonal Influenza Vaccine Effectiveness in the community (SIVE): protocol for a cohort study exploiting a unique national linked data set. *BMJ Open*. 2012;2(2):e001019. doi:10.1136/bmjopen-2012-001019.
183. Fukushima W, Hirota Y. Basic principles of test-negative design in evaluating influenza vaccine effectiveness. *Vaccine*. 2017;35(36):4796-4800. doi:10.1016/j.vaccine.2017.07.003.
184. Jackson ML, Nelson JC. The test-negative design for estimating influenza vaccine effectiveness. *Vaccine*. 2013;31(17):2165-2168. doi:10.1016/j.vaccine.2013.02.053.
185. Sullivan SG, Tchetgen EJT, Cowling BJ. Theoretical Basis of the Test-Negative Study Design for Assessment of Influenza Vaccine Effectiveness. *Am J Epidemiol*. 2016;184(5):345-353. doi:10.1093/aje/kww064.
186. Halloran ME, Struchiner CJ. Causal inference in infectious diseases. *Epidemiology*. 1995;6(2):142-151. http://www.ncbi.nlm.nih.gov/pubmed/7742400.
187. Halloran ME, Struchiner CJ, Longini IM. Study designs for evaluating different efficacy and effectiveness aspects of vaccines. *Am J Epidemiol*. 1997;146(10):789-803. doi:10.1093/oxfordjournals.aje.a009196.
188. Foppa IM, Ferdinands JM, Chaves SS, et al. The case test-negative design for studies of the effectiveness of influenza vaccine in inpatient settings. *Int J Epidemiol*. 2016;45(6):2052-2059. doi:10.1093/ije/dyw022.
189. Belongia EA, Kieke BA, Donahue JG, et al. Effectiveness of inactivated influenza vaccines varied substantially with antigenic match from the 2004-2005 season to the 2006-2007 season. *J Infect Dis*. 2009;199(2):159-167. doi:10.1086/595861.
190. De Serres G, Skowronski DM, Wu XW, Ambrose CS. The test-negative design: Validity, accuracy and precision of vaccine efficacy estimates compared to the gold standard of randomised placebo-controlled clinical trials. *Eurosurveillance*.

2013;18(37). doi:10.2807/1560-7917.ES2013.18.37.20585.

191. Gosselin V, Généreux M, Gagneur A, Petit G. Effectiveness of rotavirus vaccine in preventing severe gastroenteritis in young children according to socioeconomic status. *Hum Vaccines Immunother*. 2016;12(10):2572-2579. doi:10.1080/21645515.2016.1189038.
192. Bray F, Parkin DM. Descriptive Studies. In: Ahrens W, Pigeot I, eds. *Handbook of Epidemiology*. Second Edi. New York: Springer Science+Business Media; 2014:188-250.
193. Soumerai SB, Starr D, Majumdar SR. How Do You Know Which Health Care Effectiveness Research You Can Trust? A Guide to Study Design for the Perplexed. *Prev Chronic Dis*. 2015;12:150187. doi:10.5888/pcd12.150187.
194. Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods. *Annu Rev Public Health*. 1995;16:61-81. doi:10.1146/annurev.publhealth.16.1.61.
195. Morgenstern H. Ecologic Studies. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. Second Edi. Philadelphia: Lippincott Williams & Wilkins; 2008:511-530.
196. Wagner a. K, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. *J Clin Pharm Ther*. 2002;27(4):299-309. doi:10.1046/j.1365-2710.2002.00430.x.
197. Lopez Bernal J, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *Int J Epidemiol*. 2016:1-8. doi:10.1093/ije/dyw098.
198. Shadish WR, Cook TD, Campbell DT. Experimental and Quasi-Experimental for Generalized Designs Causal Inference. *Handb Ind Organ Psychol*. 2002;223:623. doi:10.1198/jasa.2005.s22.
199. Campbell, D. T., Stanley, J. C., & Gage NL. Experimental and quasi-experimental designs in prevention research. *NIDA Res Monogr*. 1963;107:140-158. doi:10.1016/0306-4573(84)90053-0.
200. Mikolajczyk RT, Kraut A a, Horn J, Schulze-Rath R, Garbe E. Changes in incidence of anogenital warts diagnoses after the introduction of human papillomavirus vaccination in Germany-an ecologic study. *Sex Transm Dis*. 2013;40(1):28-31. doi:10.1097/OLQ.0b013e3182756efd.
201. Pape UJ, Millett C, Lee JT, Car J, Majeed A. Disentangling secular trends and policy impacts in health studies: use of interrupted time series analysis. *J R Soc Med*. 2013;106(4):124-129. doi:10.1258/jrsm.2012.110319.
202. Schneeweiss S, Maclure M, Walker AM, Grootendorst P, Soumerai SB. On the evaluation of drug benefits policy changes with longitudinal claims data: The policy

maker's versus the clinician's perspective. *Health Policy (New York)*. 2001;55(2):97-109. doi:10.1016/S0168-8510(00)00120-2.

203. Hawken S, Potter BK, Little J, et al. The use of relative incidence ratios in self-controlled case series studies: an overview. *BMC Med Res Methodol*. 2016;16(1):126. doi:10.1186/s12874-016-0225-0.
204. Farrington CP. Relative Incidence Estimation from Case Series for Vaccine Safety Evaluation. *Biometrics*. 1995;51(1):228-235. doi:10.2307/2533328.
205. Petersen I, Douglas I, Whitaker H. Self controlled case series methods: an alternative to standard epidemiological study designs. *BMJ*. 2016;354:i4515. doi:10.1136/bmj.i4515.
206. Whitaker HJ, Hocine MN, Farrington CP. The methodology of self-controlled case series studies. *Stat Methods Med Res*. 2009;18:7-26. doi:10.1177/0962280208092342.
207. Whitaker HJ, Farrington CP, Spiessens B, Musonda P. Tutorial in biostatistics: The self-controlled case series method. *Stat Med*. 2006;25(10):1768-1797. doi:10.1002/sim.2302.
208. Porta M. *A Dictionary of Epidemiology*. Oxford University Press; 2014.
209. Gordis L. *Epidemiology, 4th Edition*. Vol 62.; 2009.
210. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci*. 2001;56(3):M146-M156. doi:10.1093/gerona/56.3.M146.
211. Brookhart MA, Patrick AR, Dormuth C, et al. Adherence to lipid-lowering therapy and the use of preventive health services: An investigation of the healthy user effect. *Am J Epidemiol*. 2007;166(3):348-354. doi:10.1093/aje/kwm070.
212. Dormuth CR, Patrick AR, Shrank WH, et al. Statin adherence and risk of accidents a cautionary tale. *Circulation*. 2009;119(15):2051-2057. doi:10.1161/CIRCULATIONAHA.108.824151.
213. Nichol KL, Nordin JD, Nelson DB, Mullooly JP, Hak E. Effectiveness of Influenza Vaccine in the Community-Dwelling Elderly. *N Engl J Med*. 2007;357(14):1373-1381. doi:10.1056/NEJMp1415160.
214. Simonsen L, Reichert TA, Viboud C, Blackwelder WC, Taylor RJ, Miller MA. Impact of influenza vaccination on seasonal mortality in the US elderly population. *Arch Intern Med*. 2005;165(3):265-272. doi:10.1001/archinte.165.3.265.
215. Simonsen L, Taylor RJ, Viboud C, Miller MA, Jackson LA. Mortality benefits of influenza vaccination in elderly people: an ongoing controversy. *Lancet Infect Dis*. 2007;7(10):658-666. doi:10.1016/S1473-3099(07)70236-0.
216. Campitelli MA, Rosella LC, Stukel TA, Kwong JC. Influenza vaccination and all-cause

mortality in community-dwelling elderly in Ontario, Canada, a cohort study. *Vaccine*. 2010;29(2):240-246. doi:10.1016/j.vaccine.2010.10.049.

217. Johnstone J, Loeb M, Teo KK, et al. Influenza vaccination and major adverse vascular events in high-risk patients. *Circulation*. 2012;126(3):278-286. doi:10.1161/CIRCULATIONAHA.111.071100.
218. Sung L-C, Chen C-I, Fang Y-A, et al. Influenza vaccination reduces hospitalization for acute coronary syndrome in elderly patients with chronic obstructive pulmonary disease: A population-based cohort study. *Vaccine*. 2014;32(30):3843-3849. doi:10.1016/j.vaccine.2014.04.064.
219. Remschmidt C, Wichmann O, Harder T. Frequency and impact of confounding by indication and healthy vaccinee bias in observational studies assessing influenza vaccine effectiveness: a systematic review. *BMC Infect Dis*. 2015;15:429. doi:10.1186/s12879-015-1154-y.
220. Nelson JC, Jackson ML, Weiss NS, Jackson LA. New strategies are needed to improve the accuracy of influenza vaccine effectiveness estimates among seniors. *J Clin Epidemiol*. 2009;62(7):687-694. doi:10.1016/j.jclinepi.2008.06.014.
221. Gail MH. Differential Error. *Wiley StatsRef Stat Ref Online*. 2014.
222. Rothman KJ, Greenland S, Lash TL. Validity in epidemiologic studies. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd editio. Philadelphia: Lippincott Williams & Wilkins; 2008:128-147.
223. Lanes S, Brown JS, Haynes K, Pollack MF, Walker AM. Identifying health outcomes in healthcare databases. *Pharmacoepidemiol Drug Saf*. 2015;24(10):1009-1016. doi:10.1002/pds.3856.
224. Rieck T, Feig M, An der Heiden M, Siedler A, Wichmann O. Assessing varicella vaccine effectiveness and its influencing factors using health insurance claims data, Germany, 2006 to 2015. *Euro Surveill*. 2017;22(17). doi:10.2807/1560-7917.es.2017.22.17.30521.
225. Giersiepen K, Pohlabeln H, Egidi G, Pigeot I. Die ICD-Kodierqualit??t f??r Diagnosen in der Ambulanten Versorgung. *Bundesgesundheitsblatt - Gesundheitsforsch - Gesundheitsschutz*. 2007;50(8):1028-1038. doi:10.1007/s00103-007-0297-4.
226. Gross G, Schöfer H, Wassilew S, et al. Herpes zoster guideline of the German Dermatology Society (DDG). *J Clin Virol*. 2003;26(3):277-289. doi:10.1016/S1386-6532(03)00005-2.
227. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol*. 2005;58(4):323-337. doi:10.1016/j.jclinepi.2004.10.012.
228. Robert-Koch-Institut (RKI). Impfquoten der Rotavirus-, Masern-, HPV-und Influenza-

Impfung in Deutschland. *Epidemiol Bull*. 2017;(1). https://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2017/Ausgaben/01_17.pdf?__blob=publicationFile.

229. Link-Gelles R, Westreich D, Aiello AE, et al. Bias with respect to socioeconomic status: A closer look at zip code matching in a pneumococcal vaccine effectiveness study. *SSM - Popul Heal*. 2016;2:587-594. doi:10.1016/j.ssmph.2016.08.005.
230. Wang E, Clymer J, Davis-Hayes C, Buttenheim A. Nonmedical exemptions from school immunization requirements: A systematic review. *Am J Public Health*. 2014;104(11):e62-e84. doi:10.2105/AJPH.2014.302190.
231. Stürmer T, Wyss R, Glynn RJ, Brookhart MA. Propensity scores for confounder adjustment when assessing the effects of medical interventions using nonexperimental study designs. *J Intern Med*. 2014;275(6):570-580. doi:10.1111/joim.12197.
232. Bruhn CAW, Hetterich S, Schuck-Paim C, et al. Estimating the population-level impact of vaccines using synthetic controls. *Proc Natl Acad Sci U S A*. 2017;114(7):1524-1529. doi:10.1073/pnas.1612833114.
233. Dupin N. Genital warts. *Clin Dermatol*. 2004;22(6):481-486. doi:10.1016/j.clindermatol.2004.07.003.
234. Simonsen L, Taylor RJ, Schuck-Paim C, Lustig R, Haber M, Klugman KP. Effect of 13-valent pneumococcal conjugate vaccine on admissions to hospital 2 years after its introduction in the USA: A time series analysis. *Lancet Respir Med*. 2014;2(5):387-394. doi:10.1016/S2213-2600(14)70032-3.
235. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology*. 2010;21(3):383-388. doi:10.1097/EDE.0b013e3181d61eeb.
236. Hawkins NG, Sanson-Fisher RW, Shakeshaft A, D'Este C, Green LW. The Multiple Baseline Design for Evaluating Population-Based Research. *Am J Prev Med*. 2007;33(2):162-168. doi:10.1016/j.amepre.2007.03.020.
237. Colizza V, Barrat A, Barthélemy M, Vespignani A. Predictability and epidemic pathways in global outbreaks of infectious diseases: the SARS case study. *BMC Med*. 2007;5(1):34. doi:10.1186/1741-7015-5-34.
238. Zumla A, Rao M, Wallis RS, et al. Host-directed therapies for infectious diseases: current status, recent progress, and future prospects. *Lancet Infect Dis*. 2016;16(4):e47-e63. doi:10.1016/S1473-3099(16)00078-5.
239. Carroll-Pankhurst C, Engels EA, Strickler HD, Goedert JJ, Wagner J, Mortimer EAJ. Thirty-five year mortality following receipt of SV40- contaminated polio vaccine during the neonatal period. *Br J Cancer*. 2001;85(9):1295-1297. doi:10.1054/bjoc.2001.2065.
240. D'Agostino RB, Pencina MJ, Massaro JM, Coady S. Cardiovascular disease risk

assessment: Insights from Framingham. *Glob Heart*. 2013;8(1):11-23. doi:10.1016/j.gheart.2013.01.001.

241. Ross MK, Wei W, Ohno-Machado L. “Big data” and the electronic health record. *Yearb Med Inform*. 2014;9:97-104. doi:10.15265/IY-2014-0003.
242. Abramson E, Kaushal R, Vest J. Improving immunization data management: An editorial on the potential of electronic health records. *Expert Rev Vaccines*. 2014;13(2):189-191. doi:10.1586/14760584.2014.870038.

APPENDIX A: Declaration

Hiermit versichere ich an Eides statt, dass ich diese Arbeit ohne unerlaubte fremde Hilfe angefertigt habe, keine anderen als die von mir angegebenen Quellen oder Hilfsmittel benutzt habe und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.
Außerdem erkläre ich, dass ich keine weiteren Promotionsversuche unternommen habe.

Kathrin Thöne Hamburg, 26. März 2018